



Progressives, Patterns, Pedagogy

*A corpus-driven approach
to English progressive forms,
functions, contexts and didactics*

Ute Römer

Studies in Corpus Linguistics ∞

JOHN BENJAMINS PUBLISHING COMPANY

Progressives, Patterns, Pedagogy

Studies in Corpus Linguistics

SCL focuses on the application of corpus method throughout language study, the development of a computational approach to linguistics, and the development of new tools for processing language.

General Editor

Elena Tognini-Bonelli

Consulting Editor

Wolfgang Teubert

Advisory Board

Michael Barlow

University of Auckland

Robert de Beaugrande

Federal University of Minas Gerais

Douglas Biber

North Arizona University

Chris Butler

University of Wales, Swansea

Sylviane Granger

University of Louvain

M. A. K. Halliday

University of Sydney

Susan Hunston

University of Birmingham

Stig Johansson

Oslo University

Graeme Kennedy

Victoria University of Wellington

Geoffrey Leech

University of Lancaster

Anna Mauranen

University of Tampere

John Sinclair

University of Birmingham

Piet van Sterkenburg

Institute for Dutch Lexicology, Leiden

Michael Stubbs

University of Trier

Jan Svartvik

University of Lund

H-Z. Yang

Jiao Tong University, Shanghai

Volume 18

Progressives, Patterns, Pedagogy: A corpus-driven approach
to English progressive forms, functions, contexts and didactics
by Ute Römer

Progressives, Patterns, Pedagogy

A corpus-driven approach
to English progressive forms, functions,
contexts and didactics

Ute Römer

University of Hanover

John Benjamins Publishing Company

Amsterdam/Philadelphia



™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Cover illustration from original painting *Random Order*
by Lorenzo Pezzatini, Florence, 1996.

Library of Congress Cataloging-in-Publication Data

Ute Römer

Progressives, Patterns, Pedagogy : A corpus-driven approach to English progressive forms, functions, contexts and didactics / Ute Römer.
p. cm. (Studies in Corpus Linguistics, ISSN 1388-0373 ; v. 18)
Includes bibliographical references and indexes.

1. English language--Tense. 2. English language--Verb phrase. 3. English language--Discourse analysis. 4. English language--Great Britain--Discourse analysis. 5. English language--Study and teaching--German speakers. I. Title. II. Series.

PE1301.R66 2005

425'.62--dc22

2005050253

ISBN 90 272 2289 4 (Hb; alk. paper)

© 2005 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

To Wolf-Dietrich Bald (1942–2004)

Table of contents

Acknowledgements	XIII
 CHAPTER 1	
Introduction: A need to take stock of progressives	1
1.1 Scope and aims of the study	1
1.2 Method of analysis	3
1.3 Structure of the book	4
 CHAPTER 2	
The theoretical basis of the study: Corpora, contexts, didactics	7
2.1 Corpus-driven linguistics (CDL)	7
2.1.1 CDL – a new theory emerging from corpus work	7
2.1.2 Corpus-based vs. corpus-driven approaches	8
2.2 Contextual approaches to the study of language	11
2.2.1 John R. Firth	12
2.2.2 John McH. Sinclair	13
2.3 Pedagogic and didactic grammar	14
2.3.1 Definitions	14
2.3.2 The Mindtian approach – empirical grammars	16
2.3.3 The present approach	17
 CHAPTER 3	
Progressives in theoretical studies and grammars of English	19
3.1 Problems of definition and terminology	19
3.2 The diachronic dimension: Progressives on the rise	21
3.3 The English progressive in two influential theoretical studies	22
3.3.1 Comrie 1976	22
3.3.2 Williams 2002	24
3.4 The progressive in recent linguistic and empirical grammars	25
3.4.1 Quirk, Greenbaum, Leech, Svartvik 1985	25
3.4.2 Biber, Johansson, Leech, Conrad, Finegan 1999	27
3.4.3 Mindt 2000	28
3.4.4 Huddleston, Pullum 2002	29

3.5	Previous empirical findings on the use of the progressive	31
3.5.1	Frequencies	32
3.5.2	Functions	33
3.5.3	Contexts	34

CHAPTER 4

Progressives in spoken British English 37

4.1	Corpus selection	37
4.1.1	Why spoken British native-speaker English (BrNSE)?	38
4.1.2	Availability of spoken British native-speaker English corpora	40
4.1.3	Corpus size and representativeness	41
4.1.4	Corpora used in this study	42
4.2	The empirical method: BNC and BoE data collection, processing, and evaluation	47
4.2.1	Verbs under analysis	47
4.2.2	The collection of corpus data: Query strategies	48
4.2.3	Data filtering	54
4.2.4	Data processing and encoding: The construction of an <i>Access</i> database	56
4.2.5	Data evaluation	58
4.3	The use of progressives in spoken English (I) – contexts	60
4.3.1	Distribution of different tense forms	61
4.3.2	Tense form contractions	66
4.3.3	Progressives and subjects	68
4.3.4	Progressives and objects	70
4.3.5	Progressives and prepositions	71
4.3.6	Progressives and negation	72
4.3.7	Progressives and other lexical-grammatical phenomena	74
4.3.8	Adverbial specification	75
4.3.9	Summary of the findings [spoken English – contexts]	79
4.4	The use of progressives in spoken English (II) – functions	80
4.4.1	Time reference	81
4.4.2	Two central function features: Continuousness and repeatedness	86
4.4.3	One central function or several central functions?	90
4.4.4	Central functions and time reference	93
4.4.5	Additional functions of the progressive	95
4.4.6	Additional functions and time reference	106
4.4.7	Summary of the findings [spoken English – functions]	110
4.5	Verbs and progressives – How lexical is grammar?	111
4.5.1	Distribution and restrictions: 100 verbs and 9,468 concordance lines [BNC/BoE]	113
4.5.2	Verbs and tense form distributions [BNC/BoE]	117

4.5.3	Verbs and subjects [BNC/BoE]	123
4.5.4	Verbs and objects [BNC/BoE]	130
4.5.5	Verbs and prepositions [BNC/BoE]	136
4.5.6	Verbs and negation [BNC/BoE]	136
4.5.7	Verbs and other lexical-grammatical phenomena [BNC/BoE]	143
4.5.8	Verbs and adverbial specification [BNC/BoE]	146
4.5.9	Verbs and time reference [BNC/BoE]	151
4.5.10	Verbs and central functions of the progressive [BNC/BoE]	157
4.5.11	Verbs and additional functions of the progressive [BNC/BoE]	161
4.5.12	Summary of the findings [spoken English – verbs and progressives]	169

CHAPTER 5

Progressive teaching (?): Progressives in the German EFL classroom	171
5.1	Learning problem “progressive” 172
5.2	Selection of teaching materials 174
5.3	The German English as a Foreign Language Textbook Corpus (GEFL TC) – a collection of EFL textbook language 175
5.3.1	Corpus design and composition 175
5.3.2	Corpus compilation 177
5.4	The empirical method: GEFL TC data collection, processing, and evaluation 178
5.4.1	Verbs under analysis 178
5.4.2	Data collection: Querying GEFL TC with <i>WordSmith Tools</i> 179
5.4.3	Data filtering 180
5.4.4	Data processing and encoding: The addition of GEFL TC concordance lines to the Access database 180
5.4.5	Data evaluation 180
5.5	The use of progressives in “school” English (I) – contexts 181
5.5.1	Distribution of different tense forms 182
5.5.2	Tense form contractions 184
5.5.3	Progressives and subjects 187
5.5.4	Progressives and objects 188
5.5.5	Progressives and prepositions 189
5.5.6	Progressives and negation 190
5.5.7	Progressives and other lexical-grammatical phenomena 191
5.5.8	Adverbial specification 192
5.5.9	Summary of the findings [GEFL TC – contexts] 195
5.6	The use of progressives in “school” English (II) – functions 196
5.6.1	Time reference 196
5.6.2	Two central function features: Continuousness and repeatedness 199

5.6.3	Central functions	201
5.6.4	Additional functions of the progressive	203
5.6.5	Summary of the findings [GEFL TC – functions]	205
5.7	Verbs and progressives in GEFL TC – How lexical is EFL textbook grammar?	205
5.7.1	The distribution of progressive verb forms in GEFL TC	206
5.7.2	Verbs and tense form distributions [GEFL TC]	207
5.7.3	Verbs and subjects [GEFL TC]	210
5.7.4	Verbs and objects [GEFL TC]	212
5.7.5	Verbs and prepositions [GEFL TC]	214
5.7.6	Verbs and negation [GEFL TC]	216
5.7.7	Verbs and other lexical-grammatical phenomena [GEFL TC]	218
5.7.8	Verbs and adverbial specification [GEFL TC]	218
5.7.9	Verbs and time reference [GEFL TC]	222
5.7.10	Verbs and central functions of the progressive [GEFL TC]	225
5.7.11	Verbs and additional functions of the progressive [GEFL TC]	227
5.7.12	Summary of the findings [GEFL TC – verbs and progressives]	229
5.8	Progressive progression (?) – When and how are progressives introduced in the textbooks?	230
5.8.1	The progression in Learning English Green Line New (volumes 1–6)	231
5.8.2	The progression in English G 2000 A (volumes 1–6)	234
5.8.3	Summary and discussion	237

CHAPTER 6

Progressives in real spoken English and in “school” English: A comparison		243
6.1	Progressives and context phenomena	244
6.1.1	Distribution of different tense forms	244
6.1.2	Progressives and subjects	246
6.1.3	Progressives and objects	248
6.1.4	Progressives and prepositions	249
6.1.5	Progressives and negation	250
6.1.6	Progressives and other lexical-grammatical phenomena	252
6.1.7	Progressives and adverbial specification	253
6.2	Progressives and function phenomena	256
6.2.1	Progressives and time reference	257
6.2.2	Progressives and central functions	260
6.2.3	Progressives and additional functions	266
6.3	Summary of the findings: Progressives in use vs. progressives in the books	268

CHAPTER 7

Pedagogical implications: True facts, textbooks, teaching	275
7.1 Corpus-driven linguistics and language teaching	276
7.2 The case for more authenticity in the classroom	277
7.3 Improving communicative competence: Teach the typical	280
7.4 Teaching progressives in natural contexts	282
7.5 Focussing on frequent functions of progressives	283
7.6 Shifting emphasis to lexis	285
7.7 Towards a corpus-driven communicative didactic lexical grammar of progressives	287

CHAPTER 8

Conclusions: Corpus, practice, theory	293
8.1 From corpus to practice	294
8.2 From corpus to theory	297
 Notes	 299
References	309
Index	325

Acknowledgements

I am very much indebted to my academic teacher and supervisor Wolf-Dietrich Bald, who was always supportive of my work, enthusiastic about my ideas, and a constant source of motivation. I would have wished he could have seen this study completed. To him the book is dedicated. Many thanks are also due to Rainer Schulze and Bernhard Kettemann for their advice, support, and encouragement, especially when things got tough.

I am grateful to the University of Cologne, the German Academic Exchange Service (DAAD), the European Commission, and the German Research Foundation (DFG) for providing me with scholarships and travel grants, and to HarperCollins publishers and The University of Birmingham for permission to access The Bank of English. Special thanks must go to David Oakey for countless helpful comments and suggestions on the whole manuscript and to John Sinclair for his invaluable feedback on parts of the book.

I would also like to thank the conference delegates at IVACS 2002 in Limerick, TaLC 2002 in Bertinoro, ICAME 2003 in Guernsey, ICAME 2004 in Verona, and TaLC 2004 in Granada for their helpful comments, stimulating questions, and supportive feedback after my presentations. The same applies to the participants at the Tuscan Word Centre courses in April 2000 and October 2001. Particular thanks go to the organisers of these expert training courses, Elena Tognini-Bonelli and John Sinclair, and to the 'guru' of the April 2000 course, Stig Johansson, for helping me find my way and for keeping me on track. It is a privilege to be introduced to corpus linguistics by some of the most original and open-minded thinkers in the field.

The following colleagues and friends deserve a big thankyou for their encouragement and support, for their interest in my work, and for countless "keep going" emails: Michael Barlow, Linda Bawcom, Ylva Berglund, Pernilla Danielsson, Anna Mauranen, Nadja Nesselhauf, David Oakey, Hanna Pishwa, Norbert Schlüter, Colin Wilcockson, and Chris Williams. On a similar note, my thanks go to a number of friends and colleagues in the English Department at the University of Cologne, not only from the linguistics but also from the literature side.

Last, but certainly not least, I wish to thank my parents Anita and Peter, my sister Anja, and my partner Karsten for their constant support, their endless patience, and for always being there.

CHAPTER 1

Introduction

A need to take stock of progressives

1.1 Scope and aims of the study

The English progressive, i.e. the combination of a form of the verb *TO BE* and the present participle (*-ing form*) of a verb (e.g. *is changing*, *were working*), has been variously described as an area of terminological inconsistencies, definitional and descriptive problems, and semantic or functional confusion (cf. Bybee, Perkins, & Pagliuca 1994; Declerck 1991; Hataş 1993; Landman 1992; Ljung 1980; Nehls 1974; Palmer 1988; Recktenwald 1975; Scheffer 1975; Williams 2002). Numerous theoretical studies have dealt with forms and functions of the “continuous”, “expanded”, “durative”, or “periphrastic” form, as the progressive is sometimes called, and large parts of any available grammar book are dedicated to a detailed description of progressive aspect. These accounts may differ considerably from each other, and there is no real consensus among scholars about the way in which progressive forms and functions are best dealt with. As Williams (2002: 13) notes, issues related to the use of progressive and non-progressive forms have been raised in a large number of publications, none of which, however, provides “an exhaustive or wholly accurate account of what it is that makes a speaker or writer choose” one form or the other.

What most of the existing studies on the progressive lack, I argue, is a broad empirical basis. More often than not, theoretical statements are derived on the basis of a limited collection of example sentences which are not attested instances of language use but rather intuitive examples which have been invented for the purpose of illustrating a particular point, or which have been taken from earlier linguistic studies and grammar books. Quantitative data on the distribution of different progressive forms or functions, or on co-occurrences with context features such as temporal adverbials, are also often missing from these studies. I think that a systematic empirical analysis, based on large amounts of language data, would be a sounder basis to theorise from. Also, it could perhaps help to sort out some of the still existing descriptive problems and contribute to clearing up part of the confusion on the function side.

One of the aims of the present study is hence to give a detailed synchronic empirical account of progressive verb forms in contemporary spoken British English. The study will address the following rather basic questions: How are progressives actually used in spoken English? How are their different forms distributed? In which lexical or syntactical contexts do they usually occur? What do different progressive forms typi-

cally express? and Is it possible to identify a “generally valid” behaviour of progressives, or do different verbs show largely dissimilar context and function patterns? To tackle these questions, I analyse large amounts of language data – data which occurred in natural communicative situations and which were then stored on the computer, forming what is well-known under the term “corpus”, i.e. an electronic collection of bits and pieces of spoken and/or written language used for linguistic analysis. A detailed empirical analysis of 10,171 progressive verb forms in context will form the basis for the development of a new and challenging approach to English progressives.

In addition to my linguistic curiosity, my interest in the topic is of a pedagogical nature. It has been noted by grammarians and applied linguistic researchers that in a foreign language teaching context, progressives, like other aspect-related phenomena, count among the grammatical “troublemakers” for learners and teachers of English. Williams (2002: 18) notes that the progressive “constitutes one of the most basic and ubiquitous problems facing language teachers”, and several other researchers highlight the difficulty that learners usually have in handling this grammar feature (cf. e.g. Hahn et al. 2000; Mindt 1997a; Nehls 1988; or Westergren Axelsson & Hahn 2001). Especially learners of language backgrounds in which the progressive does not exist as a grammatical category (e.g. German, Norwegian, Polish, or Swedish) have constant problems with its appropriate use (cf. Johansson & Stavestrand 1987; Lenko-Szymanska 2004).

In this context I consider it important to find out whether learners have problems with the appropriate use of progressives because progressives as such are more difficult to use than other language items or because their use is inadequately described in their coursebooks and reference grammars. In other words, the question is “What came first, the language problem or the faulty description?” To answer this question and to find out more about the presentation of progressives in teaching materials, a systematic account of the language of coursebooks will be the next step in my analysis. The question that I am going to address is “How is the progressive presented in a German EFL teaching context?” It is hoped that the study may thus contribute to a clarification of some of the problems connected with teaching and learning how to use progressive forms.

The next task is then to find out whether there are any significant differences between the use of progressives in real English and in so-called “school” English. This means that the results of a corpus linguistic analysis of functions and contexts of the progressive in natural English, as used in actual communicative situations, will be compared with the results of an investigation of the same features in English as it is used in the foreign language classroom. The present study can thus be seen as an example of the indirect use of corpora in language pedagogy, the merits of which have been described in a number of works by some of the leading researchers in the field of corpus linguistics (cf. Aston 2000; Aston, Bernardini, & Stewart 2004; Barlow 1996, 2003; Biber, Conrad, & Reppen 1994, 1998; Hunston 2002a; Kettemann & Marko 2002; Partington 1998; Sinclair 2004a; Tognini-Bonelli 2001). Some of the most important strengths of corpora in a pedagogical context are that (i) they cover the actual language

use of thousands of expert speakers, (ii) they give objective evidence since individual speaker preferences are equalled out in a large collection of text, (iii) they provide us with masses of natural language examples, and (iv) they highlight what is common and typical in the language.

Using the potential offered by corpora and corpus-analytic means, I will put myself in a mediator position and try to at least partially bridge the still existing gap between corpus linguistic research and language teaching. As a follow-up to the question about differences between authentic English and classroom English, it will be asked whether changes of teaching materials, and if so what kind of changes, ought to be made in order to achieve a higher degree of naturalness or authenticity in language teaching. The aim will be to develop a new concept of teaching the English progressive – a concept which takes empirical findings into account.

As mentioned above, the study will deal with the verbal construction *form of TO BE + present participle*, which is here referred to as “the progressive” of a verb. Progressives can in theory be formed of every English verb and of every tense form of the verb.¹ From all the theoretically possible forms, this study only deals with the PRESENT PROGRESSIVE (e.g. *are feeling*), the PAST PROGRESSIVE (e.g. *were feeling*), the PRESENT PERFECT PROGRESSIVE (e.g. *have been feeling*), and the PAST PERFECT PROGRESSIVE (e.g. *had been feeling*). Incomplete or fragmentary instances of these four tense forms (e.g. ... *how you feeling?*) have also been included and labelled accordingly in the database. Progressive infinitives (e.g. *be feeling*), modal progressives (e.g. *will be feeling*), and all passives (e.g. *has been being felt*) have been excluded from the analysis. These constructions are rather infrequent and thus there is insufficient data from which to make confident statements beneficial to learners.

1.2 Method of analysis

What do you imagine that it means? I have no data yet. It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts.

(*The Adventures of Sherlock Holmes*, Conan Doyle 1892:7)

As mentioned above, the central aims of this study are to carefully examine the ways in which the progressive is used in spoken British English and to relate the findings of this examination to a pedagogical context. The approach taken in order to achieve these aims can be described as *corpus-driven* (for a detailed description of this approach see Section 2.1), meaning that the analysis starts from corpus evidence and may lead to new insights into the language. Instead of first formulating a theory and then illustrating its applicability to occurrences of language use, I will put the evidence before the system and observe the data before I will derive the theoretical findings.²

When I studied chemistry, I learnt that the experiment usually comes first, even if all kinds of speculations on the object of research may precede it. You have an idea

about an interesting phenomenon worth investigating, carry out an experiment, make observations, describe carefully what you see, and then try to find a model which supports your findings, sometimes realising that the existing models are not entirely useful as they do not fully mirror reality. My analytic method is comparable to the procedure I have just described. I do not base my research on any particular pre-existing model or theory of the progressive. This means that many of the functional categories found in existing non-corpus-driven theoretical studies on the topic are not unreflectedly accepted as they may not account for everything that is likely to arise from the data. New categories had to be found and existing ones had to be newly defined. I believe that if empirical analyses are based on pre-existing categories and theories, corpus linguistics cannot be much more than a methodology (cf. Leech's 1992: 105 definition) and thus cannot lead to the formulation of a new theory about the language, but instead only to confirmations or slight alterations of already established frameworks. What we need to do, I argue (following researchers such as Sinclair, Teubert, and Tognini-Bonelli), is reject any preconceptions and existing theories when we start our analyses, since we may otherwise "run the risk [...] of distorting the data by creating a self-fulfilling corroboration of underlying assumptions." (Teubert 1999: 7)

Critical readers (and/or natural scientists) might argue in this context that no kind of observation can ever be theory-neutral (cf. Young 1971; as quoted in Stubbs 2000: 31: "all facts are theory-laden") and that starting an analysis from a particular verb form already implies a certain theoretical basis. They are right. If you do not have at least a rough idea of what you are looking at, you probably will not find much. I will therefore stick to many of the traditionally accepted grammatical terms like "verb", "preposition", "subject", "question", etc., categories that have proven useful in language analysis, even if they are not entirely theory-neutral; few chemists would ever seriously question the usefulness of the periodic table. I have tried, however, to keep the amount of theory-reliance to a minimum in order not to hinder the process of discovering new progressive-related phenomena so far not covered in the literature on the topic.³ In Sinclair's (1994: 25) terms, I have tried to "only apply loose and flexible frameworks" and to trust the data. Hence I knew what I was looking *at* (i.e. progressive forms of English verbs) and to a certain extent also what I was looking *for* (i.e. typical functions and contexts of progressives), but I did not know what I was going to find in the data (i.e. which different functions the progressive can have and which contexts it usually appears in). The method described here may appear radical (esp. when compared with more traditional ways of linguistic analysis), but it helps to avoid circularity and can lead to new findings about the language (cf. Tognini-Bonelli 2001: 11).

1.3 Structure of the book

The present book consists of an introduction, six main chapters, and a conclusion. Following this introductory chapter, Chapter 2 specifies the theoretical basis of the study, giving accounts of what is referred to as corpus-driven linguistics (CDL), as opposed

to corpus-based linguistics (CBL), introducing the reader to Firthian and Sinclairian contextual approaches to language analysis, and describing the concepts of pedagogic and didactic grammar.

Chapter 3 discusses how the topic “progressive aspect” is treated in the theoretical literature and in some recent important grammars of English. Included here is a short description of existing definitional and terminological problems in this field of linguistic analysis. A brief look at the historical development of the English progressive will serve to support the importance of aspect-related research.

Chapter 4 leads us to the first part of the empirical analysis which is at the centre of the study. It will be investigated how progressive verb forms of the 100 most frequent verbs are used in spoken British English. After a description of corpus selection criteria, the chapter will deal with methodological aspects, such as the collection, processing, encoding, and evaluation of corpus data, before illustrating in some detail which contexts and functions of progressives could be found in the chosen corpora, i.e. in the 10 million word spoken part of the British National Corpus (BNC) and in the 20 million word British spoken English (brspok) subsection of The Bank of English (BoE) respectively. To find out whether it is justified to talk about a “lexical grammar”⁴ of a language instead of sticking to the traditional, in a pedagogical context still highly influential separation of language into grammar and lexis, the question will be asked “How lexical is grammar?” What I want to find out is whether there is such a thing as a *purely grammatical progressive* or whether it would be more sensible to talk about a *lexical-grammatical progressive*. Starting from the analysed lexical items, i.e. starting separately from each of the 100 selected verb forms, I search for relations between individual verbs and certain syntactic and semantic phenomena, such as typical collocations, colligations, or combinations of verb forms with certain progressive meaning features.

In Chapter 5 the attention is turned to the treatment of the progressive in foreign language teaching, exemplified by EFL teaching on German secondary school level. The empirical analysis focussed on here runs parallel to the one described in the previous chapter. While Chapter 4 deals with BNC and BoE data, the corpus used in this second part of the empirical investigation is GEFL TC (the German English as a Foreign Language Textbook Corpus), a small electronic collection of texts taken from two best-selling EFL textbook series widely used in German secondary schools. In addition to the discussion of contexts, functions, and lexical-grammatical features of progressives, parts of Chapter 5 will also consider the grammatical progression in the textbooks selected and deal with some differences between the two textbook series *Green Line New* and *English G 2000 A*.

Chapter 6 compares the results of Chapters 4 and 5 and shows whether, and if so in what ways, the presentation of the progressive in German EFL teaching materials differs from the use of progressive verb forms in spoken British English.

Chapter 7 directly responds to the similarities and differences found between “school” English and real English and deals with a new concept of teaching the English progressive – a concept which goes back to the findings of this corpus-driven study.

Its seven subchapters serve to discuss a few pedagogical implications of the present research, some of which might present a challenge to applied linguists, textbook writers, and EFL teachers alike. Finally, the chapter provides a first draft of a new concept of teaching progressives.

Chapter 8 concludes the book, attempts to summarise major findings, and stresses the importance of further comparative and corpus-driven research on lexical-grammatical phenomena, and especially on such phenomena which cause constant problems in language learning and teaching. The chapter looks at relations between corpus work and corpus findings on the one hand, and practical applications (in our case in language pedagogy) and theoretical linguistic considerations on the other, and gives an outlook on related future research.

CHAPTER 2

The theoretical basis of the study

Corpora, contexts, didactics

2.1 Corpus-driven linguistics (CDL)

As featured in the title of this book and as mentioned above, the approach taken here is a *corpus-driven* one. In Section 1.2 it has been said that the investigation is highly committed to the data it starts from and that it tries to derive observational and theoretical findings from there, always trying not to lose contact with the corpora. This is – in a nutshell – what corpus-driven linguistics (henceforth CDL) is about. The following two paragraphs give an account of some central features of CDL, describe differences between this approach and other corpus-based approaches, and show what it is that makes CDL so powerful and thus in a way superior to other ways of analysing language.

2.1.1 CDL – a new theory emerging from corpus work

This headline may come as a surprise to the reader, especially as corpus linguistics is usually referred to as a methodology or as one of the possible data-gathering options a linguist can choose from when she or he needs evidence (alongside with informant asking or relying on her or his intuition). Articles and introductory textbooks on corpus linguistics tend to tell us that what they describe is neither a separate branch of linguistics, like morphology or syntax, nor one of the “hyphenated branches” of linguistics, like socio-linguistics or text-linguistics (cf. Leech 1992:105). Kennedy (1998:7), for instance, considers it

misleading [...] to suggest that corpus linguistics is a theory of language in competition with other theories of language such as transformational grammar, or even more that it is a new or separate branch of linguistics.

How then, the reader might ask, can CDL as a discipline within corpus linguistics be more than a methodology, a domain of study, a theory even?

It is certainly true that, as Leech (1992:105) notes, “corpus linguistics combines easily with other branches of linguistics”.⁵ In doing so, however, corpus linguistics not only adds a statistical dimension to our object of study, but often also confronts us with rather surprising findings which existing frameworks fail to account for (cf. e.g. Sinclair 1994:25). As Mindt was able to show in connection with his studies on future

expressions in English (see e.g. Mindt 1985, 1987, 1991, and 1992), corpus linguistic research often involves a necessity “to redefine linguistic classes, regroup cases or re-classify items.” (Mindt 1991:194) Corpus analysis may in fact, if taken seriously, make it necessary to redefine or regroup traditional linguistic concepts. The work done and the results obtained by researchers involved in the COBUILD project at the University of Birmingham is a case in point here.⁶ These researchers’s corpus-driven work has shown that the traditional division of the language system into grammar and lexis is rather inadequate in the light of natural data. Lexis and grammatical structures are closely linked and depend on each other. As observed by Hunston and Francis (2000:251),

if we note that *want* is often followed by a to-infinitive clause, it is not possible to decide whether this is a fact about lexis (the collocation of *want* and *to*) or a fact about grammar (the distribution of to-infinitive clauses). (*italics in original*)

In this context the authors discovered that each of the language patterns they analysed in their pattern grammar approach occurred “with a restricted set of lexical items, and each lexical item occur[red] with a restricted set of patterns.” (Hunston & Francis 2000:3; see also Hunston 2002b). This makes it very questionable whether we still ought to follow the tradition and distinguish grammar and lexis. Insights provided by research in the corpus-driven tradition prove that corpus linguistics can under certain circumstances in fact be more than a new empirical way of investigating a language and that corpus linguistic analysis can lead to a new theory of that language. It can be argued that the new developing theory is the logical result of using a new method on a new type of data.

As there usually is an interrelation between object, method and theory in any field of study which ought to be critically reflected by any researcher (cf. e.g. Bald 1995:104), changes on the object and method side are likely to result in changes on the theoretical side too (cf. Hunston & Francis 2000:250; Stubbs 1996:232). According to Tognini-Bonelli (1996a:54) “[a] corpus can be used in different ways in order to validate, exemplify or build up a language theory”. The following section will serve to show under which circumstances such a new theory build-up is likely to occur.

2.1.2 Corpus-based vs. corpus-driven approaches

The corpus-based and the corpus-driven approach can be regarded as two different opposing disciplines within corpus linguistics.⁷ It is assumed by the pioneers in corpus-driven linguistics that their relatively new approach⁸ (in contrast to corpus-based approaches) is capable of leading to a new theoretical framework of the language. What it actually is that makes CDL more theory-prone than CBL becomes obvious when we look at the ways in which corpus-based and corpus-driven linguists respond to the following three questions which are central to corpus research:

1. What is the status of the data and how and when (i.e. at which stage in the research) is the corpus approached?

2. Does corpus annotational material, i.e. any kind of information which can be added to the plain text (e.g. part-of-speech tags), rather have positive or negative effects on the research? and
3. Do we as researchers have to allow alterations of the system and should we be prepared to change existing theories in the light of corpus evidence?

It will now in turn be explained how these questions are dealt with in CBL and CDL.

The term “corpus-based” is often used by corpus researchers as a general or cover term to refer to any kind of corpus-informed or corpus-inspired research, i.e. any studies using corpus data in one way or another. In what follows here, “corpus-based” will be used in a more restricted sense. Following Tognini-Bonelli’s fundamental work in this area, I will use the term to distinguish corpus-driven approaches from non-corpus-driven approaches. It may well be, however, that a few linguists, who label their work “corpus-based” in fact use a corpus-driven approach to their object of study.

As Jan Aarts (2000:26) states, a corpus is nowadays often used “as a testing ground for hypotheses”. This statement provides an insight into the role of corpora in CBL. Corpus-based linguists do not put the corpus at the centre of their research but see it as a welcome tool which provides them with frequency data, attested illustrative examples, or with answers to questions of grammaticality or acceptability. Such linguists may be regarded as “instrumentalists” as they use corpora as instruments alongside other research strategies and other types of data. According to Tognini-Bonelli (1996a:65), “the corpus-based approach refers to a type of methodology where the commitment to the data as a whole is not ultimately very strict or systematic.” Researchers go to a corpus with pre-formulated ideas and fixed categories in mind in order to prove a certain hypothesis or to “exemplify existing theories” usually “theories that were formulated before large corpora became available” (Tognini-Bonelli 1996a:55).⁹ It is therefore very unlikely that an alteration of a theory is suggested by linguists whose work is corpus-based, no matter how significantly the data observed differ from their pre-conceived ideas (for strategies of accommodating theory and data in such a case see Tognini-Bonelli 1996b, Section 2.4 and 2001, Ch. 4).

Corpus-based linguists usually favour corpus-annotation. As they want to empirically test a theory, many researchers have their data annotated according to this theory to then easily get a quantification of their categories. As the research carried out by Geoffrey Leech and his colleagues at Lancaster University has shown, annotated corpora can indeed be very useful tools, for instance if one aims to automatically distinguish items which have the same form but which belong to different word classes, e.g. instances of the verb form “meeting” and of the noun “meeting”. McEnery and Wilson (2001:32) thus argue that “the utility of the corpus is considerably increased by the provision of annotation.” The commitment to annotated data can, however, also be considered dangerous, since the annotation of text means an abstraction of the data to certain categories (e.g. word classes). These categories seem to be more important than the actual data and the actual meaning of a lexical item may be obscured in this an-

notational process. Corpus-based linguists are thus further away from their data than corpus-driven linguists.

Corpus-driven linguists are aware of the possible negative effects corpus annotation can have on their research. It is important for them to try not to let other researchers's ideas and existing theories guide or influence them too much in approaching the data. For instance, in the BNC "it" is always tagged as <w PNP>it (i.e. as "personal pronoun"), regardless of whether it is a pronoun which refers cataphorically to a noun, as in *it's just a little article about recycling in Egypt*, or a dummy subject, as in *it's obvious that the people in Harlow like that type of thing*. So even though the corpus is tagged, it is not possible to search for "it" as dummy subject. On the other hand, the programmer of a corpus tagger may have decided to use several different tags for an item that the corpus user would always classify in only one way. In such a context annotation can indeed be seen as problematic: "With annotated corpora you are using other people's views of the language." (Michael Barlow, personal communication) Annotation does not necessarily mean an enrichment of the data. It may thus be considered a safer option to keep the text as clean as possible (cf. Sinclair 1991:9). Instead of relying on annotation, linguists who take a corpus-driven approach search by word-form (e.g. for the string "meeting") and then sort the data by superficial similarity (e.g. the left-hand context of the word) (cf. Hunston & Francis 2000: 19). Sinclair (2000a:37) describes this procedure as follows:

In corpus-driven linguistics you do not use pre-tagged text, but you process the raw text *directly* and then the patterns of this uncontaminated text are able to be observed. (my emphasis)

This quote also hints at the centrality of the text in CDL, i.e. the data stored on the computer. CDL means taking corpus evidence seriously. Findings are directly derived from the data; no filtering through existing concepts is supposed to take place. Tognini-Bonelli states that

[i]n a corpus-driven approach the commitment of the linguist is to the integrity of the data as a whole, and descriptions aim to be comprehensive with respect to corpus evidence. (1996a:69)

and

[...] the attempt of the corpus-driven linguist is to keep in constant touch with the evidence and to build up the theory step by step in the presence of massive corpus testimony (1996a:70).

This strong focus on attested instances of language use does not imply that CDL seeks to banish intuition (cf. Tognini-Bonelli 2001:91, 185). The process of data evaluation still requires the researcher's subjective judgements which are based on previous experience as a language learner and user. However, corpus data should not only supplement intuitive data but precede them. It follows from all this that working corpus-driven means that the evidence must never be ignored but has to be accepted and reflected in new theories (cf. Sinclair 1991:4). New theoretical statements are derived

from the data and have to be fully consistent with the corpus evidence. This proves that, given the situation that the data does not fit the theory, corpus-driven researchers take a different stance from corpus-based linguists.

Even though he did not use this label at the time of writing, John Sinclair in his 1991 seminal monograph *Corpus, Concordance, Collocation* presented a clear account of the corpus-driven approach, describing it as a “new view of language” (1991:1) and showed “how the corpus evidence can stimulate new linguistic hypotheses” (1991:9). In this context the work of Michael Barlow also fits in the corpus-driven framework as it describes the use of corpora “as a fundamental part of theory construction” (1996:2). To answer the third question formulated above, it can be said that corpus-driven linguists are certainly willing to alter existing theories or even work towards the development of a new language system.

I will now briefly respond to the question “What can CDL do that other corpus linguistic approaches cannot do?” Certainly, the major advantage of the corpus-driven approach is its ability to lead to new theoretical insights into language. As Hunston and Francis (2000:249) argue, the study of language phenomena in a large corpus of authentic texts “will lead to observations about the language that it has not been possible to make before.” We learn how language really works, how it is used in communicative situations, and, consequently, find out how it ought to be taught. I do not see any good reason why language learners (and their teachers) should not be presented with the actual facts about the language. Thus, if a large group of people is supposed to benefit from CDL, the insights corpus-driven approaches lead us to must not be kept secret among linguists but have to be shared with language practitioners.

It will, however, probably take some time before the investigations of CDL will have a direct impact on language teaching. Unfortunately, language teaching classrooms so far have not been very much affected (if at all) by the changes that are taking place in contemporary linguistic theory. Hopefully, by taking a corpus-driven approach to one selected language phenomenon, i.e. the English progressive, the present study will be able to contribute to affecting language teaching in a positive and forward-looking way.

2.2 Contextual approaches to the study of language

According to Aarts (2002:6), the corpus-driven approach as described in the previous chapter can “be looked upon as an operationalization of Firth’s ideas about the essential role of context in linguistic description” (cf. also Tognini-Bonelli 1996a:73). Hence an analysis that follows the principles of CDL by definition has to refer back to the work of J. R. Firth and other British contextualists who were inspired by Firth’s ideas. This chapter will therefore deal with a few central concepts of contextual approaches to linguistics as proposed and developed by John Firth and one of his followers, John Sinclair, respectively.

2.2.1 John R. Firth

John Rupert Firth (1890–1960), who was appointed the first chair of General Linguistics at an English university in 1944 (cf. Bazell et al. 1966: v, Honeybone Forthcoming, and Seuren 1998: 169), can be regarded as the “father” of the British contextualist tradition of language analysis (cf. Palmer 1998: 1062). A quote from his paper “A synopsis of linguistic theory, 1930–55” shows that Firth is clearly an empiricist: “Attested language text duly recorded is in the focus of attention for the linguist.” (Palmer ed. 1968: 199) The importance of “actual” language for Firth’s linguistic work also becomes evident in an earlier paper:

We must take our facts from speech sequences, verbally complete in themselves and operating in contexts of situation which are typical, recurrent, and repeatedly observable. (Firth 1957: 35)

Although roughly 70 years have passed since the formulation of this statement (in 1935), Firth’s focus on “typical, recurrent, and repeatedly observable” situations still sounds very modern to the corpus linguistic researcher today, who is constantly looking for language typicalities and repeated occurrences, not necessarily of contexts of situation in Firth’s sense, but of words and phrases in context.

Firth sees language as a whole, embedded in culture and society, which means that any kind of context is of major importance, be it a broader cultural or a much narrower lexical one. He suggests that the study of linguistics should centre around “what people say and what they hear and in what *context of situation* and experience they do these things” (Palmer ed. 1968: 171; my emphasis). This important phenomenon of “context of situation”, a notion Firth borrowed from the Polish anthropologist Malinowski (cf. Seuren 1998: 169), is often described as one of the key concepts in Firthian linguistic theory (see e.g. Palmer 1998: 1063). Referring back to Wittgenstein’s (1958: 20) famous quote “the meaning of a word is its use in the language”, Firth stresses the necessity of presenting words “in their commonest collocations” (Palmer ed. 1968: 179) and puts forward a suggestion which became his probably most frequently quoted statement: “you shall know a word by the company it keeps!” (Palmer ed. 1968: 179) Included in this quote is a definition of an important notion which was not, as people tend to think, coined by Firth (it was used before him, for example by Jespersen) but which became famous and widely known through his work, the notion of *collocation*, now central to almost any kind of corpus-driven research. Whenever we are interested in a word’s collocational behaviour, we analyse what company it keeps with other words, i.e. which other words repeatedly or habitually occur close to this word, usually “within a short space of each other in a text” (Sinclair 1991: 170). Barnbrook (1996: 87) very fittingly refers to collocation as the “sociology of words”. An example of a collocational relationship would be the association of “cows” with “milk” or “milking” (cf. Palmer ed. 1968: 180). Another typical example Firth mentions is the collocability of “night” and “dark” (cf. Firth 1957: 196). Collocations, as defined by Firth, “are quite

simply the mere word accompaniment, the other word-material in which they are most commonly or most characteristically embedded.” (Palmer ed. 1968: 180)

A second important notion used (and coined) by Firth is that of *colligation*. What collocation is on a lexical level of analysis, colligation is on a syntactic level. The term does not refer to the repeated combinations of concrete word forms but to the way in which word classes co-occur or keep habitual company in an utterance. An example would be the frequent colligation of personal pronouns in 2L position, i.e. two positions to the left, of a progressive verb form (e.g. *I think she ’s keeping you on.* BNC_spoken). Firth himself uses this term to deal with meaning on the level of syntax and defines colligation “as the interrelation of grammatical categories in syntactical structure” (Palmer ed. 1968: 183).

The concepts of collocation and colligation can also be found in the work of John Sinclair. In the next section it will be shown how Sinclair further developed Firthian concepts and how he applied them to corpus-driven analyses.

2.2.2 John McH. Sinclair

John McHardy Sinclair can be seen as one of the most innovative and visionary figures in contemporary English linguistics and certainly counts among the leading corpus linguistic researchers. His works clearly reflect a strong Firthian influence. For Sinclair *context* is all important in any kind of linguistic study. One of the “precepts” he formulates in an article on the impact of “[c]orpus evidence in language description” is “[i]nspect contexts” (1997: 34).¹⁰ For Sinclair (1987b: 320) “[i]t is clear that words do not occur at random in a text” (cf. also Sinclair 1991: 110), i.e. when they are combined in an utterance or in a text, words do not behave according to an open-choice principle or a slot-and-filler model. This observation led to the formulation of the *idiom principle*, a model of language interpretation which expresses that language is much more repetitive than most people think or perceive it to be:

The principle of idiom is that a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices.

(Sinclair 1987b: 320)

It implies that the study of language must include an examination of “the way in which patterns recur” (Sinclair 1966: 410). The idiom principle can be seen as an exemplification or illustration of Firth’s concepts of collocation and colligation. In other words, typical co-occurrences of words and word classes show the idiom principle in use and make the phenomenon of co-selection an important one for neo-Firthian linguists (cf. Stubbs 1993: 14).

Sinclair adds the notions of *semantic preference* and *semantic prosody* as further steps in contextual analysis to Firth’s concepts of collocation and colligation. A summary of the central concepts of his Firthian analytic approach to language is presented by Sinclair’s method of contextual analysis as described in Tognini-Bonelli 2001 (p. 19) and as taught by Sinclair himself on his intensive corpus linguistics courses at the

Tuscan Word Centre, Italy.¹¹ I would like to refer to Sinclair's analytical method as the *3C2S-method*, the five letters standing for the starting point and the four steps in the process of analysis:

Core → Collocation → Colligation → Semantic preference → Semantic prosody.¹²

In this analysis you start with a lexical item of your choice (either a single word or a multi-word item) which forms the core (or the node in your concordance). You then observe the core's collocational and colligational patterns, i.e. you examine the lexical and syntactical co-selection of items. In the next step you search for associations of relations of repeated words in the surrounding text to a certain semantic field and thus determine the semantic preference of the core, "where the only guaranteed co-selection is semantic, and where both the actual words and the word classes can vary." (Sinclair 1999:26) The final analytic step leads to insights into the meaning of the patterns and concerning the core's semantic prosody, "which gives some indication of the communicative intent, and has an attitudinal or pragmatic orientation." (Sinclair 1999:26) Sinclair "justifies" and exemplifies the four steps (CCSS) in a detailed and convincing examination of the expression "naked eye", providing a clear example of the corpus-driven approach (cf. Sinclair 1996:6ff.).¹³

The fact that the present analysis is embedded in the above-mentioned contextualist framework implies that I will try to take a holistic approach to the English progressive. Throughout the analysis, the BE + *V-ing* structure is not treated as an isolated syntactic phenomenon in its own right but always seen in the lexical-grammatical and syntactic context of a concordance line. I thus take an approach to progressives which differs from many previous accounts which try to explain the meaning of the verb form by looking at isolated (and often constructed) examples.

2.3 Pedagogic and didactic grammar

2.3.1 Definitions

As Leech (1988:5) rightly observes in his state-of-the-art article on English grammar and English grammars, "grammars come in many different varieties" and differ with regard to features like purpose and target audience. One possible distinction which can often be found in the literature on the topic is that between linguistic (or scientific or theoretical) grammars on the one hand and pedagogic (or didactic or applied) grammars on the other. A linguistic grammar, mainly used by the linguist and less frequently by language teachers or learners, is usually perceived to be rather comprehensive and complex and often features one specific language model (see e.g. Haegeman & Guéron 1999, who give an account of English grammar from a generative perspective), whereas a pedagogic grammar, mainly used by language learners and teachers, is generally less

comprehensive and more simplified and can combine aspects from more than one model of the language under investigation.

A lot of research on the relationship between linguistic grammar (LG) and pedagogic grammar (PG) has been done in Germany, especially in the 1970s and 1980s, by scholars like Bausch, Börner, Dirven, Hüllen, Jung, Klein, Mindt, Piepho, Ungerer, and Zimmermann,¹⁴ but the LG/PG differentiation goes as far back as to the late 19th century when Henry Sweet (1899: 138) stated that

there is an obvious distinction to be made between a grammar – whether for beginners or advanced students – which is to be assimilated completely so that the learner at last practically knows it by heart, and one which is only for reference,

and it was picked up again in the 1960s by Chomsky (1966: 10) for whom

[a] grammar, in the traditional view, is an account of competence. It describes and attempts to account for the ability of a speaker to understand an arbitrary sentence of his [or her] language and to produce an appropriate sentence on a given occasion. If it is a *pedagogic grammar*, it attempts to provide the student with this ability; if a *linguistic grammar*, it aims to discover and exhibit the mechanisms that make this achievement possible. (my emphasis)

It becomes obvious later in the book and in some of his other writings that Chomsky is not so much concerned with the pedagogical side but more interested in linguistic grammatical accounts and in the discovery of rule systems that languages are based on. Opposed to this view, the present study will deal with the often neglected pedagogical side of grammar.

If we assume that there is a relationship between LGs and PGs and that a PG is somehow based on an LG, one of the central questions that arise is “What is the best linguistic grammar to base a pedagogic grammar on?” This problem has been discussed by several linguists (see for instance Arndt 1969; Chalker 1994; Greenbaum 1988; Hudson 1992; Leech 1988; Mindt 1981a; Szałek 1982) with the overall outcome that there is no single most appropriate LG which can serve as a basis of PGs and that a mixture of different models or the development of an entirely new model will probably be the best option. Researchers agree that linguistic grammars should not be *directly* used in the classroom (cf. Mindt 1981a: 178; Szałek 1982: 43). There is, however, no consensus on the question how linguistic findings can be applied to language teaching or, put differently, how the information stored in LGs has to be processed before it can be included in a PG.

To pedagogicise or didacticise grammar Corder suggests a “predigestion” or partial digestion of the material found in LGs according to pedagogical guidelines:

In other words, the authors of such grammars have already organized the description or presented the data in a form which the teacher can use more or less directly in presentation to his [or her] own pupils. (Corder 1973: 328; see also 1988: 127)

Another way of reducing complexity and of making grammatical descriptions more suitable for the hand of the teacher (or learner) is the use of a pedagogic or didactic

filter. This filter model is described and promoted by Jung 1975, Bausch 1979, and Börner and Vogel 1976 but rejected by Mindt 1981a as the model cannot meet the requirements of a grammar for foreign language teaching (cf. Mindt 1981a:179) and as it “has no empirical basis in the form of an adequate corpus” (Dirven 1990:2).

In addition to “pedagogic grammar” some of the above mentioned German linguists use the concept “didactic grammar” (DG) to refer to grammars used in a language learning and language teaching environment. To my knowledge, the first person who used the term didactic grammar (or, to be more precise, the German equivalent “didaktische Grammatik”) was Werner Hüllen. According to Hüllen such a grammar has to be problem-oriented, semantically based, and relevant for language use (cf. 1973:5, 14). A didactic grammar is supposed to make foreign language teaching more effective and should facilitate the language learning process (cf. Hüllen 1976:52). Unfortunately, there are no clear agreed-upon definitions of the concepts PG and DG. One scholar’s didactic grammar may be another researcher’s pedagogic grammar. While Dirven (1990:1) treats “pedagogic grammar” as a general “cover term for any learner-oriented description or presentation of foreign language rule complexes” and includes various types of grammar in his definition, Mindt uses his DG as a general basis on which he can then develop specific pedagogic grammars for use in different learning contexts, at different school levels, and by different groups of pupils (cf. Mindt 1985:172). The next subchapter will deal with the approach Mindt takes.

2.3.2 The Mindtian approach – empirical grammars

Mindt takes a clearly empirical and thus innovative approach to pedagogic and didactic grammars. He stresses the importance of corpora in language analysis and description and bases his new model of a didactic grammar on corpora which have to be selected according to certain didactic criteria, e.g. the teaching aims (see Section 4.1 for further discussion of this point). In the Mindtian approach it is not a particular language model or linguistic grammar which comes first in the development of didactic and pedagogic grammars but a corpus. As Tognini-Bonelli (2001:91) states

Mindt and others systematically derive generalisations only from tested phenomena and refuse to superimpose on the data categories which have not been derived from, and substantiated by, the corpus evidence.

The first step in Mindt’s analysis is therefore a detailed corpus-driven description of the language that is learnt and taught. Of major importance is information on frequencies, i.e. on the relative distribution of language items and structures (cf. Mindt 1985:172).

According to Mindt, a didactic grammar for LT has to solve three major problems, the problems of (i) selection, (ii) progression, and (iii) presentation, on which there usually is not much information included in linguistic grammars (cf. Mindt 1981a:179).¹⁵ Frequency data from a corpus are needed to respond to problems (i) and (ii). They can help to make decisions about the selection of language items for teaching and about the sequencing of these selected items which determines the lin-

guistic progression in the course. To deal with the third problem (presentation), issues of co-occurrence and the co-selection of lexical-grammatical structures and contextual elements are important. These issues, for instance the co-occurrence of certain adverbials with certain tense forms and systematic empirical investigations of such phenomena, have been largely neglected in applied linguistic research (though see Crystal 1966 for an early study on this topic). A grammar which is descriptive, which includes information about frequencies and co-selection derived from corpus data, and which deals with selection, progression, and presentation criteria is called a didactic grammar (cf. Mindt 1981b:32) and presents a useful tool for teachers and authors of teaching materials. From this didactic grammar Mindt's procedure leads to several pedagogic grammars which can then directly be used by learners in the foreign language classroom. What such a customised PG looks like depends on factors like school type, course design, learner profile, or cultural context (cf. Mindt 1987:34, 1985:172).

Applications of Mindt's empirical grammar model can be found in his investigations of future expressions in English (cf. Mindt 1985, 1987, 1992, 1996) and modal verbs (cf. Mindt 1995, 1996). Also worth mentioning in this context are the works of some of Mindt's colleagues at Berlin, especially Haase, Schlüter, and Tesch, who all used a Mindtian empirical approach to compile didactic grammars of conditionals, the present perfect, and the indefinite pronouns *some* and *any*, respectively (cf. Haase 1995; Schlüter 2000, 2002a, 2002b; Tesch 1990). A more comprehensive didactic grammatical study is Mindt 2000 which will be dealt with later in some detail (see Section 3.4.3).

2.3.3 The present approach

The approach taken in the present study is strongly influenced by Mindt's empirical approach. Part of what I attempt to do in this analysis is work towards a corpus-driven communicative didactic lexical grammar of English progressives. The grammar is *corpus-driven* in Tognini-Bonelli's sense, in that it puts the corpus first, takes the evidence seriously, and derives findings in the presence of large amounts of naturally occurring language data. It is *communicative* in Leech's sense, putting an emphasis on spoken language (see Ch. 4) and on aspects of learning how to communicate successfully. Leech (1988:14) sees a communicative grammar as the "most appropriate kind of grammar to use as a basis for language teaching" and defines it as

a description of the grammar of a language in which grammatical forms or structures are related to the meanings which they express, and to the conditions under which they are used. It is a grammar of performance, as well as of competence. It may also be characterized as a study in which grammar is related to semantics, pragmatics, and style.¹⁶

Besides, it is *didactic* in Mindt's sense in that it focuses on language use, starts with an analysis of corpus data, and pays attention to the distribution of forms, functions, and contexts of progressives. The present analysis tries to respond to this definition by

looking at connections between language structures and their meanings or functions in real-life contexts.

CHAPTER 3

Progressives in theoretical studies and grammars of English

Even though I said above (see Sections 1.2 and 2.1) that working corpus-driven means working, as far as possible, unaffected by existing frameworks or theoretical descriptions, it makes sense to provide some information on previous publications on progressives in English and see how other linguists, who take different (non-corpus-driven) approaches, deal with the topic. The “newness” of a new approach can probably best be perceived against the background of what has been said about the progressive by some of the leading figures in tense and aspect related research before. Thus, in the following subsections, I will first discuss some general definitional and terminological problems researchers have to deal with (3.1) before I will have a closer look at the positions of a few individual scholars and give an account of the treatment of progressives in some landmark grammars of English (3.3 and 3.4). Also, the recent development of the progressive will be briefly described (3.2), and a summary of the results of previous empirical studies on progressives will be given (3.5).

As this would divert the reader’s attention from the central concern of the study, I will not provide an exhaustive overview of all existing studies on the English progressive or even give a historiographic account of the literature on the subject, but only account for some representative works. For further information on different theoretical approaches to the topic and overviews on relevant literature, the reader is referred to Scheffer 1975, Schulze 1985, and Williams 2002.

3.1 Problems of definition and terminology

It has already been noted in the introduction (Section 1.1) that different terms can be used to label the verbal construction here referred to as “the progressive”. A glance at the bibliography of any work on aspect in English (including the present volume) will suffice to support this observation. Titles including expressions like “expanded form”, “expanded tense”, “the continuous”, “temporary aspect”, “the periphrastic form”, “progressive aspect”, and “the progressive” basically all deal with the same construction: a form of TO BE + the present participle of a verb, e.g. *are looking*. This clearly hints at a confusing situation (particularly for students and novice linguists) which can be regarded as a terminological Babel.

Comrie, back in 1976, made a similar observation when he stated that “in discussions of aspect, as opposed to many other areas of linguistics, there is no generally accepted terminology” (1976: 11; cf. also Klein 1974: 76). In recent years, however, linguists more and more seem to favour the term “progressive” and only rarely refer to the “expanded” or “periphrastic” form. I therefore decided to choose this most frequently found label. I will mostly refer to “progressive forms” or “progressives” because my approach takes individual lexical items (-*ing* forms of high frequency English verbs) as the starting point of the analysis. Sporadically, mainly in connection with the review of other studies on the topic, reference will be made to “progressive aspect”. Like the majority of linguists and grammarians, I see the progressive rather as an aspectual phenomenon than as a tense. Utterances in the past progressive and past simple, for instance, do not necessarily express different time orientations but only put a different emphasis on things like continuousness or speaker attitude. It could be argued that my use of expressions like “progressives” or “progressive aspect” in the present study implies that I have implicitly accepted a pre-existing theory before approaching my data, and that this is therefore not a truly corpus-driven study. However, my decision to use the term “progressives” instead of a less theory-bound (but surely more clumsy) label, such as “verb forms ending in -ing, used in a particular context” was made for reasons of clarity. Apart from the borrowing of the term “progressives”, no other theoretical assumptions have been taken from previous studies. The corpus data have driven every part of the analysis.

As Tobin (1993: 4) states “[a]spect has almost as many definitions as there are linguists who have attempted to deal with it” (see also Klein 1995: 141; Bublitz 1995: 135). This quote leads us directly to the problem of defining the progressive. How do different linguists define progressive aspect differently and how do they account for its function or functions? Only a few selected, at times not entirely transparent, definitions can be presented here. For more extensive overviews and more detailed literature reviews the reader is referred to Hatcher 1951, Scheffer 1975, Ljung 1980, Sasse 2001, and Williams 2002.

For Ota “[a]spect means the signaling of the mode of action by some grammatical device.” (1963: 2; emphasis in original) Instead of in the “mode of action” Strang sees the crucial point in the indication of the “*manner* in which the “action” denoted by the verb is considered as being carried out” (1968: 143; my emphasis). Other linguists refer to aspect as the “character” of an event or the different “points of view” or “standpoints” of the speaker or writer. Dahl (1998: 64), for example, considers “the speaker’s perspective” as highly important but also the “structure of a situation” which is described differently with progressives and non-progressives. As for the main function of progressive aspect, Scheffer is probably right when he states that “[o]f all the different basic meanings attributed to the progressive that of duration is found most often” (1975: 21; emphasis in original). Imperfectivity and incompleteness also appear to be key terms when accounting for the nature of the progressive (cf. the discussion of the works of Comrie and Williams in Section 3.3). The fact that happenings in the progressive are “NOT NECESSARILY COMPLETE” is mentioned by Quirk et al. (1985: 188;

small caps in original; cf. Section 3.4.1; see also Quirk & Greenbaum 1990:53). In calling the progressive “the temporary aspect”, Joos (1964:107) focuses on another meaning feature which a number of researchers consider central: temporariness. Van Ek, for instance, describes “the basic meaning of the progressive as heightened temporary relevance” (1969:585; emphasis in original), centering on temporariness but also including subjective or emphatic function features.

In the present study I am not blindly subscribing to any of the definitions mentioned here but instead try to find out more about the nature of progressives by looking at their actual occurrence in discourse. It seems to me that the terminological and definitional problems briefly dealt with in this section call for a better, preferably empirical, analysis and clearer description of the topic.

3.2 The diachronic dimension: Progressives on the rise

A major reason why so many linguists see a necessity to deal with the progressive may lie in the historical development of this verbal construction. In this study I have thus decided to include a diachronic dimension. However, in my brief historical account I will not go back several centuries like a number of researchers before me did (see e.g. Jespersen 1931; Mossé 1938; Dietrich 1955; Elsness 1994; Görlach 1991; Nehls 1974; Recktenwald 1975; Scheffer 1975; Smitterberg 2005) but instead focus on the more recent changes related to frequencies and functions of the English progressive.

In the last few years a couple of linguists showed a growing interest in what Renouf (2002:29) refers to as “modern diachronic corpus linguistics”, a welcome trend within corpus linguistics which stresses the significance of investigating recent language developments on a morphological, lexical, and semantic level, and which thus contributes to a better understanding of language as a changing phenomenon. Of key importance in this context was the compilation of the FLOB and Frown corpora, the early 1990s counterparts of the LOB and Brown corpora of early 1960s English¹⁷ by a research team around Christian Mair at Freiburg University. The two newer 1 million word corpora offer the researcher a short-term diachronic perspective and together with the older Brown and LOB corpora enable comparisons of samples of written British and American English over a thirty-year time span.

While Scheffer (1975:110–113) comments on the increasing use of the progressive in Modern English, providing quantitative evidence from some best-selling 19th and 20th century novels, Mair and Hundt (1995:111) 20 years later support his findings in a paper based on data from the four above-mentioned matched corpora and state that “progressives have become more frequent over the past thirty years”.¹⁸ Further empirically-based confirmation of the increase phenomenon was given by Rydén in 1997, by Smitterberg et al. in 2000, and recently by Smith 2002. Extending the analysis of Mair and Hundt, Smith systematically used the LOB, FLOB, Brown, and Frown corpora to investigate ongoing changes related to the progressive and found increasing

frequencies of the present progressive in British and American English of about 30% (cf. Smith 2002:317).

As for a reason why progressives are on the rise, Potter (²1975:120) suspects that people have a growing “desire, often vaguely and subconsciously, to make what they say more lively and vivid”. More empirically founded are the two possible explanations that Mair and Hundt give: either the increase can be treated as a “textlinguistic or stylistic” phenomenon and “be regarded as a symptom of the “colloquialisation” of written English” or “[i]t is treated as a grammatical phenomenon. [...] in cases in which the simple form can be used alongside the progressive, the latter tends to be chosen with increasing frequency.” (Mair & Hundt 1995:118) This would mean that the expansion in frequency of use of progressives does not go together with an expansion of the functional repertoire. The described increase in use of the form implies an increasing importance of dealing with the progressive. Language description must not lose track of language development, especially not if we are concerned with pedagogical aspects. Learners and teachers deserve to get adequate, true, and up-to-date descriptions of the language, including information on recent trends in the use of central items or structures. It is therefore particularly important to further investigate such language phenomena which undergo a significant development.

3.3 The English progressive in two influential theoretical studies

Two important monographs that deal extensively with the progressive are Comrie 1976 and Williams 2002. While Comrie’s book is probably one of the most influential and most frequently cited works on aspect of the second half of the 20th century, Williams’s much more recent publication covers later developments and theories in aspectual research and offers a new approach to dealing systematically with the opposition between progressives and non-progressives. The central arguments of both books will be presented and evaluated in Sections 3.3.1 and 3.3.2. Let me stress again that the purpose of this study is not to list and discuss a large number of theories on the subject but to offer a fresh data-guided approach to progressive forms. The present chapter may hence not be as exhaustive as some readers might expect it to be.

3.3.1 Comrie 1976

Although published almost 30 years ago, Bernard Comrie’s book *Aspect* still presents an influential study in theoretical linguistics and traditional grammar. In his preface the author describes the book as an “introduction to verbal aspect and related concepts” in general which “is not concerned with any one particular language” (p. vii). A lot of information, however, is given on the progressive as a central aspectual category and large parts of the argumentation focus on English.

Comrie sees “aspects” as “different ways of viewing the internal temporal constituency of a situation” (p. 3; see also Comrie 1985 and 1998) and treats English constructions like *he was reading* as means of expressions of imperfective aspect, in this case “the periphrastic Progressive” (p. 9), with “imperfective” indicating “a situation in progress” (p. 19). The author states that

English has two aspectual oppositions that pervade the whole of the verbal system, that between Progressive (verb *to be* and verbal form of *-ing*) and non-Progressive, and that between Perfect (verb *to have* and Past Participle) and non-Perfect. (p. 124)

With reference to the “unusually wide range” (p. 33) of the English progressive and its extension of meaning (cf. p. 38) Comrie discusses different uses of the construction, for instance the according to several scholars (see also Jespersen 1931:524) “typical” framing context in which “the Progressive indicates a situation [...] that frames another situation”, as in “*when I visit John, he’ll be reciting his latest poems*” (p. 30), or more specific cases in which progressives refer “to a habitual situation that holds for a relatively limited period, as in “*we’re going to the opera a lot these days*” (p. 37, italics in original). The author also refers to the progressive function of adding “greater emotive effect” (p. 37) than the non-progressive form and gives a typical example in which the form co-occurs with the adverbial *always*: “*she’s always buying far more vegetables than they can possibly eat*” (p. 37). To Comrie these last two “specific uses” of the progressive appear to be idiosyncrasies which constitute deviations from its basic meaning of indicating a contingent situation (cf. p. 38).

Concerning the question which verbs go with the progressive and which do not, the author first refers to the traditional classification of verbs into the non-overlapping categories of statives (e.g. BE) and nonstatives (e.g. RUN) but then notes that “verbs are treated sometimes as stative, sometimes as nonstative, depending on the particular meaning they have in the given sentence.” (p. 36) He concludes that “the general rule seems to be that lexically stative verbs can be used nonstatively and appear in the Progressive” and provides a few examples of stative verbs in progressive use, e.g. “*Fred is being silly*” and “*I’m understanding more about quantum mechanics as each day goes by*” (p. 36).

On the whole it can be said that Comrie’s account is quite detailed and informative but in large parts rather speculative. His ideas are not supported by any massive evidence, the selected examples he provides are invented rather than authentic, and there is no information on frequencies of occurrence of the different special progressive functions Comrie mentions. It thus remains difficult to see *how* special (or idiosyncratic) functions like the emotive use of the progressive in fact are. The author does not manage to find a system which can give a fully comprehensive characterisation of progressive uses in (contemporary) English and which could account for all described “idiosyncrasies”. What Comrie provides the reader with, though, is an interesting discussion of the central concepts and features that are used in the non-empirical (and partially also in the empirical) literature on aspect in general and the progressive in

particular. By means of reviewing a more recent publication on the topic, the next subsection will try to show how some of these central concepts can be used as starting points for the adoption of a new, maybe more comprehensive, theoretical approach to English progressives.

3.3.2 Williams 2002

In his monograph entitled *Non-progressive and Progressive Aspect in English* Christopher Williams gives a detailed account of the differences between simple and progressive forms. Starting from the observation that “the theoretical issues raised so far by scholars do not provide an exhaustive or wholly accurate account” (p. 13) of the use of progressives in English, the author provides many new observations which lead him to formulating his own theory concerning progressive aspect: the theory of “susceptibility to change”.

Strongly influenced by Renaat Declerck’s publications on the subject, Williams gives a general overview of the state-of-play of progressive aspectual research while concentrating on the “more ‘difficult’ areas where linguists’ views often differ markedly” (p. 15), for instance the compatibility of progressives with apparently permanent situations or the progressive with future time reference. Williams’s interpretation of the progressive differs in several points from what linguists before him found (e.g. Hirtle 1967 or Scheffer 1975). For instance, he does not subscribe to the common view that non-progressive “always implies perfectivity” (p. 36). Also, imperfectivity for Williams is not restricted to the progressive but as well covers continuousness and habituality. As “intrinsic features of progressive aspect in English” the author lists (p. 50):

- i. that a situation be deemed as being ‘in progress’;
- ii. that a situation be deemed as being in some way incomplete;
- iii. that any situation conveyed by using the progressive form implicitly contains a ‘piece’ of that same situation prior to the moment in which it is viewed as occurring.

Williams then puts forward a new hypothesis of explaining the use of progressive (and non-progressive) aspect which says that “the use of the progressive form implies that the situation referred to in the verb may be susceptible to change in some way” (p. 87), i.e. it is easily affected by some internal or external factor; it “leaves “room for something to come”” (p. 88). He draws upon a number of convincing examples to show that his theory can also account for some of the trickier cases of progressive use, e.g. those which describe apparently permanent situations or those which refer to the future. However, only some of Williams’s examples are authentic “used” language examples (in Brazil’s 1995 terms);¹⁹ the rest of them have been invented or taken from previous theoretical studies. It may be questioned how appropriate it is to describe language on the basis of made-up instead of real examples. As an overall evaluation we can say that the author offers a more comprehensive and less speculative approach and manages

to explain more of the rather problematic progressive aspect related phenomena than other researchers before him (including Comrie). Still, it is doubtful whether what is postulated in his study can be considered the “final truth”, mainly because insufficient effort is made to cover larger amounts of natural data. It cannot be taken for granted that the examples given by the author provide a sound test bed for his theory of susceptibility to change. To make clear in how far the present approach differs from Williams’s (and other linguists’s) accounts, I will refer back to his book (and other studies) in Chapter 4.

3.4 The progressive in recent linguistic and empirical grammars

After I have taken a brief look at the presentation of progressive aspect in two important theoretical studies, I will now turn to four recent important grammar books. Sections 3.4.1 to 3.4.4 discuss the treatment of the progressive in grammars which represent different degrees of “theoreticalness” and “empiricalness”, with the cases “fully theoretical” and “fully empirical” seen as endpoints of an imaginary scale. I will focus mainly on findings related to different functions of progressives and on the grammar books’s commitment to natural (corpus) data. All grammars are “post-corpus-boom” grammars, i.e. grammars that were published after the first electronic corpora of English had become available to linguistic researchers.

I have selected Quirk et al. 1985 as the probably most comprehensive and most widely used reference grammar of the 20th century, Biber et al. 1999 as a first entirely corpus-based grammar of different English language registers, Mindt 2000 as the first corpus-driven empirical grammar of the English verb system, and Huddleston and Pullum 2002 as a most recent voluminous publication which claims to be “[t]he most up-to-date English grammar” and “[t]he grammar for the 21st century”.²⁰

3.4.1 Quirk, Greenbaum, Leech, Svartvik 1985

The volume *A Comprehensive Grammar of the English Language* (henceforth CGEL) by Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik, published in 1985, certainly presented *the* reference work on English grammar for researchers and teachers in the late 20th century and can still be regarded as an invaluable source of information about the English language and an indispensable reference work today. As a culmination of several decades of joint work of the four authors, who also produced *A Grammar of Contemporary English* (Quirk et al. 1972), *A Communicative Grammar of English* (Leech & Svartvik 1975, ²1994), and *A University Grammar of English* (Quirk & Greenbaum 1973, ²1980), CGEL attempts to give an authoritative and comprehensive account of English grammar which replaces previous grammars.

Within the main chapter on “The semantics of the verb phrase”, 16 pages are dedicated to “progressive aspect”. The authors define “aspect” as “a grammatical category

which reflects the way in which the verb action is *regarded* or *experienced* with respect to time.” (188, emphasis in original) Progressive aspect, they state, “indicates a happening IN PROGRESS at a given time” (188, small caps in original) and can be decomposed functionally into three meaning components,

not all of which need to be present in a given instance:

- (a) the happening has DURATION
- (b) the happening has LIMITED DURATION
- (c) the happening is NOT NECESSARILY COMPLETE (198, small caps in original).

Quirk et al. then go on to deal with the compatibility (or incompatibility) of different verb senses (states, events, and habits) with the progressive and claim that stative verbs like *OWN* are usually unacceptable in a progressive construction but that when they occur, “temporariness rather than permanence” is expressed, as e.g. in “[w]e *are living* in the country.” (199, emphasis in original) Progressive forms of event verbs such as *BLOW* in “[t]he referee *is blowing* his whistle” according to CGEL convey “the idea that an event has duration and has not yet come to an end” (199). The example “[t]he professor *is typing* his own letters while his secretary is ill” is used to illustrate the fact that the progressive can also impart habituality, but only if the habit referred to is temporary. In this context the authors mention the “subjective feeling of disapproval” (199) progressives may express in combination with adverbials of the *always* type (e.g. *always, constantly, continually*).

The CGEL approach to progressive aspect is characterised by a high degree of sensitivity to progressives and verb semantics. The authors repeatedly note that depending on the aspectual form (progressive or non-progressive) present in a certain sentence or utterance, different interpretations of the function of the verb are required. The “non-normal progressive” use of the stative verb *HOPE* in “I am hoping you will come”, for instance, adds the special effect of tentativeness or politeness to the utterance (202). To sum up what is said about progressive functions, CGEL clearly supports a “one major use plus several special uses” view. While progressives mainly express temporariness, they may also be used to refer to future actions or to convey tentativeness (210).

As for the level of empiricalness of Quirk et al., the grammar can probably be best described as corpus-informed. It includes occasional information on frequencies of grammatical structures based on the Survey of English Usage (SEU), Lancaster-Oslo/Bergen (LOB), and Brown corpora, but does not present any systematic detailed distributional data or statistics. Although the authors mention some corpora they have been working with, there is no clear attribution to sources. It remains unclear which examples (and how many of them) actually come from the SEU, Brown, or LOB and which have been constructed or taken from previous studies. Particularly in cases in which one grammatical construction is contrasted with another (e.g. “Joan *sang* well” vs. “Joan *was singing* well”, 197, emphasis in original), it becomes evident that the examples have been invented for illustrative purposes. On a scale of theoreticalness I would see Quirk et al. 1985 somewhere in the middle, as they neither follow any

particular existing theory (unlike Universal Grammar) nor are they fully committed to accounting for natural data. Where the 1999 *Longman Grammar of Spoken and Written English* can be found on the same scale and in how far it follows (or differs from) the CGEL, will be dealt with in the next subsection.

3.4.2 Biber, Johansson, Leech, Conrad, Finegan 1999

The 1999 *Longman Grammar of Spoken and Written English* (henceforth LGSWE) by Douglas Biber, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan, can certainly be regarded as a landmark in linguistic analysis and description. As the authors state in their preface, the book is not meant to compete with the above described large-scale CGEL but to complement it (cf. viii). Even though LGSWE “does not attempt to aspire” (viii) to the comprehensiveness of Quirk et al., it extends the scope of CGEL by being entirely corpus-based. In focussing on register-dependent linguistic choices and in putting an emphasis on spoken English (and its differences to written English), Biber and his colleagues go beyond CGEL. Their account is more functional and discourse-oriented and less structural than Quirk et al.’s.²¹

In LGSWE a five-page subsection of chapter six “Variation in the verb phrase: tense, aspect, voice, and modality” deals exclusively with progressive aspect in different spoken and written registers of British and American English, while reference to progressives is also made in a few other parts of the book, e.g. in connection with future time or relative clauses. Aspect, according to LGSWE, “relates to considerations such as the completion or lack of completion of events or states described by a verb” (460). The definition of progressive aspect closely follows that in CGEL which stresses the being *in progress*, “usually for a limited duration” (470), of the respective situation at some point in time. The authors distinguish between present progressive aspect, which functions to describe “events that are currently in progress or are about to take place in the near future” and past progressive aspect, used to refer to “events that were in progress or about to take place at some earlier time” (470), and omit perfect progressives because of their rarity in the Longman Spoken and Written English Corpus (LSWEC). Progressives, and particularly present progressives, they note, are most often found in conversation while written registers overall contain lower numbers of the form (cf. 462).

Under the heading “Lexical associations of progressive aspect” Biber et al. first list and then discuss different types of verbs that are more or less likely to occur in progressive constructions. In this respect the authors follow the example of CGEL where verb semantics (as noted above) plays an important role. LGSWE, however, goes further in providing rough frequencies (e.g. “more than ten times per million words”, 472) for the occurrence of particular verbs with the progressive aspect and in stressing that, contrary to what many researchers state, “both dynamic and stative verbs are included among the most common verbs in the progressive” (472). Also, LGSWE pays more attention to context and lexical-grammatical features than CGEL. As crucial factors that help determine the use of (stative and dynamic) verbs in the progressive, the authors

mention context features, e.g. human subjects as agents. They thus manage to account for the apparently contradictory use of stative verbs in progressive constructions as e.g. in “*I was looking at that one just now*” (473) where the human subject “I” controls the visual perception. On the other hand, a stative verb like *SEE* is only rarely used with progressives with animate subjects because it describes a state which cannot be actively controlled by the agent. On the whole, the discussion of progressives and their functions in LGSWE is less comprehensive in not accounting for as many special cases of progressive meanings but much more empirically founded than in CGEL.

As we could see, LGSWE is empirical and descriptive rather than theoretical and prescriptive. Ample evidence from natural used language stored in LSWEC is given throughout the book. On our imagined theoreticalness scale Biber et al. would hence be found further towards the empirical end than Quirk et al. However, as LGSWE largely follows CGEL in its terminology and descriptive framework and does not develop many new categories or descriptive labels to meet new types and larger amounts of data than were analysed by Quirk and his colleagues, the approach taken can not be called corpus-driven and still deviates in some points from the approach described in Section 2.1 of this study.

3.4.3 Mindt 2000

Dieter Mindt’s *An Empirical Grammar of the English Verb System* (henceforth EGEVS), published in 2000 as “the result of ten years’ work” claims to use “a new approach to English” and, unlike LGSWE, to make “no borrowing from previous grammars” (6). The two most important characteristics of EGEVS are (i) that the grammar is “fully corpus-based” and (ii) that it is “especially geared to the requirements of ELT” (Mindt 2002:91).

In his description of the English verb system, which this grammar is restricted to, Mindt uses an inductive approach “from language to grammatical generalization” rather than a deductive one “from pre-stated rule to example” (6). This means that, although the author uses the label “corpus-based”, according to Tognini-Bonelli’s definition and my outline in Section 2.1, the approach taken in EGEVS is, in fact, rather corpus-driven. The fact that Mindt’s grammar “makes use of a number of grammatical categories which evolved from the analysis of the data” (7) is certainly a typical feature of CDL. In that respect Mindt’s procedure clearly resembles the analytic method applied in the present study (cf. Section 1.2 above).

Progressives are dealt with in EGEVS in some detail (pages 248–265) within a chapter on “finite verb phrases”. Instead of giving some introductory definitional information on progressive aspect, the author starts straight away with a functional analysis of the form. According to Mindt, “[t]here are nine meanings of the progressive”, listed in EGEVS in order of frequency of occurrence (cf. 248; see also Mindt 1997a: 230–232).

1. incompletiion
2. temporariness

3. iteration/habit
4. highlighting/prominence
5. emotion
6. politeness/downtoning
7. prediction
8. volition/intention
9. matter-of-course

The author gives corpus examples for all nine meanings and then notes that the first three (incompletion, temporariness, iteration/habit) are progressive, the other six non-progressive meanings (cf. 250). All nine meanings “can occur in isolation with the exclusion of other meanings” (250) but quite often meanings are combined in a single progressive utterance. In this context it may be considered problematic that the examples given under meanings 1 to 9 are not always clear-cut cases but include different meaning features. One of the examples given to illustrate “iteration/habit”, for instance, “*she was crying and **jumping up and down***” (249, emphasis in original), does also express incompletion (still going on at the time of reference) and temporariness (lasting only for a limited period of time). EGEVS also provides information on verbs that are frequently used with progressives and lists a lot of details concerning the lexical-grammatical contexts in which progressive forms are typically found. These findings will be further discussed in Section 3.5 and, in comparison with my own empirical findings, throughout Chapter 4.

Concerning the grammar’s degree of theoreticalness, it has already been noted that EGEVS follows a corpus-driven approach of analysis and description and can therefore be found very close to the empirical endpoint of the theoreticalness scale. In not assuming any descriptive frameworks other than the one evolving from the corpus data, Mindt’s book is probably the least theoretical (and most empirical) of the grammars discussed so far. Unfortunately, it is not entirely clear which subsets from which corpora were used during which stage of Mindt’s research. On this subject, the author only states the following: “During the final stage of our work we had access to more than 240 million words of English” (6) and gives a list of corpora used at the end of the volume (596).²²

3.4.4 Huddleston, Pullum 2002

The fourth and last grammar under discussion is *The Cambridge Grammar of the English Language* (henceforth CamGEL to avoid confusion with Quirk et al. 1985) by Rodney Huddleston and Geoffrey Pullum, published in 2002. CamGEL claims to represent “a major advance over previous grammars by virtue of drawing systematically on the linguistic research carried out on English since the 1950s” and to be “based on a sounder and more consistent descriptive framework than previous large-scale grammars” (cover blurb). The grammar thus obviously competes with the 17 years older CGEL, although it acknowledges Quirk et al.’s book as an “indispensable source

of data and ideas” (xvi). What the authors of CamGEL state to do is “give a synchronic, descriptive grammar of general-purpose, present-day, international Standard English” (2).

The CamGEL approach to grammatical description is more specifically theory-focussed than that of CGEL and much less empirically oriented than Biber et al.’s and, even more so, Mindt’s (cf. 18–19). As for the data used in connection with writing this grammar, the authors mention a wide range of sources, mainly of an introspective kind or referring back to other linguists’s works, but also including three, by today’s standards, small corpora of American, Australian, and British English: the one million word Brown, ACE (Australian Corpus of English), and LOB corpora. Empiricists may find it doubtful whether three million words of written English from the 1960s (Brown and LOB) and the 1980s (ACE) can actually provide a sound basis for “the most up-to-date English grammar, reflecting decades of change” (CamGEL promotional description), even if the authors are not aiming at a statistical account of grammar including frequencies (like Biber et al.) but instead focus more on the theoretical level and only draw upon corpus evidence for illustrative purposes.²³ As also noted with respect to Quirk et al.’s volume, the sources of the individual examples in CamGEL are unclear and the reader is left guessing which examples come from which source.²⁴

One of the main chapters of the grammar entitled “The verb” includes a detailed subchapter on tense, aspect, and mood, which in turn has a 10-page subsection exclusively dedicated to progressive aspect. “The term **aspect**”, the authors write, “applies to a system where the basic meanings have to do with the internal temporal constituency of the situation” (117, emphasis in original). Echoing Comrie’s definition of perfective and imperfective aspect, they note that

[t]he difference [between progressive and non-progressive] is a matter of how the speaker views the situation. The progressive takes an internal view, looking at it from the inside, as it were, as something ongoing, in progress. The unmarked, non-progressive, version takes an external view (117).

In large parts the account of progressives closely follows the work of Comrie (cf. Section 2.3.1). According to CamGEL, the progressive has one basic use, namely “the expression of progressive aspectuality” (162), and two specialised uses, both related to future time reference (“The progressive futurate” and “*Will* + progressive”, 171). Progressive aspectuality requires certain meaning features: the situation described is presented as in progress, as durative, as dynamic, and as having limited duration. It is viewed imperfectively, and the reference time is “a mid-interval within” the situation time (163). These meaning features are then discussed and exemplified in turn. With regard to the duration feature Huddleston and Pullum, like Comrie, stress “the frequent use of the progressive for a situation presented as a frame or background for a perfective situation” (166). It remains to be seen whether “framing” really represents such a central function of progressives in real language use (see Section 4.4.5).

One longer subsection of the chapter on progressive aspect deals with the compatibility of certain verb classes “falling around the boundary between states and

occurrences" (169) with progressives, e.g. "verbs of perception and sensation" such as SEE, HEAR, or FEEL. Throughout the tense and aspect chapter the authors put an emphasis on different kinds of situations, e.g. processes vs. achievements or occurrences vs. states, the latter of which do not normally occur with progressives (cf. 119). With respect to verb-progressive compatibilities CamGEL makes a number of informally statistical claims expressing "how easily" some verbs take the form, whether they are "possible" with it, or what verbs tend "to favour the progressive" (170, 171). A real statistical analysis, however, has apparently not been carried out. The reader certainly gets a lot of information on different interpretations of progressive constructions but, unfortunately, no empirical confirmation for the claims made.

Coming back to the issues of theoreticalness and empiricalness, it should have become clear from the above discussion that, of the four grammars described, CamGEL is the one closest to the theoreticalness end of the scale. Like CGEL, the grammar is also merely corpus-informed, using corpora as example-providing instruments, but starts off with a more fixed theoretical framework than Quirk et al. and does not seem to permit alternative approaches or account for gradience in the way CGEL does. Grammar thus appears to be a very determinate system.

3.5 Previous empirical findings on the use of the progressive

At the turn of the millennium Jan Aarts (2000:34) critically remarked that

in spite of the fact that English is perhaps the most written-about language, most English grammars written in this century do not have a very firm foundation in empirical fact. (2000:34)

What Aarts observes with reference to grammars and language phenomena in general is also true for 20th century books and articles on one particular area: progressive aspect. Most studies on the subject are of a mainly theoretical nature and try to develop new frameworks to better explain the use of progressives on the basis of previous research and usually small amounts of mostly intuitive data. There are, however, a few exceptions. The following sections will deal with some selected empirical studies on the progressive and summarise their findings concerning overall frequencies of occurrence (3.5.1), functions (3.5.2), and contexts (3.5.3) of the form. I will refer back to these in later parts of the study to see whether they are comparable with the results of my own larger-scale corpus-driven analysis.

My "collection" of studies which include quantitative information on the use of progressives in English comprises (sorted by year of publication) Ota 1963, Joos 1964, Allen 1966, Scheffer 1975, Pürschel 1981, Legenhausen 1985, Levickij & Romanova 1997, Virtanen 1997, Biber et al. 1999, Mindt 2000, and Smith 2002. Not all authors make statements on all three aspects (frequencies, functions, contexts); some only deal with questions of general frequencies or adverbial modification. It should also be noted that, as for the empirical basis of the individual analyses, data collections of

very different types and sizes have been selected and used as corpora. The list of empirical studies includes a few somewhat dated publications from before the advent of electronic corpora. Although carried out in pre-computer times, these studies contain some interesting findings on the use of progressives and make some significant contributions to the field of aspectual research. With the exception of Mindt 2000 whose work is corpus-driven (see Section 3.4.3), all authors take corpus-based approaches to the topic, starting from certain hypotheses and with sets of categories which are then applied to the data. Joos can be considered as a borderline case between CBL and CDL as the author on the one hand “occasionally draw[s] upon transformational grammar” but on the other hand generally “tries to make contributions to general linguistic theory while describing a body of data in one professional way” (Joos 1964:5).

3.5.1 Frequencies

Information on the frequencies of progressives compared with other non-progressive forms can be found in five of the eleven studies mentioned above (Ota, Joos, Allen, Biber et al., and Smith).

Ota's work on tense and aspect in the early 1960s has been described as “an early example of corpus research which broke new ground” (Kennedy 1998: 122). As a basis of his investigations he used a small corpus²⁵ of (mainly) spoken and (less) written American English from 1950 to 1961, consisting of radio conversations, TV play scripts, and academic research articles and including 17,166 finite verb forms (cf. Ota 1963:21). Of these verb forms 6.1% are progressives. Joos found a lower progressive percentage in his analysis of verb forms in written British English. His corpus, however, consists of a single novel only, Sybille Bedford's *The trial of Dr. Adams*, first published in 1958, and contains 8,038 verb form tokens of which progressives make up 4.9% (cf. Joos 1964:76–79). Two years after Joos, Allen published his empirical study of verb forms in written American English from the 1920s to the 1960s. Of about 4,800 so-called “verb-clusters”, randomly selected from novels, non-fiction works, plays, and periodicals, 207 (or 4.3%) are progressives (cf. Allen 1966: 136). In their early quantitative approaches to the subject all three researchers found rather low relative frequencies of progressive forms as compared with non-progressives. A similar observation was made by Turner (cf. 1981: 195) who also discusses a number of studies from the 1960s.

Some newer empirical investigations, based on larger amounts of more recent corpus data, basically support these early findings. Biber et al. do not give exact figures but note that progressive aspect verb phrases are much less common than their simple counterparts and “slightly less common than perfect aspect verb phrases” which show relative frequencies between 5% and 10%, depending on the register (1999:461).²⁶ Finally, Smith in an article on recent changes in the use of progressives in written British English gives a slightly different type of frequency information, namely the ratio of present progressives to simple present forms, and arrives at values of 3.3% for the LOB corpus and 3.9% for FLOB (cf. Smith 2002:320). His counts support previous findings on the relative rarity of progressive forms, particularly in written text. As I did

not set out to investigate the entire English verb system and look at all possible verb forms (as Ota or Joos did for instance), the present study will not include any numerical information comparable to that reported on in this section but instead focus on frequency information related to the contexts and functions of progressives in the selected corpora. In Section 4.3.1, I will however present frequency findings on the different progressive forms, i.e. present progressive, past progressive, present perfect progressive, and past perfect progressive. For the time being, it will have to suffice to know that, concerning their overall frequencies and according to the findings of a number of scholars, progressives are rare in all varieties and registers of English but comparatively more frequent in spoken data, and that they are, in Mair and Hundt's (1995) terms, "becoming more frequent" on the whole.

3.5.2 Functions

Seven of the eleven selected studies include empirical information on the functions of the progressive (Ota, Scheffer, Pürschel, Legenhausen, Virtanen, Mindt, Smith). However, only three of them give actual frequency numbers for the different meanings²⁷ progressives can express.²⁸ The other authors mainly refer to meaning tendencies, present rather tentative results, and do not attempt to actually interpret every single progressive form and to quantify the different cases. Ota, for instance, does not provide any explicit quantitative data on the distribution of progressive functions but just observes that

[i]n a majority of cases progressive forms indicate an action in process now or at some time in the past [...] but there seem to be cases when the present progressive form indicates "timeless" events. (1963:63)

Basing his findings on a corpus of seven hours of spoken British English material from BFBS radio programmes,²⁹ Pürschel talks about a preferred use of the present progressive in situations in which speakers know a lot about the topic under discussion and want to emphasise certain aspects (cf. 1981:89). He also found repeated instances of the temporal frame meaning in his data but considers the described emotional use of the progressive to be more important. Another rather tentative statement concerning the basic function of progressives based on an empirical investigation of a small corpus of British radio reports (football, tennis, etc.) can be found in an article by Legenhausen. The author states that progressives basically have a scene-setting function and are mainly used to provide a background for the main actions described (cf. 1985:128). He compares this important use with Weinrich's *Reliefgebung* (²1971) but suspects that it may be a particular characteristic of the special text type (reports on the radio). Virtanen, who analyses the use of progressives in British and American students' argumentative essays (in comparison with the writings of Swedish and Finnish learners), says that, in the light of her findings, it seems

that the main function of the progressive in the British data is to add duration to the situation coded in the sentence and/or to indicate the constantly growing nature of a process under attention. (1997:306)

A similar observation, this time with reference to explicit frequency data, is made by Scheffer in a study on the progressive in a 375,000 word corpus of six, at the time of writing, contemporary British novels. Scheffer (1975:60) claims that “[m]ore than 60% of the verbs in the progressive in the corpus express duration, sometimes mixed with some other element of meaning” but fails to mention how the rest of the progressives can be distributed across other functions. Smith in his corpus-based study also touches upon the issue of progressive functions and looks, among a few other verb-class-related phenomena, at expressions of future meaning. He identifies only very small numbers of present progressive examples with clear future reference in LOB and FLOB. Even when indeterminate cases in which future and present time reference merge are included in the counts, just 4.5% of the present progressives in LOB and 3.9% in FLOB refer to future events.

None of the studies described so far provides 100% satisfactory information on the distribution of progressive functions in authentic spoken British English. The only author who systematically accounts for the shares of different meanings of progressives is Mindt. His findings are, however, based on a collection of several corpora of spoken and written British and American English. There is no differentiation between the registers and regional varieties concerning meaning distributions. The percentages for the nine meanings with the progressive that Mindt (2000:256) isolated are as follows:³⁰ incompleteness c. 60%, temporariness c. 36%, iteration/habit c. 12%, highlighting/prominence c. 9%, prediction c. 7%, volition/intention c. 5%, other meanings (emotion, matter-of-course, politeness/downtoning) c. 5%. As noted above (see Section 3.4.3), it is not fully clear what data Mindt’s findings are based on. It could also be questioned whether his set of meanings is an ideal one to capture the functional spectrum of the progressive, given for instance that an example listed under the “iteration/habit” meaning also expressed incompleteness and temporariness (see my comments in Section 3.4.3). The analysis described in the present study (cf. Section 4.4) will try to provide a more detailed and systematic account of progressive functions based on a clearly defined large set of data from two massive corpora of spoken British English.

3.5.3 Contexts

With the exception of Joos, Allen, and Biber et al. the studies listed above (Section 3.5) contain empirically retrieved information on language items that tend to co-occur with the progressive. An important feature which most authors report on, though with different degrees of explicitness and detail, is the collocation of progressives and (mainly temporal) adverbials. Ota, who describes his work as “a study of the co-occurrences or collocations (as Firth calls them) of individual items with each verb

form” (1963:14), determines the share of temporal adverbial modification of all progressives in his predominantly spoken American English material. Of 1,047 phrases, 140 (i.e. 13.4%) contain a so-called “time-indicator” such as *now*, *at this moment*, *today*, or *these days* (cf. 1963:21–24).

Scheffer finds a much higher percentage of progressive specification with temporal adverbials. In his corpus of six novels 21.2% “of all the verbs in the progressive [...] are specified by temporal adverbials” (1975:55).³¹ Although this sounds as if the author is referring to types (i.e. different verbs used in the progressive), other passages in his book make it clear that he actually means tokens (i.e. occurrences of verb forms; cf. e.g. 1975:120). Following Crystal (1966), Scheffer distinguishes different groups of adverbials supplying answers to the questions “How often?”, “How long?”, and “When?”. Adverbials which give a time-reference (group 3, answering the question “When?”), e.g. *now*, *today*, *tomorrow*, or *at this moment*, are with 73.6% most frequently used in his data (cf. 1975:52). Smitterberg, Reich and Hahn confirm Scheffer’s findings. In an empirical analysis of progressives in 19th and 20th century academic and political debates they note that “[a]dverbs expressing time references are the most frequent type of adverbial modification.” (Smitterberg et al. 2000:114)

Other studies on adverbial collocations of progressives do not (like Scheffer) give any detailed statistical information but just make rather general statements. Pürschel simply writes that time adverbials (like *right now* or *at the moment*) very rarely co-occur with present progressives in his BFBS radio programmes and questions the signal status of these items for the choice of the right tense form (cf. 1981:89). Legenhausen found that in sport reports on the radio some adverbial expressions (e.g. *in the opening minutes*, *this afternoon*, *tonight*) call for the progressive and never collocate with a simple form (cf. 1985:130). According to him, the adverbials *now*, *at the moment*, and *just* also mainly occur in progressive verb phrases. The author does not determine the shares of adverbial modification of all instances of the progressive but instead calculates the percentages for different verb forms in relation to all occurrences of a particular adverbial. It is thus difficult to compare his results with the findings of other researchers.

Looking at “the combinability of adverbs and verbs” in a 1.5-million word corpus of written British and American English from different text types (poetry, prose, press), Levickij and Romanova also observe a statistically significant (chi-square tested) co-occurrence between progressives and adverbials of time (cf. 1997:136) but do not go into any detail concerning different types of time adverbials. Virtanen considers adverbials in the immediate context of progressive forms in her student essays and states

[r]epresentative instances in the British subcorpus include temporal adverbials such as *continuously*, *already*, *still*, *since*, *over the years*, *no longer*, or *today* and *at the moment*. Moreover, we also find the adverbial *increasingly* collocating with *become* and of course a host of lexical items appearing in the immediate context of this verb which are modified by *more* or *more and more*. (Virtanen 1997:306)

Unfortunately, as is the case with most of the other studies too, explicit numerical data on adverbial collocation is missing in Virtanen's account. Mindt does not specify the exact percentage of adverbial modification of progressives either but lists a number of adverbs which frequently co-occur with the progressive. In his mixed (spoken and written) data, the six items *still*, *now*, *also*, *already*, *just*, and *always* cover 65% of all cases of progressive-adverbial collocation (cf. Mindt 2000:265). In addition to his investigations on adverbials in the context of progressives, Mindt looks at other features such as their combinability with subjects. He first distinguishes between intentional and non-intentional subjects and then states that the progressive much more frequently combines with members of the former group (e.g. *she*, *I*, *people*, making up about 86% of all cases) than with the latter (e.g. *something*, *it*, *the earth*, c. 14%; cf. 2000:263).

Smith also looks at subjects but applies a different, more traditional system of classification to quantify instances of certain progressive-subject collocations. He subdivides all occurrences of present progressives in LOB and FLOB according to person of subject and finds highest figures for the 3rd person (71.4% LOB, 73.4% FLOB) and much lower numbers for 1st and 2nd person subjects (LOB: 17.3% 1st person, 11.4% 2nd person; FLOB: 18.1% 1st person, 8.5% 2nd person; cf. 2002:324).

Summing up the information given about progressives and their context phenomena in previous empirical studies, it is rather problematic to find a common core. The individual accounts differ remarkably with respect to what they actually analyse when they examine collocates of progressive forms and with respect to the type of data they use as their analytic basis. They hence provide very different and largely incomparable results. In Chapter 4.3 of this book I will attempt to give a systematic account of typical items that appear in the context of progressives in spoken British English. Features under investigation comprise subjects, objects, prepositions, adverbials, and others. Whenever appropriate, reference will be made to the findings of other scholars.

CHAPTER 4

Progressives in spoken British English

In the introduction I said that one of the aims of the present study is to give a detailed synchronic account on the use of progressive verb forms in contemporary English. As stated above (cf. Section 1.3), this chapter will explore in which contexts and functions progressives are found in two corpora of spoken British English. The chapter constitutes the first part of the empirical analysis which is at the heart of the study. The approach is different from most previous studies on the progressive which base their observations on limited amounts of, often invented, data. Sinclair (1991:100) has an important point when he says that “the language looks rather different when you look at a lot of it at once”. I will therefore look at large amounts of data in order to see what the language is really like.

Fortunately, there is nowadays not much doubt about the important role that corpora and corpus-analytic techniques play in linguistics. The advantages of corpus linguistic analysis are widely acknowledged by leading scholars from different linguistic areas. As Nelson (2000:225) rightly notes, “[t]here is a profound sense among corpus linguists these days that the use of corpora has “arrived””. However, this has not always been the case. The linguistic mainstream from the 1950s to the 1980s was largely rationalist rather than empiricist, and even today a number of linguists working in generative frameworks along the lines of Chomsky and his followers are still not very much in favour of corpus work and, to use Sinclair’s (1991:1) words, prefer “to look inwards to the mind rather than outwards to society”. Fortunately, today the mainstream of linguistics is moving more and more towards empiricism where outward looks are fostered.

The purpose of the present chapter is not to further elaborate on the importance of corpus analysis in linguistics but, taking the centrality of corpora for granted, to discuss an example of a large-scale corpus analysis of English progressives.

4.1 Corpus selection

The first step in any kind of corpus linguistic research project, no matter whether it is corpus-based or corpus-driven, has to be the selection of an appropriate corpus (or appropriate corpora) or, if no suitable sources are available, the compilation of a new corpus. Sinclair’s (1991:13) “[t]he results are only as good as the corpus” is a

key statement in this context and calls for the greatest possible care in corpus choice and creation.

Corpora are usually compiled for a particular purpose (e.g. dictionary making, studies in language acquisition, research on specialised languages), which means that whatever goes into the corpus has to meet certain selective criteria and that corpus size may vary considerably, depending on the compiler's aims. Researchers who want to base their studies on existing corpus resources should keep this in mind. The following sections deal with issues of corpus selection, such as corpus availability, size, and representativeness, and give an account of the corpora that have been used in the present analysis of progressives. The first question that has to be answered, though, is a question about the particular type of English this study is based on: spoken British native-speaker English.

4.1.1 Why spoken British native-speaker English (BrNSE)?

This headline combines the following three questions related to corpus selection:

1. Why spoken and not written English?
2. Why British English and not any other variety?
3. Why native-speaker and not non-native speaker English?

One argument for the selection of *spoken* material as an empirical basis of a study of progressive forms lies in their higher frequency of occurrence in speech than in writing (cf. Allen 1966:136; Biber et al. 1999:462; Quirk et al. 1985:198). As Williams (2004) found in an empirical study on legal discourse, the progressive form practically disappears in highly formal contexts. Smith (2002:318) also sees a particular need to investigate spoken corpus data “as this is the most likely locus of innovation in progressive use.” Besides, despite the quantitative dominance of speech over writing in everyday life, most existing grammars are still biased towards the written language – a fact which calls for more explorations of data from real-life conversations where “we shall find the key to a better understanding of what language really is, and how it works” (Sinclair & Coulthard 1975:3, quoting Firth 1935). A similar argumentation can be found in Sweet's writings in which he makes a strong claim for starting any linguistic analysis from spoken data:

all study of language, whether theoretical or practical, ought to be based on the spoken language. [...] We see, then, that the advantage as regards clearness and definiteness is on the side of the spoken language: by starting from the spoken language we have less to learn, and we learn it accurately. (Sweet 1899:50, 52)

The main reason for choosing spoken English as a descriptive basis, however, was the emphasis which language teaching curricula usually put on teaching spoken language. The so-called *Richtlinien* for ELT in North Rhine-Westphalia,³² i.e. the official guide-lines for teachers which include the teaching syllabus, stress the pre-eminence of speech over writing and call for teaching materials which help to improve the com-

municative competence of the pupils and their ability to use the foreign language in actual life situations. As the syllabus in the first phase of ELT centres on speech, it makes sense to focus on spoken corpus data as a target. This view is supported by researchers in applied linguistics who state that

language teaching which aims to foster speaking skills and natural spoken interaction should be based upon the grammar of spoken language, and not on grammars which mainly reflect written norms. (Carter & McCarthy 1995: 141)

The central argument for the selection of *British* English lies in the fact that ELT in Germany is mainly oriented towards the British standard (rather than AmE or other varieties; cf. Görlach 1999: 18). It is probably wise, at least at an introductory stage of language instruction, to choose one particular regional model and not to confront learners with a number of different Englishes straight away.³³ As Görlach (1999: 18) notes, “[i]t cannot be the aim to teach an active competence in more than one variety”. Later in the course, though, pupils ought to be trained in dealing with variation. They should learn that often the choice of a lexical item or construction depends on the selected genre or regional code.³⁴ Also, from a data access perspective, it has to be noted that the availability of spoken corpora of other regional varieties than British English is rather limited. The existing and publicly accessible corpora of spoken American English, for instance, are either very small or specialised.³⁵

Given the pedagogical background of the study, authentic *native-speaker* English has been chosen as the basis for comparison with “school” English. I consider spoken BrNSE a suitable norm (even though it may not be the only possible one) in a foreign language teaching context and treat it as an ideal kind of learner input (cf. also Römer 2004b). However, in the last few years, a number of linguists started to criticise the predominant concentration in research and teaching on English as a native language (ENL) and question the model status of native-speaker English in language pedagogy. I am here referring to researchers working in the field of English as a *lingua franca* (ELF, e.g. C. Gnutzmann, J. House, J. Jenkins, K. Knapp, A. Mauranen, B. Seidlhofer), a type of English which is used in interactions between speakers of different first languages (cf. Seidlhofer 2001a: 76–77; House 1999: 74). Mauranen (2004: 101), for example, notes

[a]s English is increasingly used as a lingua franca among non-native speakers, it is becoming more and more obvious that sensible targets and norms of usage cannot be those designed for a native speaker.

(see, e.g. Knapp & Meierkord 2002; Seidlhofer 2000)

Researchers in the areas of global or international English tend to regard lingua franca English as a superior type of English which is more collaborative than NS English.³⁶ They stress the importance “to realize that native-speaker language use is just *one* kind of reality” (Seidlhofer 2001b: 138, emphasis in original) and expect ELF as a new, more relaxed teaching model to be better suited to meet the growing diversity of Englishes in a global context. At present, however, research on lingua franca English is still at a rather early stage. At the time of data collection there were no ELF corpora available

that could have been used in the present study. Seidlhofer herself comments on the restricted access to alternative descriptions to native-speaker English but states that “this is going to change” (2002:208; see also Seidlhofer 2001b:133). Worth mentioning in this context are two spoken ELF corpora which are currently being compiled at the universities of Vienna (Austria) and Tampere (Finland): VOICE (the Vienna-Oxford International Corpus of English; project coordinator: Barbara Seidlhofer)³⁷ and ELFA (the English as a Lingua Franca in the Academia corpus; project coordinator: Anna Mauranen). Once these and other resources of (spoken) lingua franca English are available, it will certainly be worth carrying out an ELF-based, or rather ELF-driven, follow-up analysis to this study on progressives. Like Seidlhofer and others I see a danger in denying diversity and think that there is a need to extend corpus analysis to ELF research in order to complement native language investigations. I am not sure, however, whether we really need new and more relaxed non-native speaker target models. As recently pointed out by Mukherjee (2004), a non-native target norm may not be what learners want. For the time being, I will stick to the “old” target and regard (spoken) native-speaker English as the language teaching norm.

The choice of spoken BrNSE corpora of course implies that the results of the analysis will only be valid for one small fraction of language use. A grammar of English as a whole would have to include descriptive accounts of a variety of different registers (spoken, written, “in-between” media such as email or e-chat language), text types, and regional variants. It has to be kept in mind, though, that one of the main purposes of this study is a pedagogical one. The analysis is therefore restricted to spoken British native-speaker English corpus data.

4.1.2 Availability of spoken British native-speaker English corpora

Nowadays, a constantly growing number of corpora of various different types is made available to the linguistic research community. Although day after day people produce much larger amounts of spoken than written language, written corpora dominate the scene, mainly for practical and economical reasons. The compilation of a corpus of writing is in general a less difficult, less time-consuming, and less expensive undertaking than that of a speech corpus. Any transcription of speech recordings, especially if it includes the phonetic and prosodic levels, is much more costly in terms of time and money than the conversion of written material into electronic format by means of keying, scanning, or re-formatting.

The construction of the first 220,000 word corpus of spoken British English in the early 1960s has to be regarded as a particularly remarkable achievement, if we consider that “[t]ape-recorded monologue or dialogue typically contains 7,000–9,000 words per hour” and that “[i]t may take about 10 hours to transcribe orthographically one hour of recording” (Kennedy 1998:81). This first machine-readable speech collection was produced by a team of linguists around John Sinclair (first at Edinburgh and later at Birmingham) as part of the Birmingham Collection of English Text (cf. Sinclair 1995:99, 2001:viii).³⁸ The second computer-held corpus of spoken English ever

compiled is the London-Lund Corpus (LLC), an electronic version of the spoken component of the pre-electronic Survey of English Usage³⁹ (cf. Svartvik & Quirk 1980). Compiled in the late 1970s, the LLC consists of 500,000 words of prosodically transcribed BrNSE from various spoken genres, such as face-to-face conversations, sports commentaries, or telephone conversations (cf. Svartvik 1990).

The next much larger computerised collection of spoken data that was released in 1995 is the 10 million word spoken subsection of the British National Corpus (BNC, see Section 4.1.4). A larger set of transcribed spoken material (more than 20 million words) is only accessible as part of The Bank of English (BoE, see 4.1.4). All other spoken corpora of British English that have been made available for research are much smaller than these last two. There are, for instance, the very small Lancaster/IBM Spoken English Corpus of about 52,000 words (cf. Knowles et al. 1996) and the 600,000 word spoken part of ICE-GB, the recently published British component of the International Corpus of English (cf. Nelson et al. 2002). Unfortunately, there is only restricted access to the much larger CANCODE corpus, the Cambridge and Nottingham Corpus of Discourse in English, which comprises about five million words of spoken BrNSE. Worth mentioning are also two other large commercial corpora, or rather groups of corpora, which include reasonable amounts of spoken material but are not available for academic institutions: the Cambridge International Corpus and the Longman Corpus Network.⁴⁰

4.1.3 Corpus size and representativeness

One of the central issues in corpus linguistics, particularly in corpus design and selection, is the question of corpus size. Closely linked to the size problem are questions of corpus representativeness. A large corpus can generally be regarded more representative of the type of language it consists of than a small corpus which contains the same kind of language. Of course, any small corpus is better than no corpus at all, but if the choice is between a small and a large corpus of the same (or similar) kind of material, I would always go for the latter.

Although the size of a corpus can be measured and given in absolute numbers, corpus size has to be seen as a relative phenomenon, at least from a diachronic perspective. A corpus of one million words (such as the Brown corpus of written American English) was considered “large” when it was compiled in the 1960s, and it was probably “a big thing” for the computers of the time to handle that much data. Today, with constantly developing electronic resources and computers which can process billions of words without a problem, and with existing (mainly written) corpora of several hundred million words in size (e.g. The Bank of English, the Cambridge International Corpus, or the Reuters Corpus⁴¹), a one-million word corpus appears rather tiny. Sinclair’s “we associate with a corpus the attribute of it being large” (1995: 103) is even more valid today than when he said it. Some corpora, however, do not need to be as large as others in order to prove useful linguistic resources:

This notion of size is obviously conditioned by various factors. So, for example, if the corpus is of a very specialized kind of language, then it would not necessarily have to be as big as it would be if it was of a very general type of language.

(Sinclair 1995: 103–104)

The mini-corpus of German EFL textbook texts that was compiled for this study (see Chapter 5), for example, is such a specialised corpus. It contains only a bit more than 100,000 words but still reveals a lot about the particular type of language it represents: the language of German EFL textbooks.

Corpus representativeness is certainly a controversial topic, as its treatment in the literature and repeated discussions on the CORPORA mailing list show.⁴² An important question that has been raised refers to absolute representativeness of corpora: Can there be a fully representative corpus? or, put differently, Can there be a corpus which leads to absolutely generalisable findings? To be able to answer these questions, we first have to know what the corpus in question is supposed to be representative *of*. It is not particularly difficult to compile a corpus that is representative of Shakespeare's poetry – all one has to do is collect all poems ever written by Shakespeare and put them on the computer – but if our aim is the construction of a representative corpus of spoken BrNSE, things are less straightforward. It is logistically impossible to include all spoken material that is being produced day after day by native speakers of British English all over the world in a representative corpus. Researchers thus need to apply certain sampling strategies to select some pieces of spoken data for their corpora in order to obtain highest possible representativeness.⁴³ As Kennedy (1998:66) rightly states,

any corpus, however big, can never be more than a minuscule sample of all the speech or writing produced or received by all of the users of a major language on even a single day.

Thus, the answer to the question “Can there be a fully representative corpus of spoken BrNSE?” must be negative. However, the larger the corpus and the cleverer the sampling of included spoken data are, the higher the degree of representativeness and the more generalisable the findings. Also, if two independent randomly collected data-sets from a corpus provide comparable results, this indicates that the corpus is representative of the text type it contains. Another useful strategy in this context is to test the results obtained from one corpus against a second corpus of the same type. This issue is further discussed in the next section.

4.1.4 Corpora used in this study

The main criteria that determined the corpus selection for this study were corpus size and representativeness. As the aim was to arrive at more generalisable and more reliable results than previous empirical studies on progressives, I decided to work with the largest and most representative spoken BrNSE corpora available at the time of data collection. A major argument for the selection of the largest available corpora also lies in the corpus-driven approach taken in this study. According to Sinclair (2000a:38),

“[c]orpus-driven linguistics demands extremely large corpora because of its need for multiple occurrences of all the items it handles”. By using large corpora it was ensured that a representative amount of progressive verb form tokens of different types could be retrieved and patterns could emerge from the data. These considerations led to the selection of the two largest accessible corpora of spoken BrNSE: the spoken component of the British National Corpus (henceforth BNC_spoken) and the spoken British subsection of The Bank of English (henceforth BoE_brspok).

Readers might wonder why the analysis was not simply based on one large corpus (either BNC_spoken or BoE_brspok). Each of the two corpora alone could certainly provide the researcher with enough relevant data on progressives. The motivation behind choosing two similar corpora was, again, the achievement of highest possible generalisability of the findings. Back in 1951, before the advent of computers and electronic corpora, Harris (1951: 13) remarked that

the analysis of a particular corpus becomes of interest only if it is virtually identical with the analysis which would be obtained in like manner from any other sufficiently large corpus of material taken in the same dialect. If it is, we can predict the relations among elements in any other corpus of the language on the basis of the relations found in our analyzed corpus. When this is the case, the analyzed corpus can be regarded as a descriptive sample of the language.

Similar statements can still be found in the corpus linguistic literature today, more than 50 years on. Stubbs (2001b: 45), for example, states that “[f]indings from one corpus should be checked against an independent corpus” and that

it is essential to compare findings from different independent corpora (since all corpora have gaps and biases), and to cross-check corpus findings with data from different sources. (Stubbs 2001b: 72)

Partington uses a nice simile when he says that a set of results which are not compared with results from somewhere else is “like the sound of one hand clapping.” (personal communication) The two corpora, or “hands”, that were selected as the basis of this study are described in the following.

The British National Corpus spoken subcomponent (BNC_spoken). The British National Corpus (BNC) is an electronic collection of over 4,000 British English text samples (i.e. not full texts but parts of texts) from a variety of spoken and written sources. With an overall size of over 100,000,000 words (i.e. tokens), the BNC is one of the largest English language corpora world-wide.⁴⁴ Its spoken subcomponent makes up approximately 10%, the rest of the corpus consists of written data. The BNC covers a wide range of written text-types from different domains (e.g. arts, leisure, social science) and media (e.g. books, periodicals), and spoken material from informal conversations and more formal sources, such as meetings and lectures.

The corpus is available to researchers in two versions: the original BNC version 1.0 (released in 1995) and the technically enhanced and error-corrected BNC World Edition, which was launched in December 1999.⁴⁵ In addition to the full version, a BNC

Sampler is available, including roughly 2% of the total collection, half from written and half from spoken sources. The latest member of the BNC family is “BNC-baby” which covers four 1-million word corpora (imaginative writing, academic writing, newspaper texts, spontaneous conversations) extracted from the BNC World Edition (cf. Burnard 2003:2). All analyses for the present study are based on the 1995 version of the BNC (version 1.0).

Compiled in 1990–1994 in a collaborative project between researchers at Lancaster and Oxford University and commercial partners (Chambers Harrap, Longman, Oxford University Press), the BNC aims at giving “maximal coverage of the varieties of modern British English” (BNC release announcement, 21 May 1995, Humanist discussion group, Vol.9, No. 0034). In 1998 Aston and Burnard describe the corpus as “a microcosm of current British English in its entirety” (29). Four years later, however, Burnard (2002:68) states that “[f]rom being a sample of the whole of language, the BNC was rapidly re-positioned as a repository of language variety.” That means that the corpus is maybe not as representative of the entire English language as originally assumed but that it can certainly still function as a valuable linguistic resource, a welcome source for the construction of specialised corpora, and a basis of studies in text-type comparison.

The spoken component of the BNC consists of two parts: a demographic and a context-governed one. Researchers involved in the compilation of the former part adopted an approach of “demographic sampling” (cf. Crowdy 1995).⁴⁶ 124 British native-speakers, aged 15+ from across the country, were equipped with portable cassette-recorders and asked to tape all conversations they participated in over a period of a few days (cf. Crowdy 1993:259; Burnard 2000:13). On the recording practice Crowdy (1993:260) remarks:

It is our intention to record all conversations as non-intrusively as possible, so that the material gathered approximates closely to natural, spontaneous speech. In many cases the only person aware that the conversation is being taped is the person carrying the recorder.

The result of this demographic sampling was a corpus of 4,211,216 words (cf. Burnard 1995:7) of conversational English. Of course, not all spoken language is conversational. Hence, to complement the demographic BNC_spoken component, a second subcorpus was created which consists of less informal speech data. This context-governed part includes monologues (40%) and dialogues (60%) from four different “contextually based categories: educational, business, public/institutional, and leisure.” (Crowdy 1993:262) The subcorpus has a size of 6,153,761 words and contains, for instance, samples of lectures, interviews, sales demonstrations, business meetings, parliamentary proceedings, sports commentaries, and phone-ins (cf. Burnard 1995, 2000).

On the whole, BNC_spoken consists of transcripts from 1,200 hours of recordings or 10,365,464 words (cf. Crowdy 1995:224; Burnard 1995:7). It is difficult to say how representative the corpus really is of spoken BrNSE. As Crowdy (1995:224) notes, “[w]ith a corpus of spoken language there are no obvious objective measures that can

be used to define the target population or construct a sampling frame.” However, the spoken part of the BNC certainly provides a good sample of British speech in the early 1990s, particularly as it covers a reasonable amount of language variation (including speakers from different age groups, regions, and social classes).

The spoken British part of The Bank of English (BoE_brspok). The Bank of English (BoE), which was first launched in 1991 (cf. Krishnamurthy 2002a) by COBUILD,⁴⁷ is currently the largest available corpus of a range of spoken and written texts from different registers and different English language varieties (mainly BrE, but also AmE and AusE). The corpus “is considered to be a sample of contemporary English – no more, no less.” (Francis & Sinclair 1994: 190) In his book *Corpus, Concordance, Collocation* Sinclair (1991: 26) uses the expression “bank of English”, not (yet) referring to the COBUILD corpus but to English monitor corpora in general which would allow for “new kinds of access to the patterns of the language which bombard readers every day” (1991: 26).

At its release the BoE had a size of approximately 20 million words, but since the latest update in 2004, it contains roughly 524 million words (HarperCollins online).⁴⁸ Opposed to the fixed-size BNC, the BoE is a dynamic, so-called “monitor corpus” (cf. Sinclair 1982).⁴⁹ The overall size of the corpus is periodically increased, as is the range of included text-types, and older material is replaced with newer to keep the corpus up-to-date (cf. Krishnamurthy 2002a). High degrees of representativeness are probably easier to achieve with a monitor than with a non-dynamic corpus, because if we include diachronic aspects in our considerations, “[a]ny corpus that is not regularly updated rapidly becomes unrepresentative, in the sense that it no longer represents the language as currently written or spoken” (Hunston 2002a: 30). Concerning the BNC, there seem to be regrets on the compilers’s sides about not having designed the corpus as a monitor corpus:

Clearly, the design of the BNC entirely missed the opportunity to set up a grand monitor corpus, one which could watch the river of language flow and change across time.
(Burnard 2002: 68)

As Table 1 shows, the size of the BoE has increased steadily from its release to the present. My research relates to the early 2001 version of the corpus. The exact overall size at the time of data collection was 418,449,873 words, with a distribution over 19 spoken and written subcorpora as given in Table 2. For the present study only the

Table 1. The increasing size of the BoE (cf. Krishnamurthy 2002a, 2002b)

date	size	date	size
1991	20 million words	1996	323 million words
1993	121 million words	2000	418 million words
1994	167 million words	2001	448 million words
1995	211 million words	2002	450 million words

Table 2. The composition of the BoE (as of March 8th 2001)

name of subcorpus	word count	classification
today	26,606,537	UK Today newspaper
times	31,110,198	UK Times/Sunday Times newspaper
brephem	4,655,650	UK ephemera
usspok	2,023,487	US spoken press briefings/meetings
usnews	10,110,450	US regional newspapers
econ	15,897,437	Economist magazine
oznews	34,651,116	OZ newspapers
usbooks	32,471,311	US miscellaneous books
bbc	18,690,390	UK BBC World Service broadcasts
newsci	7,901,182	UK New Scientist magazine
sunnow	31,786,908	UK Sun/News of the World newspaper
npr	22,250,592	US National Public Radio broadcasts
brspok	20,078,901	UK spontaneous speech
usacad	6,341,959	US academic textbooks
brbooks	43,333,620	UK books
guard	32,339,864	UK Guardian newspaper
indy	30,386,339	UK Independent newspaper
usephem	3,513,282	US ephemera
brmags	44,300,650	UK general interest magazines

“UK spontaneous speech” subsection (brspok) was used. This spoken BrNSE subsection was extended from a size of about 4 million words in 1993 to over 20 million words in 1996. The later version contains 2,669 mainly post-1990 texts with an average text length of 7,000 words. Texts include informal face-to-face conversations from all parts of Britain, phone calls, business meetings, public talks, university lectures, and radio programmes. No spoken data has been added to the corpus since 1996 when the BoE had an overall size of 323 million words (Ramesh Krishnamurthy personal communication).⁵⁰ On the data collection for BoE_brspok, Sinclair notes:

[w]ithin the spoken component, the most difficult kind of language to collect was, as always, the informally recorded conversations of people going about their daily lives, without thought of their language being preserved in a corpus.
(Cobuild ³2001:x)

Despite the described difficulty, a huge amount of informal conversational data has been collected to form “a crucial element of the corpus” (Payne 1995:203).

Although BNC_spoken and BoE_brspok were compiled following rather different design strategies, they both consist of a range of spoken BrNSE material and both represent this language type to a relatively high degree.

4.2 The empirical method: BNC and BoE data collection, processing, and evaluation

After the selection of appropriate corpora, the actual corpus-driven work began. The following sections highlight the steps carried-out in the collection, processing, and evaluation of the corpus data from BNC_spoken and BoE_brspok.

4.2.1 Verbs under analysis

From a practical point of view, it would be rather difficult, if not impossible, to carry out a detailed empirical investigation of progressive forms of all verbs in the English language. The sheer quantity of thousands of different verb types and hundreds of thousands of verb tokens in large corpora like the ones used here prohibits any kind of comprehensive close analysis. Hence, as a first step in the corpus-driven procedure, individual verbs had to be selected for inclusion in the analysis.

The selection criterion was frequency of occurrence in spoken British English. The study was intended to be based on a set of frequently used verbs which belong to the core English vocabulary. In order to retrieve a list of the most frequent verbs in

Table 3. 100 high-frequency verbs selected for analysis (in alphabetical order)

1. accept	26. find	51. mean	76. spend
2. add	27. finish	52. meet	77. stand
3. agree	28. follow	53. move	78. start
4. ask	29. forget	54. need	79. stay
5. be	30. get	55. pay	80. stop
6. become	31. give	56. pick	81. suggest
7. believe	32. go	57. play	82. support
8. bet	33. happen	58. provide	83. suppose
9. bother	34. have	59. pull	84. take
10. bring	35. hear	60. put	85. talk
11. buy	36. help	61. read	86. tell
12. call	37. hold	62. remember	87. think
13. carry	38. hope	63. ring	88. try
14. change	39. imagine	64. run	89. turn
15. check	40. keep	65. save	90. understand
16. come	41. know	66. say	91. use
17. cost	42. learn	67. see	92. walk
18. cut	43. leave	68. seem	93. want
19. deal	44. let	69. sell	94. watch
20. do	45. like	70. send	95. wear
21. draw	46. listen	71. set	96. win
22. eat	47. live	72. show	97. wonder
23. expect	48. look	73. sit	98. work
24. explain	49. make	74. sort	99. worry
25. feel	50. matter	75. speak	100. write

spoken BrNSE, the “frequency list” feature of BNCweb was used. On the project website BNCweb is described as “a user-friendly, web-based client program for searching and retrieving lexical, grammatical and textual data from the British National Corpus (BNC).”⁵¹ The program was developed by a team of researchers at the University of Zürich, Switzerland, including Sebastian Hoffmann, Hans-Martin Lehmann, and Peter Schneider.

The frequency list I used is based on occurrences of verbal infinitives in BNC_spoken. The first 97 items from this list were selected, plus the verbs *BE*, *HAVE*, and *DO*. These three verbs occur more often in BNC_spoken than any of the listed verbs but do not feature in this particular BNCweb frequency list as they have separate word class tags attached to their different forms (VBI [infinitive of *BE*], VHI [infinitive of *HAVE*], and VDI [infinitive of *DO*] instead of VVI [infinitive of any other verb]). Table 3 displays the 100 high-frequency verbs that have been selected for inclusion in the following detailed examination of progressive forms. Large amounts of corpus data were collected for the -ing forms of these 100 verbs. Section 4.2.2 describes the query strategies that were used in analysing BNC_spoken and BoE_brspok with the help of different corpus-analytic software tools.

4.2.2 The collection of corpus data: Query strategies

As Stubbs (2000:27) aptly notes “technology increases our powers of observation” and “concordance software can reveal patterns invisible to the naked eye.” In the collection of corpus data featuring progressive forms, extensive use was made of technology.

Large computerised corpora would be rather useless tools if the researcher did not have the appropriate corpus-analytic programs at hand. Among other things, well-

```

1      he environmental. They had a working committee did n't they
2      do we have, we do n't have a working group for children. N
3      se in Stortford Twelve is a working number. Oh well I hop
4      ew working party, which is a working party that 's erm basi
5      out, to, to people, erm and working with people, erm, an
6      amateur companies apparently working as a casual member of
7      tival. Some of you I know are working but some now Roger ar
8      time to assess how things are working out. We shall be anal
9      at the, the level that we are working at it is really coun b
10     what we do, you know we are working at one particular area
11     refore at any one time we are working, I mean that 's what
12     oo close. I will actually be working to start with in this
13     s that Caroline actually, er working on with Warwick Univer
14     duction as far as ordinary er working conditions are concern
15     e ca often work sideways from working with this door open?
16     you more stupid Mm you know working type chairman or somet
17     l as, as socialize while I 'm working erm you 've got to fin
18     ng it here and I 'm, so I 'm working on the assumption that
19     we decided because it was n't working altogether an Edinburg
20     e terms like team working net working erm. Has he been on a

```

Figure 1. An example of a KWIC concordance (searchword *working*), sorted to the left (1L, 2L)

chosen software searches millions of words within seconds and enables the retrieval of datasets from corpora. The central function of software packages for corpus analysis is the concordance function which serves to compile long lists of textual examples of a word or phrase from a corpus, usually with the searchword or searchphrase in the centre of the computer screen and the context in which the item appears displayed on the left and on the right. This so-called key-word-in-context (KWIC) format, as illustrated in Figure 1, is the most common and probably the most useful format of a concordance.

To retrieve a large number of progressives in context from the spoken parts of the British National Corpus and The Bank of English, extensive use was made of two software packages especially developed for corpus work: *WordSmith Tools* and *Lookup*.

The BNC and WordSmith Tools. The original version of the BNC (version 1.0) was distributed on three CD-ROMs that contained the compressed corpus files and a corpus-analysis program called “SARA” (SGML Aware Retrieval Application). SARA is a concordance software specifically designed for the BNC. It is not the most user-friendly program for corpus analysis. For instance, subcorpora (such a BNC_spoken) have to be defined anew for every single query and searches in larger datasets (especially for high-frequency items) can take quite a long time (cf. also Aston & Burnard 1998: 46). There are also some restrictions concerning the search functions available with SARA, e.g. the lack of wildcard features which would enable the retrieval of all forms ending in or beginning with a particular letter combination.⁵²

As I was sketching my project and thinking about the most appropriate BNC data collection strategies, a few other software packages became commercially available. Programs like *MonoConc Pro* (Michael Barlow, Athelstan) and *WordSmith Tools* (Michael Scott, 1996, Oxford University Press, henceforth WST)⁵³ suddenly offered researchers more sophisticated and faster means for corpus analysis. I familiarised myself with WST, using a random collection of text files on my hard drive as a mini-corpus, and preferred the search facilities of this program over the SARA version available with the BNC (1.0). As a next step, a way of using the BNC with WST on an ordinary PC had to be found. It was at this early stage in my research that I found out about and learned to treasure the help of subscribers to the Bergen-based CORPORA mailing list.

An “innocent” query of mine to the list in early 2000, a time when I was still planning to include other learning problems like future expressions in my analysis and hence asked about options to trace constructions of the “BE + V-ing” kind,⁵⁴ evoked a strong response from the list and put me in touch with researchers who helped me solve my query problems and showed me (and the CORPORA community) a way of using the BNC on the PC with non-SGML/XML-based concordancers such as WST or *MonoConc Pro*. In one of his messages (to the list and me) Chris Tribble described a procedure of unpacking the BNC from disk 1 of the original 3-CD-ROM distribution package (the 3-CD pack required a UNIX operating system for installation) and explained how the corpus can be used on a PC/Windows computer as a collection of plain text files and how the corpus can be divided into subsets which can then be

searched separately.⁵⁵ I followed Tribble’s corpus installation steps and separated the 10 million word spoken BNC component from the 90 million word written part. I was then able to apply the full range of the WST corpus analysis features to BNC_spoken.

To briefly describe the software, WST “is an integrated suite of programs for looking at how words behave in texts.” (Scott 1997:7; cf. also Berber Sardinha 1996) The individual tools (“WordList”, “Concord”, and “KeyWords”) can be used to carry out a number of simple (and a few more complex) corpus- or text-analytic procedures such as the compilation of frequency wordlists or the extraction of recurring word clusters. For the retrieval of concordances of progressive forms from BNC_spoken, the WST “Concord” tool was used. WST Concord is a program, known as a concordancer, which serves to generate concordances, i.e. listings of the occurrences (tokens) of a particular word or phrase (type) in a corpus, displayed in KWIC format (see Figure 1).

BNC_spoken concordances were compiled for the -ing forms of each of the 100 selected high-frequency verbs (see Section 4.2.1). The verb forms under analysis are

Table 4. 100 -ing forms under analysis (in alphabetical order)

1. accepting	26. finding	51. meaning	76. spend
2. adding	27. finishing	52. meeting	77. stand
3. agreeing	28. following	53. moving	78. start
4. asking	29. forgetting	54. needing	79. stay
5. becoming	30. getting	55. paying	80. stop
6. being	31. giving	56. picking	81. suggesting
7. believing	32. going*	57. playing	82. support
8. betting	33. happening	58. providing	83. supposing
9. bothering	34. having	59. pulling	84. taking
10. bringing	35. hearing	60. putting	85. talking
11. buying	36. helping	61. reading	86. telling
12. calling	37. holding	62. remembering	87. thinking
13. carrying	38. hoping	63. ringing	88. trying
14. changing	39. imagining	64. running	89. turning
15. checking	40. keeping	65. saving	90. understanding
16. coming	41. knowing	66. saying	91. using
17. costing	42. learning	67. seeing	92. walking
18. cutting	43. leaving	68. seeming	93. wanting
19. dealing	44. letting	69. selling	94. watching
20. doing	45. liking	70. sending	95. wearing
21. drawing	46. listening	71. setting	96. winning
22. eating	47. living	72. showing	97. wondering
23. expecting	48. looking	73. sitting	98. working
24. explaining	49. making	74. sorting	99. worrying
25. feeling	50. mattering	75. speaking	100. writing

* In addition to multiple instances of *going* in progressive constructions in BNC_spoken, we also find 11,968 entries for *gonna*. This form has not been included in our study, since we were not interested in different pronunciation or spelling variants of the items under investigation, but in their typical functions and contexts. With respect to these features, significant differences between the two variants were not to be expected. As Berglund (2000: 163) has found in an empirical study based on BNC_spoken data, “*gonna* and *going to* are to a large extent found in similar syntactic and lexical environments, that is, the two variants co-occur with the same word-classes and lexical items in similar patterns.”

listed in Table 4. For those items that occurred more often than 200 times in the corpus (i.e. most of the verb forms in the chosen set) the “random set” option in the Concord settings was activated to retrieve a random selection of concordance lines from a maximum number of different BNC_spoken files, and not just examples from a few texts. The datasets of a maximum of 200 concordance lines were saved as text files for further analysis with a contextual span of 200 characters per line. Before the concordances were saved, however, they were sorted “1L, 2L”, which means that the items in positions one and two on the left of the searchword (e.g. *working*) were arranged in alphabetical order. Such a sorted concordance facilitates the investigation of repeated co-occurrences of particular words and enables the researcher to spot significant collocations or clusters at a glance. Although with verb forms it is usually “more revealing to sort to the right, since most verbs have complementation patterns which follow them” (Hunston & Francis 2000: 37), I decided to work with left-sorted concordances. In the case of -ing forms left-sorting highlights typical subjects and makes it easier to distinguish progressives (e.g. *two consultants are working on ward four*) from non-progressives (e.g. *it’s a working account*).

Three of the selected -ing forms (*having*, *supporting*, and *taking*) yielded somewhat problematic results. A large number of the 200 examples in each of the three concordances showed exactly the same line; a rather unusual phenomenon, considering that there should not be several copies of one text in a corpus. Figure 2 displays part of the concordance of *supporting*. The lines which appeared several times in the *having*, *supporting*, and *taking* datasets are part of the BNC header, a long paragraph that precedes each text in the BNC. This paragraph describes the contents of the individual files and includes some general information about the corpus and text availability (cf. Burnard 2000: 36). A small part of this header is provided in Figure 3.

In order to cut the headers from all BNC_spoken files, the WST settings had to be adjusted. This was done by selecting “only part of file” in the “tags” section in the

```

1  , without the state actually supporting the monarchy, the mon
2  great, helping others. And supporting. Supporting him yeah,
3  SERVICES, from whom forms and supporting materials are availabl
4  SERVICES, from whom forms and supporting materials are availabl
5  SERVICES, from whom forms and supporting materials are availabl
6  SERVICES, from whom forms and supporting materials are availabl
7  SERVICES, from whom forms and supporting materials are availabl
8  SERVICES, from whom forms and supporting materials are availabl
9  SERVICES, from whom forms and supporting materials are availabl
10 SERVICES, from whom forms and supporting materials are availabl
11 SERVICES, from whom forms and supporting materials are availabl
12 SERVICES, from whom forms and supporting materials are availabl
13 SERVICES, from whom forms and supporting materials are availabl
14 SERVICES, from whom forms and supporting materials are availabl
15 SERVICES, from whom forms and supporting materials are availabl
16 SERVICES, from whom forms and supporting materials are availabl
17 SERVICES, from whom forms and supporting materials are availabl
18 SERVICES, from whom forms and supporting materials are availabl

```

Figure 2. *Supporting* in BNC_spoken, including headers

```
<bncDoc id=BDK6J n=100102>
<header type=text creator='dominic' status=new update=1994-11-27>
  <fileDesc>
    <titStmt>
      <title>
        Tutorial lesson--- an electronic transcription
      </title>
    <respStmt>
      <resp>
        Data capture and transcription
      </resp>
      <name>
        Longman ELT
      </name>
    </respStmt>
  </titStmt>
  <ednStmt n=3>
    Header automatically generated by mkhdr 0.31
  </ednStmt>
  <extent kb=234 words=12013>
  </extent>
  <pubStmt>
    <respStmt>
      <resp>
        Archive site
      </resp>
      <name>
        Oxford University Computing Services
      </name>
    </respStmt>
    <address>
      13 Banbury Road, Oxford OX2 6NN U.K.
      Telephone:      +44 491 273280
      Facsimile:      +44 491 273275
      Internet mail:  natcorp@ox.ac.uk
    </address>
  <idno type=bnc n=100102>
    100102
  </idno>
  <avail region=world status=unknown>
    THIS TEXT IS AVAILABLE THROUGHOUT THE EUROPEAN UNION only as
    part of the British National Corpus at nominal charge FOR
    ACADEMIC RESEARCH PURPOSES SUBJECT TO A SIGNED END USER
    LICENCE HAVING BEEN RECEIVED BY OXFORD UNIVERSITY COMPUTING
    SERVICES, from whom forms and supporting materials are
    available.
```

Figure 3. Text sample taken from the header of BNC_spoken text K6J

Concord “settings” menu. In the new pop-up window “select only sections of text files” under “sections to cut” I selected “starting at: start of file” and “ending with: </header>” and checked “activated” (see Figure 4). It was then possible to repeat the queries for *having*, *supporting*, and *taking* and to retrieve results only from the transcripts proper.

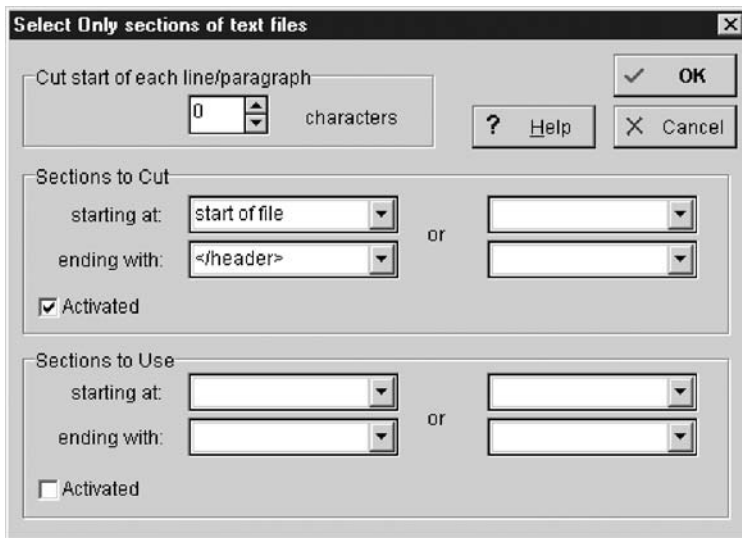


Figure 4. Adjustment of WST settings to cut BNC headers

The BoE and Lookup. *Lookup* is the corpus analysis program which serves to search The Bank of English.⁵⁶ Sinclair (1997: 39) describes the system as “an exceptionally efficient and user-friendly means of access for language researchers to the riches of the corpus.” Indeed, *Lookup* is easy to use and very fast. Bernardini (2000: 123–124) also mentions “[s]peed and flexibility” as “the most striking features” of the software but calls the program “not [...] particularly user-friendly”, referring to the complex range of its functions. *Lookup* can be accessed via a telnet connection with titania.cobuild.collins.co.uk. Once a researcher is logged onto this system, she or he can select the “interactive corpus access tool” from a menu and then choose a BoE subcorpus or, alternatively, search the whole database.

To retrieve data illustrating the use of progressives in spoken BrNSE, the “br-spok” subsection was selected and queries on the 100 -ing forms listed in Table 4 were performed. For each verb form the system was asked to display a random set of a maximum of 100 concordance lines in KWIC format. After the manual processing of the saved BNC_spoken concordance lines had been very time-consuming (cf. 4.2.3), I decided that sets of 100 (instead of 200) lines for each of the 100 -ing forms had to be large enough to enable a comparison with the BNC data. Like the BNC datasets, the BoE concordances were sorted to the left of the searchword, and the concordance lines were saved as text files with a contextual span of 200 characters per line. The next section deals with further steps in the processing of these “raw” sets of concordance data from BNC_spoken and BoE_brspok.

4.2.3 Data filtering

From the saved concordances of -ing forms, first of all those instances had to be deleted in which the searchword was not part of a progressive construction but functioned, for example, as noun (as e.g. in *she'll play about with some helping to get some answers to these fractions*, BNC_spoken) or adjective (as e.g. in *students need a helping hand*, BNC_spoken). This filtering process was carried out manually. Figures 5 and 6 show two samples taken from an unfiltered and a filtered concordance of the item *working* in BNC_spoken.

```

1      t group. We also get requests from work from the organisational review working
party, which is a working party that 's erm basically looking to progress to the D M D,
Decentralization of Microtizat 15.162 c:\bnc\d95 83
2      rds, if half the interest is er is erm deemed to be the male 's and therefore
taxable, if it 's a working account, it 's never likely to get much interest on it anyway,
No. cos the money will come 8.139 c:\bnc\g4h 25
3      is a thing now, social security. The the guy ca n't get that cos I had a few
shillings more after working two years. Oh I see. Aye. That 's not fair r really is it. No
well that 's the way the coun 3.371 c:\bnc\heu 35
4      go. Now we 're back to six people again and they 're all working for me in a very
small way. Again working with the same sort of people, but in a very small and for our own as
well. So that 's the na 658 c:\bnc\gyn 10
5      in your national organization and fight the government. You ca n't do so locally,
if you 're all working separately. You 've got to unite and get in the one gives us strength
to fight them. Is er 12.114 c:\bnc\gyu 88
6      messing around with the Tax Officer for quite a while because I was on demob leave
and I was also working so he was n't gon na let me have two lots of sa wages. How had th the
transport changed then 6.835 c:\bnc\hdl 51
7      what we 'll do we 'll get those set up at the sort of the appropriate working
temperatures, and working humidity and then we will have to deal with people on an individual
basis, if basically, the 3.895 c:\bnc\fls 31
8      bronchi, and you can get these enormous casts of mucrotic material er, starting in
the pharynx and working their way right down in the bronchi. It 's been known for a very long
time that from these c 3.039 c:\bnc\f8s 49
9      with learning disability. I mean it you know that would give you the sort
working in a team and working with people and erm Mm. Mm. Mm. I du n no. It depends though. If
you want to work with young 9.486 c:\bnc\hdy 93
10     go on the dodgems. Good. Er and then full day erm Good. a full day programme on
the Thursday and working up until erm mid-afternoon on the Friday. Er and Jane has sent us in
some proposals, which e 19.016 c:\bnc\fu j 88
11     fing levels at the hospital where killer nurse Beverley Allitt worked. Only two
consultants are working on ward four at Hospital a year after a health authority report called
for increased supervi 7.554 c:\bnc\fxt 40
12     ve Marine Laboratory, Tyne & Wear tutorial ecology talking about projects
that students are working on 1994-11-27 Initial accession to corpus dominic Doctor 's
obviously a more frugal character 340 c:\bnc\f8f 24

```

Figure 5. Part of the unfiltered BNC_spoken concordance of *working* (lines 1–12 of 200)

```

5      in your national organization and fight the government. You ca n't do so locally,
if you 're all working separately. You 've got to unite and get in the one gives us strength
to fight them. Is er 12.114 c:\bnc\gyu 88
6      messing around with the Tax Officer for quite a while because I was on demob leave
and I was also working so he was n't gon na let me have two lots of sa wages. How had th the
transport changed then 6.835 c:\bnc\hdl 51
11     fing levels at the hospital where killer nurse Beverley Allitt worked. Only two
consultants are working on ward four at Hospital a year after a health authority report called
for increased supervi 7.554 c:\bnc\fxt 40
12     ve Marine Laboratory, Tyne & Wear tutorial ecology talking about projects
that students are working on 1994-11-27 Initial accession to corpus dominic Doctor 's
obviously a more frugal character 340 c:\bnc\f8f 24

```

Figure 6. Part of the filtered BNC_spoken concordance of *working* (lines 1–4 of 76 remaining lines)

Sometimes the corpus examples were rather fragmentary (due to their spoken nature) and, in very few cases, the defined context of 200 characters per line was not sufficient to determine whether a particular instance contained a progressive or not.⁵⁷ In such cases of doubt the concordance line was eliminated in order to make sure that the remaining set of -ing form occurrences was an “all-progressive” set. The overall number of collected concordance lines from the two corpora was reduced rather significantly in the process of manual filtering from originally 27,252 (17,676 BNC_spoken plus 9,576 BoE_brspok) to 9,468 (6,311 BNC_spoken plus 3,157 BoE_brspok). Table 5 lists for each of the 100 -ing forms the number of examples in the unfiltered and filtered concordances from BNC_spoken and BoE_brspok (cf. 4.5.1 for a discussion of the distribution of collected progressive form tokens across verb types). After the manual filtering process the reduced datasets were again saved as text files (e.g. “working BNC filtered.txt”) which could then be imported into an *Access* database.

Table 5. Numbers of originally collected and remaining concordance lines (after the filtering)

verb form	BNC_spoken		BoE_brspok		verb form	BNC_spoken		BoE_brspok	
	before filt.	after filt.	before filt.	after filt.		before filt.	after filt.	before filt.	after filt.
accepting	94	36	99	30	meaning	200	28	100	9
adding	166	43	100	18	meeting	200	3	100	10
agreeing	62	19	100	33	moving	200	57	100	34
asking	200	103	100	48	needing	65	19	100	23
being	200	6	100	2	paying	200	101	100	49
becoming	200	109	100	55	picking	200	80	100	31
believing	49	2	51	4	playing	200	78	100	32
betting	34	4	38	3	providing	200	35	100	24
bothering	48	25	83	31	pulling	200	71	100	28
bringing	200	65	100	25	putting	200	71	100	30
buying	200	83	100	33	reading	200	53	100	19
calling	200	74	100	50	remembering	94	16	99	14
carrying	200	72	100	35	ringing	200	75	100	51
changing	200	60	100	29	running	200	68	100	28
checking	200	55	100	30	saving	200	45	100	12
coming	200	61	100	39	saying	200	95	100	58
costing	146	80	100	47	seeing	200	40	100	35
cutting	200	49	100	16	seeming	4	1	14	1
dealing	200	89	100	39	selling	200	72	100	32
doing	200	110	100	55	sending	200	74	100	28
drawing	200	49	100	14	setting	200	38	100	16
eating	200	77	100	33	showing	200	64	100	35
expecting	200	154	100	74	sitting	200	80	100	45
explaining	102	28	100	20	sorting	110	22	100	11
feeling	200	54	100	21	speaking	200	50	100	20

Table 5. (continued)

finding	200	49	100	42	spending	200	48	100	29
finishing	132	47	100	36	standing	200	55	100	38
following	200	22	100	19	starting	200	72	100	38
forgetting	85	20	100	17	staying	200	91	100	44
getting	200	97	100	47	stopping	200	77	100	31
giving	200	86	100	40	suggesting	200	157	100	66
going	200	137	100	71	supporting	200	47	100	22
happening	200	157	100	77	supposing	104	0	100	1
having	200	48	100	22	taking	200	86	100	34
hearing	200	47	100	20	talking	200	132	100	56
helping	200	56	100	29	telling	200	106	100	54
holding	200	64	100	33	thinking	200	93	100	48
hoping	200	149	100	80	trying	200	124	100	60
imagining	18	9	27	13	turning	200	56	100	24
keeping	200	57	100	24	understanding	200	7	100	2
knowing	200	0	100	1	using	200	56	100	35
learning	200	44	100	22	walking	200	65	100	32
leaving	200	79	100	31	wanting	200	76	100	35
letting	200	68	100	28	watching	200	80	100	34
liking	34	6	65	2	wearing	200	107	100	42
listening	200	110	100	46	winning	159	39	100	11
living	200	43	100	34	wondering	200	159	100	80
looking	200	104	100	42	working	200	76	100	40
making	200	76	100	28	worrying	169	44	100	16
matterng	1	0	0	0	writing	200	40	100	22

4.2.4 Data processing and encoding: The construction of an Access database

In order to make it easier to handle the still large amount of corpus data that remained after the filtering procedure, an electronic database was created on the computer with the help of the Microsoft Access software (version 97). The following paragraphs deal with the import of data into this database and the manual annotation of the data according to a number of semantic and contextual criteria.

Entry of concordance lines. As a first step in the construction of the database, all 9,468 progressives in context, i.e. all concordance lines from the two spoken BrNSE corpora, had to be inserted into one Access file. This was done by using the “import external data” function from the file menu of an empty database. Each of the 196⁵⁸ filtered concordance files was imported separately. The program was set to regard the tabs in the concordance lines as delimiters for column assignment. That means that the characters that followed a tab were copied to a new column in the database. The result of the data import was a long table which displayed each single progressive, i.e. each BNC_spoken and BoE_brspok concordance example in a separate row. It was then

possible to add and define new columns and to annotate the progressives in some detail, selecting appropriate function and context variables and values.

Feature determination and concordance line annotation: Creating and filling boxes. In *Access*, a database consists of several parts, most notably tables, forms, and queries. The imported corpus data was first of all stored in a table. The subsequent creation of customised forms based on the same data made it easier to enter further information about each token into the database. Working with forms can help minimise data entry errors as all information about one particular corpus example in form view is displayed at a time and no horizontal scrolling through a wide table is required. A number of boxes was created in form view to allow for a detailed annotation of the contexts and functions of each progressive verb form token.

Due to the corpus-driven approach taken in this study, the creation of boxes, i.e. the determination of function and context features, and especially the determination of feature values, had to be a dynamic process. In the process of annotating several thousand concordance examples, the set of context and function features and values was constantly adjusted in the light of the data. It occurred a few times that certain features which had appeared to be ideal to describe the use of progressives in an early database-annotational stage did not fully capture the functions of concordance lines examined later. It was impossible to group all the data into the categories that had been defined early in the process of database labelling. This “lack of fit” between corpus data and existing or early-defined categories is a well-known and common problem for corpus-driven linguists. In this context John Sinclair’s experiences in the COBUILD project are a case in point. He states:

I grossly underestimated the effect of the new information that the corpus supplied, and in particular the total lack of fit between the evidence coming from the corpus and the accepted categories of English lexicography. The Cobuild team had to reconceptualise the dictionary in the light of the early evidence.

(Sinclair 2004b:9)

A considerable amount of reconceptualisation of initially determined features was also involved in the encoding of corpus data in the present study. In the end, however, an appropriate set of features and feature values had been devised to account for all functions of the large number of progressives in the database. Also, the list of context features had been expanded to yield the most central information about typical co-occurrences in progressive constructions. The eventual sets of context and function features and their possible values are displayed in Tables 6 and 7. They will be exemplified and explained in some detail in Sections 4.3 and 4.4. After all progressives in the *Access* database had been annotated according to these lists of features and values, the data evaluation could be carried out semi-automatically (cf. 4.2.5).

Table 6. Function features under analysis

Central function features	Values
time reference	past, present, future, indeterminate
repeatedness	repeated, non-repeated
continuousness	continuous, non-continuous
Additional function features	
general validity	
politeness/softening	
emphasis/attitude	
shock/disbelief	
gradual change/development	
habituality	
framing	

Table 7. Context features under analysis

Feature	Possible feature values
tense form distribution	present progressive (e.g. <i>is looking</i>) past progressive (e.g. <i>were looking</i>) present perfect progressive (e.g. <i>'ve been looking</i>) past perfect progressive (<i>had/d been looking</i>)
TO BE contraction	short form (e.g. <i>'m, 've been</i>) long form (e.g. <i>am, have been, was</i>)
subject	e.g. <i>I, she, they, the customer</i>
preposition ⁵⁹	e.g. <i>at, for, up</i>
object	e.g. <i>it, a job, the woman on the bridge</i>
time adverbial	e.g. <i>now, at the moment, last year</i>
place adverbial	e.g. <i>here, there, in the dining room</i>
other adverbial	e.g. <i>actually, desperately, quite</i>
negation	negated (+) / not negated (–)
question	+ / –
if-clause	+ / –
relative clause	+ / –

4.2.5 Data evaluation

Access query strategies. The main function used to retrieve results from the annotated database was the *Access* filter function. The definition of filters in table or form view makes it possible to display only specific records which fit certain criteria.

For instance, if we only want to examine instances of *was looking* in BNC_spoken that are followed by the preposition *at*, a complex filter (a so-called “filter by form”) can be created accordingly. After the application of the specifically created filter, *Access* shows only those records that match the criteria (in this case examples of *looking*

concordance line	searchword	corpus	subject/p	BE form	time adv	time
per Mhm. Just a normal sort of squared exercise book. I have n't got any with me, and I was just looking at graphs for someone else earlier. Yeah. Erm that 's I mean I would	looking	BNC	I	was	just	ea
s I say, that was the mood of rejection erm and there was something about the way the old girl was looking at the women on the bridge, almost as if there was this recognition	looking	BNC	the old girl	was		
ards and workshops hit. That was directly. I was just checking here in from the book by and I was looking at the chapter here on Roses Rivetus and er and it was on the	looking	BNC	I	was		
e 's, there 's ten million of them I would of thought so that 's right And it helps you Cos I was looking at with funding. you know, it does help What about Jimmy? because then they	looking	BNC	I	was		
hich is one I raised. Yes. Because that has an effect on how you can use people. Erm Yeah, I was looking at that myself yesterday actually and er if I include three, three Trident	looking	BNC	I	was	yesterday	

Figure 7. The application of an Access “filter by form” in table view (BNC_spoken query)

preceded by *was* and followed by *at*). Of course, the more complex the filter is and the more criteria are defined, the more restrictive the query and the fewer the matching results. In our example only five of the 9,468 corpus examples in the database meet the criteria searchword: “*looking*”, corpus: “BNC_spoken”, form of TO BE: “*was*”, and preposition: “*at*”. The same filter can be applied to the database in table view. Instead of a small set of five individual forms we then get a five-line extract from the full table, again including only those instances which match the defined criteria. Figure 7 gives a screenshot of the filtered database in table view.

Using filters by form, numbers for different feature combinations were retrieved separately for the selected corpora to enable comparisons (here for BNC_spoken and BoE_brspok, later also for the two EFL textbook corpora; cf. Chapter 5).

Comparing figures from different datasets. The further processing of the automatically retrieved figures from the Access database was carried out manually. To enable a BNC_spoken vs. BoE_brspok comparison of absolute numbers for a certain feature combination, relative frequencies based on the underlying set of data entries from the respective corpus were calculated.

To give an example from the context feature analysis (cf. 4.3), a customised filter rendered the following absolute numbers for negated progressives (e.g. *I'm not working for that*, BoE_brspok):

BNC_spoken: 542

BoE_brspok: 252

This means that 542 out of 6,311 concordance lines from BNC_spoken (i.e. 8.6%) and 252 of the 3,157 BoE_brspok examples in the database (i.e. 7.9%) contain a negation of the verb form. All figures and diagrams in the following sections (in Chs. 4, 5, and 6) which illustrate differences or similarities between different corpora or datasets are based on percentages, i.e. on relative frequencies.

Even though there may often be rather huge differences between the frequencies of occurrence of a particular phenomenon in two datasets, researchers often claim that it cannot simply be taken for granted that these differences are in fact significant. They might as well be a matter of chance. If we are in doubt whether our datasets are “different enough” and want to be more confident about the significance of our results, we can perform a range of statistical tests that give us this confidence.⁶⁰

A well-known and widely used statistical test in corpus analysis is the chi-square (or chi-squared) test. This test is claimed to be particularly useful if one wants to compare the frequencies of items in two or more different corpora (cf. Butler 1985:112; Abney 1996; Bald 1977; Tesch 1988). By comparing actually observed with probabilistically expected frequencies (always raw frequencies, no percentages), chi-square serves to determine whether found differences are significant or not and, if they are significant, at what probability level they are significant. In the following contextual and functional analyses of progressives in context I will always calculate chi-square values for each comparison of figures from different sets of data and comment on the results of this statistical test whenever considered appropriate. All chi-square tests were performed using the Georgetown University “Web Chi Square Calculator”.⁶¹

4.3 The use of progressives in spoken English (I) – contexts

It was pointed out above (see 2.2) that any work in corpus-driven linguistics stresses the centrality of contextual approaches to language analysis, usually along the lines of J. R. Firth and J. McH. Sinclair. For the corpus-driven researcher context in all its different shapes is an all-important phenomenon (cf. e.g. Stubbs 2001b or Teubert 1999). Also, I would like to argue that context is not only central in linguistic analysis and description but in language pedagogy too. I strongly believe that it makes sense to pay attention to collocation and colligation in language instruction and to teach lexical items in their typical syntactic and semantic contexts. This belief clearly echoes one of Sinclair’s (1997:34) earlier-mentioned data-centred pedagogical precepts, “[i]nspect contexts”, in which he “advocate[s] a much closer inspection of the verbal environment of a word or phrase than is usual in language teaching.”

A corpus-driven study of progressives, especially when it is in part pedagogically motivated, thus has to closely examine the contexts of the respective items under analysis and investigate which items are normally selected together by the compe-

tent speaker of English. The following sections will look at co-selection and present some essential findings from the close analysis of important features in the lexical-grammatical context of progressive forms in the selected datasets from BNC_spoken and BoE_brspok. Large amounts of concordance data from real spoken BrNSE have been searched to identify patterns in the combination of different types of progressives and subjects, objects, prepositions, adverbials, and other context phenomena. All parts of the contextual analysis include a comparison of figures from the two selected corpora. As discussed earlier on, the empirical findings can be called representative of spoken BrNSE with more certainty if we observe comparable results for both datasets. This “representativeness check” is the main reason for the comparison of BNC_spoken and BoE_brspok data throughout Chapter 4.

The results obtained in this part of the study will later serve as a basis for comparison with the findings drawn from the EFL textbook analysis (cf. 5.5).

4.3.1 Distribution of different tense forms

The first question I am going to ask in the contextual analysis of progressives in spoken BrNSE is one about the different types of progressive forms: What are the shares of the included progressive verb forms (present progressive [PresProg], past progressive [PastProg], present perfect progressive [PresPerfProg], past perfect progressive [PastPerfProg]) in the datasets from BNC_spoken and BoE_brspok? In order to determine these shares, absolute frequencies of occurrence for each possible form of TO BE preceding individual -ing forms were retrieved from the database, applying the query

Table 8. The co-selection of progressives and forms of TO BE⁶²

Form of TO BE	BNC_spoken (6,311 examples) absolute and relative frequencies	BoE_brspok (3,157 examples) absolute and relative frequencies
<i>am</i>	40 (0.63%)	23 (0.73%)
<i>'m</i>	846 (13.41%)	430 (13.62%)
<i>are</i>	786 (12.45%)	339 (10.74%)
<i>'re</i>	1,505 (23.85%)	737 (23.34%)
<i>is</i>	403 (6.39%)	166 (5.26%)
<i>'s</i>	865 (13.71%)	385 (12.20%)
<i>was</i>	978 (15.50%)	626 (19.83%)
<i>were</i>	568 (9.00%)	301 (9.53%)
<i>has been</i>	35 (0.55%)	4 (0.13%)
<i>'s been</i>	29 (0.46%)	31 (0.98%)
<i>have been</i>	31 (0.49%)	18 (0.57%)
<i>'ve been</i>	125 (1.98%)	75 (2.38%)
<i>had been</i>	17 (0.27%)	8 (0.25%)
<i>'d been</i>	22 (0.35%)	14 (0.44%)
none (empty box)	49 (0.78%)	0
<i>ai</i> (in <i>ain't</i>)	12 (0.19%)	0

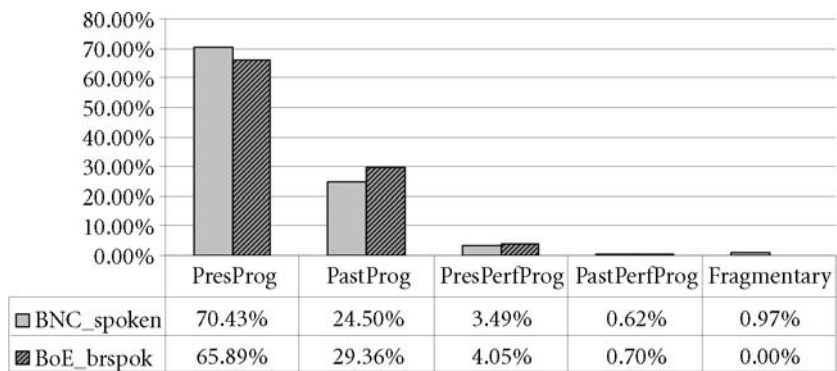


Figure 8. The distribution of different tense forms in BNC_spoken and BoE_brspok

strategies explained above (see 4.2.5). Table 8 displays the retrieved absolute numbers together with their relative frequencies in the BNC_spoken and BoE_brspok datasets.

In both corpora, highest numbers were found for the patterns *'m V-ing*, *are V-ing*, *'re V-ing*, *is V-ing*, *'s V-ing*, *was V-ing*, and *were V-ing*. All other forms (*am V-ing*, and all perfect and fragmentary forms) are comparatively rare. Apparently, with the exception of *am*, there is a clear difference in distribution between patterns that contain a simple form of TO BE (e.g. *are*, *was*) and those which are combinations of a complex auxiliary form and *V-ing* (e.g. *have been*). This result becomes even more evident when we summarise the individual TO BE-form occurrences under the traditional tense forms PresProg, PastProg, PresPerfProg, and PastPerfProg. The distribution of progressive tense forms in BNC_spoken and BoE_brspok corpus data is visualised in Figure 8. Despite their somewhat different internal composition (cf. 4.1.4), the results for the two corpora are rather similar.⁶³ By far the most frequent form in both corpora is the present progressive with slightly more than 70% in BNC_spoken and almost 66% in BoE_brspok. If we split up the present progressive into its different possible forms as displayed in Figures 9 and 10, we see that the most common PresProg pattern in both corpora is the plural form *'re/are V-ing* and that equally large shares go to *'s V-ing* and *'m V-ing* while *am V-ing* and *is V-ing* are rather infrequent. Again, the figures for the two analysed datasets are astonishingly similar.

Past progressives are far less common in the BNC_spoken and BoE_brspok datasets than their present counterparts but still make up 24.50 and 29.36 per cent (cf. Figure 8). While the relative frequency for the PresProg was about 4.5% higher in BNC_spoken than in BoE_brspok, this distribution is turned upside down in the case of past progressives. The shares of PresPerfProg (3.49% and 4.05%) and PastPerfProg (0.62 and 0.70%) are fairly low in both corpora. In addition, the BNC_spoken dataset included a couple of fragmentary progressive constructions in which the form of TO BE was either missing (e.g. in *What you wearing tomorrow night?*) or which contained the short form *ai + n't* (e.g. in *I ai n't telling you no more*), a variant of TO BE in negative contexts that can apply to all persons and hence could not be assigned to any

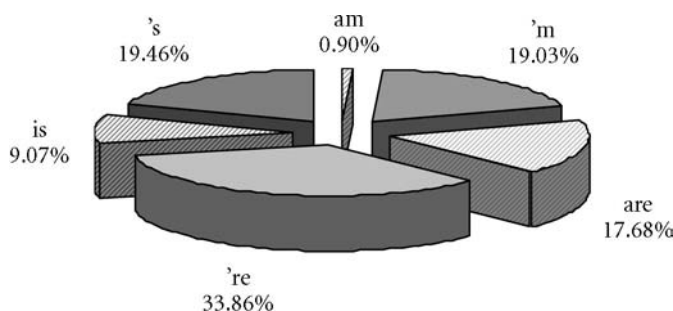


Figure 9. The distribution of forms of TO BE within the PresProg (BNC_spoken)

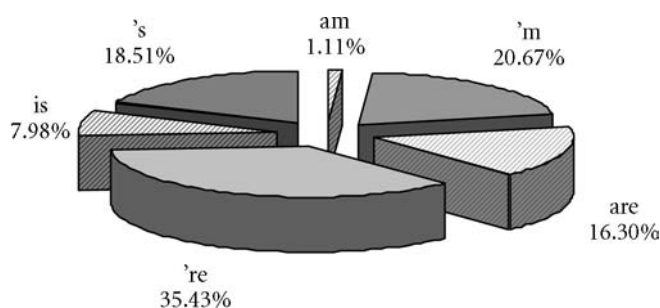


Figure 10. The distribution of forms of TO BE within the PresProg (BoE_brspok)

of the other standard present tense forms of TO BE. Such fragmentary cases were not found in the BoE_brspok data. This may be due to different transcription practices in the compilation of this subcorpus or, again, to the somewhat different kinds of spoken data included. However, this difference is not to be overemphasised as only 0.97% of all BNC_spoken examples fall into the “fragmentary” category.

Figures 11 to 16 show the distributions of the occurring forms of TO BE within the PastProg, the PresPerfProg, and the PastPerfProg in the BNC_spoken and BoE_brspok datasets. As has been the case for different PresProg patterns, the shares of the individual constructions are on the whole fairly similar across the two datasets. Significant (chi-square-tested) differences could only be determined in the cases of *has been* and *'s been* (cf. Figures 13 and 14).

If we compare the findings of this corpus-driven analysis of progressive tense forms with the empirical results of other scholars, we find both similarities and differences. Figure 17 illustrates the distribution of progressive tense forms that Ota (1963), Joos (1964), and Allen (1966) have found in their data. As we can see in this figure, the results of Ota differ significantly from those obtained by Joos and Allen which are largely comparable. These differences between Ota's findings on the one hand and Joos's and Allen's on the other can probably be mainly ascribed to register differences in their data. While the works of Joos and Allen are based on exclusively *written* British

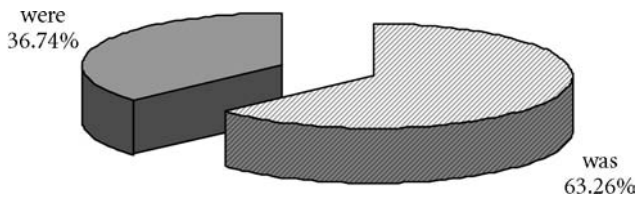


Figure 11. The distribution of forms of TO BE within the PastProg (BNC_spoken)

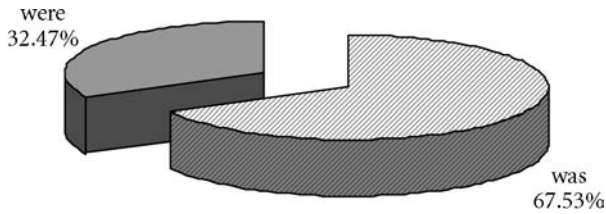


Figure 12. The distribution of forms of TO BE within the PastProg (BoE_brspok)

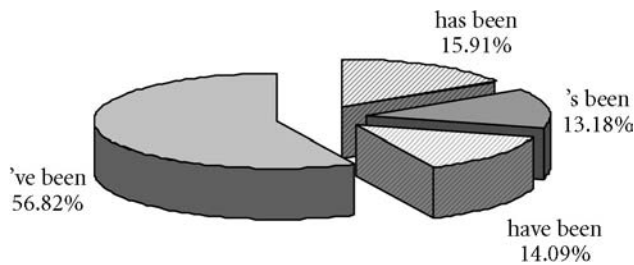


Figure 13. The distribution of forms of TO BE within the PresPerfProg (BNC_spoken)

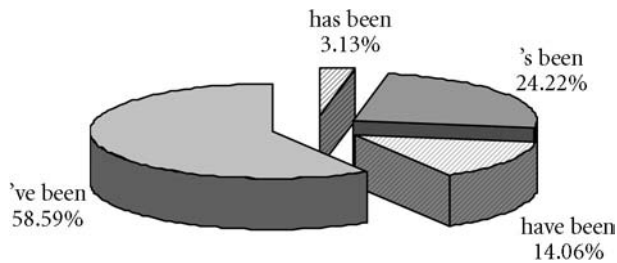


Figure 14. The distribution of forms of TO BE within the PresPerfProg (BoE_brspok)

(Joos) and American English (Allen), Ota draws his data from a corpus of mainly *spoken* American English material. It is thus not very surprising to see that Ota’s figures are much closer to my own results from spoken BrNSE (cf. Figure 8 above) than to numbers that go back to written English.

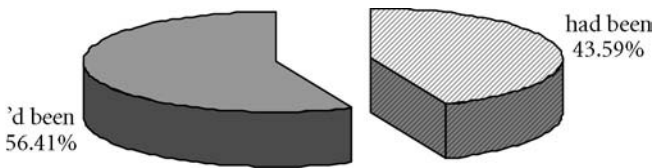


Figure 15. The distribution of forms of TO BE within the PastPerfProg (BNC_spoken)

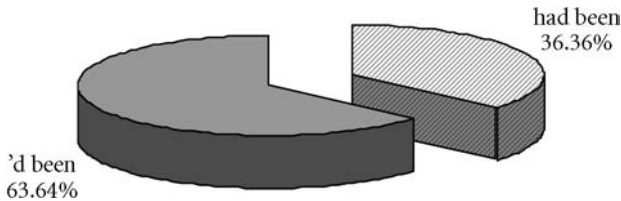


Figure 16. The distribution of forms of TO BE within the PresProg (BoE_brspok)

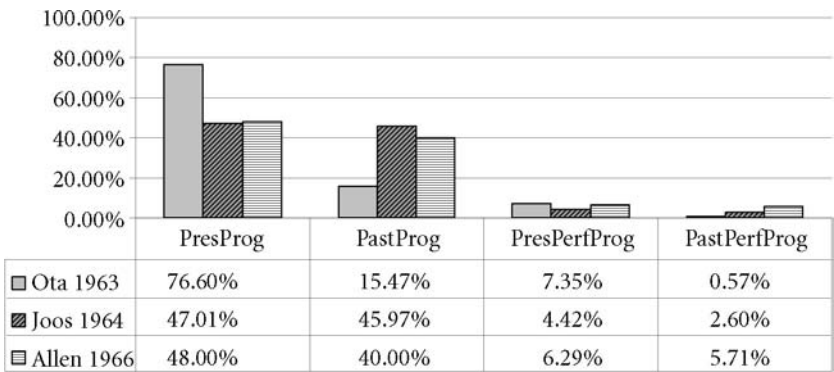


Figure 17. The distribution of progressive forms in three earlier empirical studies

Biber et al.'s (1999:462) frequency findings of present and past progressives in conversation look very similar to my figures too. Unfortunately, the authors do not give any concrete numbers but only display the verb form distribution in a diagram, a fact which makes a real comparison with my results rather difficult. From their diagram, however, I would infer that about 30% of the progressives in the CONV subcorpus are PastProgs and about 70% are PresProg forms.

As the results of my own analysis and other researchers's investigations demonstrate, there is obviously a close connection between tense form distribution and textual register. The data indicate that speech contains much higher shares of PresProgs and lower shares of PastProgs than writing and vice versa. Section 5.5 will discuss whether the spoken-type language of EFL textbooks mirrors the significant distribution of progressives found in real spoken English.

4.3.2 Tense form contractions

Progressives can be realised either by a contracted “short” form of TO BE⁶⁴ plus V-ing, as illustrated in example (1), or by a non-contracted “long” form of TO BE plus V-ing as in example (2), both from BoE_brspok.

- (1) Think it’s because I’m telling you the truth. (BoE_brspok)
- (2) I don’t mean to sound unsympathetic but I am telling you the absolute truth. (BoE_brspok)

The shares of contracted and non-contracted progressive constructions in BNC_spoken and BoE_brspok are displayed in Figure 18. Figures 19, 20, 21, and 22 show the group-internal distributions of short and long forms of TO BE (and TO HAVE + been).

As Figure 18 shows, short forms of the auxiliary TO BE plus V-ing are more common in both datasets from spoken BrNSE than long forms. The actual relative num-

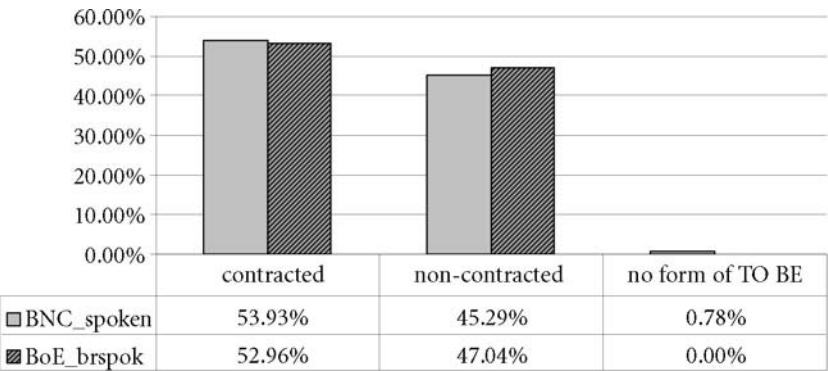


Figure 18. Contracted and non-contracted progressive forms in BNC_spoken and BoE_brspok

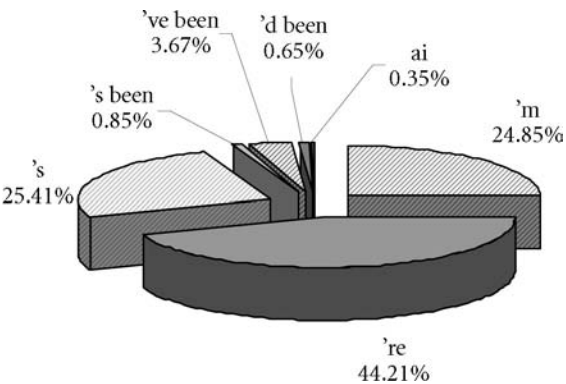


Figure 19. The distribution of different contracted forms (BNC_spoken)

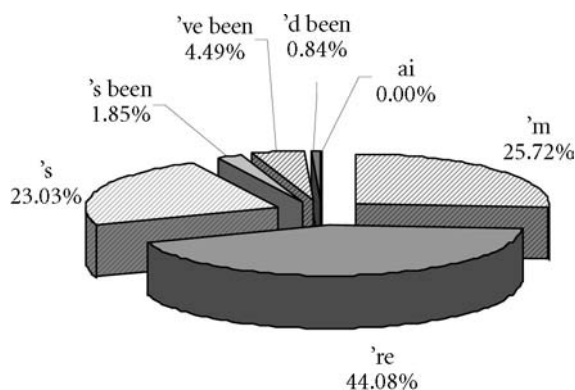


Figure 20. The distribution of different contracted forms (BoE_brspok)

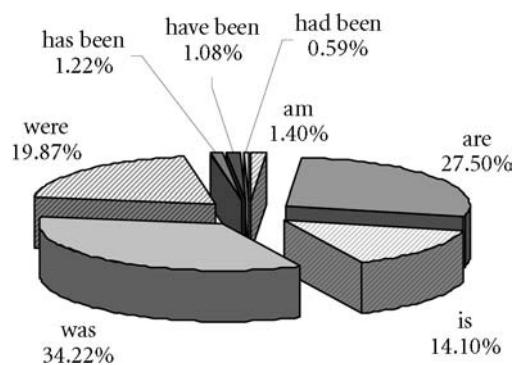


Figure 21. The distribution of different non-contracted forms (BNC_spoken)

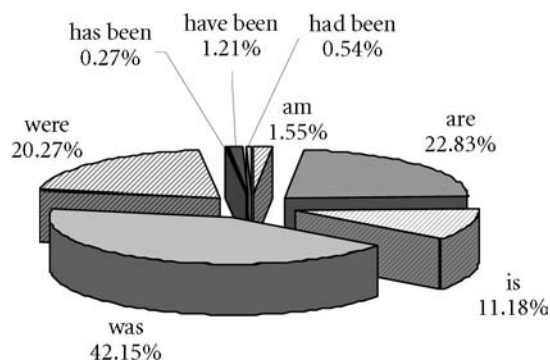


Figure 22. The distribution of different non-contracted forms (BoE_brspok)

bers of occurrence are again strikingly similar in BNC_spoken and BoE_brspok. The BNC_spoken dataset contains a few concordance lines (49 in 6,311) in which the auxiliary is missing, a finding that can be seen as another indicator of the more fragmentary and colloquial nature of BNC_spoken texts as compared to the BoE_brspok material in which there are no such occurrences of progressives without a form of TO BE.

Figures 19 to 22 illustrate largely comparable results of long form and short form distributions in the two spoken BrNSE datasets. There are some smaller differences in the cases of *'s*, *'m*, and *is*. However, the only significant differences according to the chi-square measure relate to the forms *was* and *are*. While the relative frequency of *was* in BoE_brspok is with 42.15% much higher than in the BNC_spoken data (34.22%), the long form *are* occurs relatively more often in BNC_spoken (27.50%) than in BoE_brspok (22.83%). The order of frequency of the different individual contracted and non-contracted forms, however, is exactly the same in both datasets. Despite some group-internal differences between BNC_spoken and BoE_brspok, the shares of contracted (53.93% and 52.96%) and non-contracted progressives (45.29% and 47.04%) in the two corpora are on the whole certainly nicely comparable and hence hint at a high degree of representativeness of the underlying corpora.

Assuming this high degree of representativeness of the data, it is rather puzzling to see that other researchers have arrived at very different results concerning the distribution of long and short forms. Biber et al., for instance, give a number of c. 75% for the share of TO BE contraction in their conversation corpus (cf. 1999:1129), and Mindt's counts even amount to c. 88% (cf. 2000:255). A possible explanation for the divergence between Biber et al.'s and my results is the LGSWE authors's inclusion of all kinds of TO BE contractions in their figure. This means that their 75% not only relate to contracted progressives of the *'m telling* kind but also to copula and passive constructions. It remains unclear why Mindt found a so much higher share of short forms of TO BE in spoken corpus data than Biber et al. and I did. As the author does not make explicit which corpus (or corpora) his result is based on, it is rather difficult to speculate about possible data-inherent origins for the different findings.

4.3.3 Progressives and subjects

The next context feature under examination is the co-selection of progressive forms and their subjects. As part of the corpus-driven pattern approach, it will be investigated which subjects most typically combine with progressives in spoken BrNSE.

Similar to the findings discussed in the previous sections, the results for subjects of progressives are very similar in both corpora. With relative frequencies of 76.06 and 79.25 per cent, personal pronouns (*I, you, he/she/it, we, they*) form by far the most common group of subjects in the analysed datasets from BNC_spoken and BoE_brspok (cf. Figure 24). It can be seen in Figure 23 that within this group the first person pronoun *I* is notably frequent, followed (in order of frequency) by *you* (singular and plural), *we*, and *they*. Among the third person singular pronouns (*he, she, it*), *he* is the most common one. Considering the general male bias in the English lan-

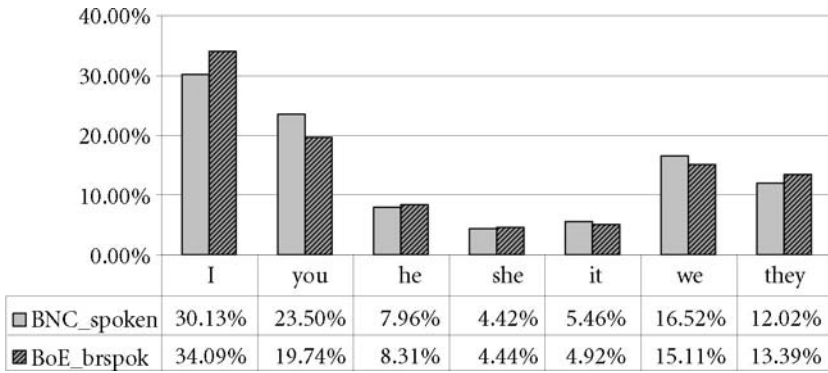


Figure 23. The distribution of personal pronouns as subjects of progressives (BNC_spoken and BoE_brspok)

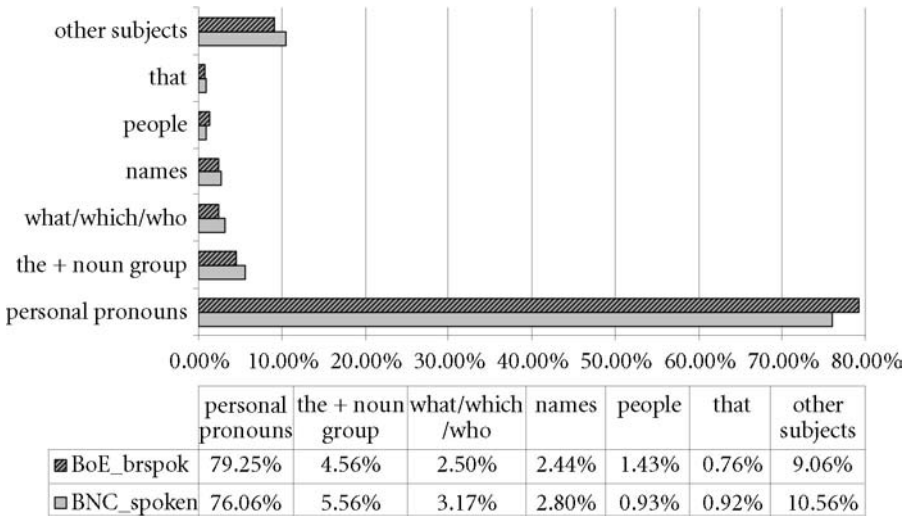


Figure 24. The distribution of subjects in progressive constructions (BNC_spoken and BoE_brspok)

guage (cf. e.g. Biber et al., Section 4.7), it is not surprising that *she* occurs much less frequently as the subject than its male counterpart, and even slightly less frequently than the gender-neutral pronoun *it*.

Other (non-pronoun) subjects are comparatively rare in combination with progressives. The frequencies of some of the relatively more common subjects in the non-pronoun group are displayed in Figure 24.

One of the more common types of subjects in both datasets is a noun (or noun group) preceded by *the*, such as *the students* or *the county council*. There are also many repeated occurrences in BNC_spoken and BoE_brspok of progressives with the sub-

jects *what/which/who*, *people*, and *that*. Names of individuals, e.g. *Diane*, *David*, or in BoE_brspok the person indicators *MX* and *FX*, are used frequently too. Examples for these significant co-selection patterns of subjects and progressives are given in (3) to (6) below.

- (3) You better get up the stairs and tell Johnny *what* 's happening. (BNC_spoken)
- (4) *People* # have been thinking about art # er automatic language analysis for th many # many years (BoE_brspok)
- (5) and *that* really is putting into everyday English (BNC_spoken)
- (6) *FX* is calling from Birmingham (BoE_brspok)

Of course, we also find a large number of other types of subjects (e.g. *her heart*, *nothing*, *anybody*), all of which are, however, rather infrequent.

In the existing empirical literature on the topic there is not a great deal of information on subject-progressive collocation. Mindt only notes that all types of subjects can be found with the progressive and distinguishes between intentional subjects (e.g. *she*, *I*, *the children*) and non-intentional subjects (e.g. *the leaves*, *something*, *it*), of which the former type combines much more frequently with the progressive (c. 86%) than the latter type (c. 14%; cf. Mindt 2000:263). As there is no information in his book on the shares of particularly frequent individual subjects, Mindt's figures cannot be compared with my findings.

In his empirical study on recent diachronic changes in the progressive, Smith (2002:324) also deals with subjects in the present progressive. He subdivides all subjects by person and observes highest frequencies for 3rd person subjects and much lower figures for 1st and 2nd person subjects (cf. also Section 3.5.3). These strong dissimilarities between Smith's results and the BNC_spoken/BoE_brspok findings here can, again, probably mainly be ascribed to different origins of data. While 3rd person subjects seem to be favoured in progressive constructions in *written* English, the most common subject in my *spoken* English data is the 1st person pronoun *I*.

4.3.4 Progressives and objects

After an examination of items preceding the selected -ing forms, i.e. the 100 search-words in my original concordances, I will now turn my attention to the right hand contexts of *accepting* to *writing*. The question here is: "Are there any typical collocations of progressives and words or phrases in object position, i.e. following the -ing form, such as *the course* in *how are you finding the course* (BoE_brspok)?"

Figure 25 displays the shares of some particularly common items to the right of the progressive form in BNC_spoken and BoE_brspok. As the second object box ("object 2") in the database was empty in most of the cases (c. 93%) and no significant patterns became apparent, I will only consider the first object in each of the 9,468 concordance examples.

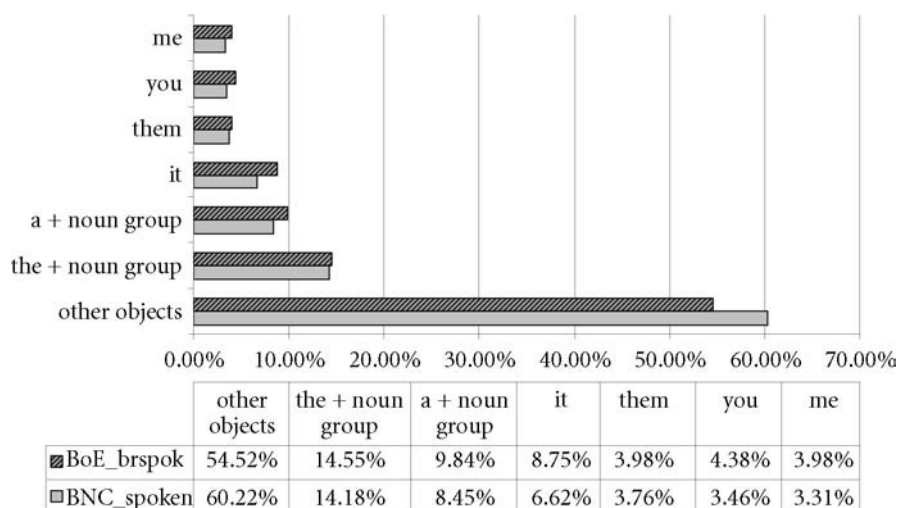


Figure 25. The distribution of objects in progressive constructions (BNC_spoken and BoE_brspok)

As Figure 25 shows, nouns or noun groups introduced by *a* and *the* (e.g. *a phone call*, *the job*) count among the rather frequent objects in the examined datasets. Quite common also are the objects *it*, *them*, *you*, and *me* which occur in examples (7) to (10) below.

- (7) he was finding *it* a bit difficult (BNC_spoken)
- (8) and we're now calling *them* (BoE_brspok)
- (9) he 's ringing *you* instead (BNC_spoken)
- (10) you're not giving *me* what I want (BoE_brspok)

Although there is quite some variation concerning the different less frequently used objects, the figures for these selected common items are roughly comparable in the BNC_spoken and BoE_brspok datasets. Unfortunately, no existing empirical study which includes information on particular progressive-object collocations could be found and used for comparison.

As it may be the case that individual verb forms trigger certain objects and form particular clusters, I consider such combinations of progressive form and object again in Section 4.5.4 when I focus more on the lexical grammar of the items under analysis.

4.3.5 Progressives and prepositions

Another salient feature in the immediate right-hand context of progressive forms are prepositions. The choice of appropriate prepositions following verbs is a particularly important topic in language learning and teaching. Personal experience of assessing

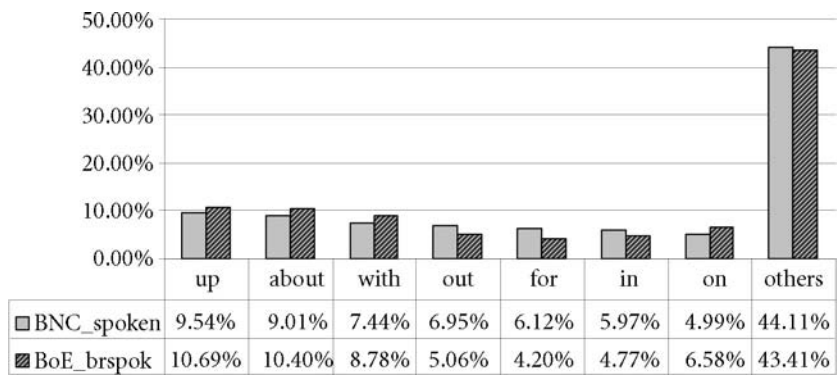


Figure 26. The distribution of prepositions in progressive constructions (1R position; BNC_spoken and BoE_brspok)

the writings of (mainly German) university students shows that learners, even on an advanced level, often have problems choosing the appropriate preposition in a given lexical context. A systematic analysis of prepositions in the context of particular verb forms may thus be a helpful first step to solving these problems.

The first thing to note is that the shares of progressives with and without prepositions in 1R position, i.e. immediately to the right of the searchword, are strikingly similar in the datasets from BNC_spoken and BoE_brspok. In the database, the “preposition” box is filled in 32.37% of the BNC_spoken and in 33.20% of all BoE_brspok examples, which means that 2,043 progressive verb form tokens from BNC_spoken and 1,048 from BoE_brspok are followed by a preposition. As for the different prepositions that occur frequently in these concordance lines, Figure 26 illustrates the distribution of the more common types that were found in 1R position of the -ing form tokens from the two corpora.

As can be seen in Figure 26, *up*, *about*, and *with* are the most frequent prepositions in 1R position in the examined progressives, followed by *out*, *for*, and *in* in BNC_spoken and by *on*, *out*, and *in* in BoE_brspok. At this stage, a more detailed discussion of these results does probably not make much sense as prepositions are highly lexically determined items, which largely depend upon the verb they postmodify, and are thus not typical features of “the progressive” as such but rather of particular verbs or verb forms. Differences between BNC_spoken and BoE_brspok datasets may hence go back to different input frequencies of certain preposition-triggering verb form tokens. I will further examine the connections between common prepositions and individual progressive forms in Section 4.5.5 of this study.

4.3.6 Progressives and negation

In Mindt’s *An Empirical Grammar of the English Verb System*, we find the following general statements on progressives and negation:

The progressive occurs mainly in affirmative contexts (c. 93%). Negative contexts are considerably less frequent (c. 7%). (Mindt 2000:263)

These figures go back to a meta-corpus of spoken and written sources (in undefined proportions). In the entirely written LOB and FLOB corpora Smith (2002:325) found lower shares of clause negation in the PresProg with values of 4.4 and 4.0 per cent. His numbers hint at another difference between speech and writing: apparently, progressives in negative contexts are more common in spoken than in written data, a finding which is supported by Biber et al.'s observations concerning the overall distribution of negation in English (cf. 1999:159).

As we see in Figure 27, the results from the BNC_spoken and BoE_brspok data analysis largely support Mindt's but deviate from Smith's findings. Of the 6,311 examples from the BNC, 542 (i.e. 8.59%) contain a negation while 252 out of 3,157 BoE_brspok progressives (i.e. 7.98%) are negated. Again, the relative frequencies based on the chosen spoken BrNSE datasets are very similar and hence imply a comparability of BNC_spoken and BoE_brspok data. Typical examples of negated and non-negated progressives from both corpora are presented in (11) to (14).

- (11) No, I 'm *not* listening at the moment.(BNC_spoken)
- (12) I 'm listening to the radio.(BNC_spoken)
- (13) I don't think I'd bother because you know it's it's *not* # worrying me.
(BoE_brspok)
- (14) And this is what's worrying me. (BoE_brspok)

I will come back to the phenomenon of negation in Section 4.5.6 below and examine whether the average values illustrated in Figure 27 are valid for all, or at least for the majority of progressive verb form types.

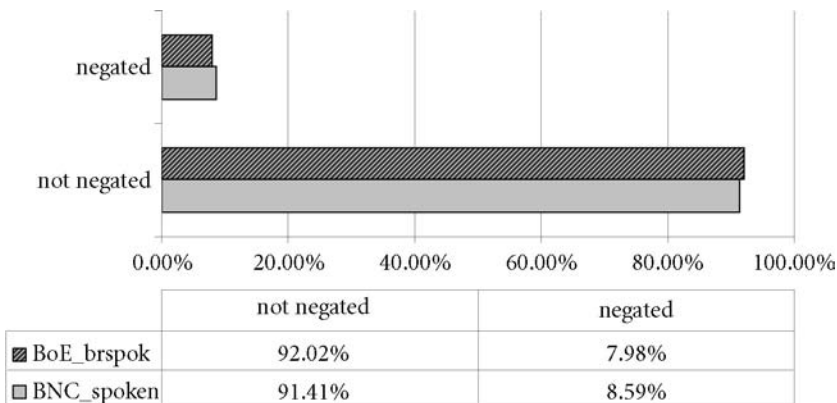


Figure 27. Shares of negated and non-negated examples in the BNC_spoken and BoE_brspok datasets

4.3.7 Progressives and other lexical-grammatical phenomena

This brief section deals with the occurrence of progressive forms in three particular types of lexical-grammatical constructions: questions, if-clauses, and relative clauses. As I decided not to rely on the corpus file transcripts in the case of question marking, the decisive criterion for labelling an example as a question was not the form (i.e. the presence of a question mark) but the function of the concordance line. Thus the “question” box in the *Access* database was ticked in example (15) although there is no question mark in the transcript.

(15) Do you think that the grammar that you’re *learning* is # helpful (BoE_brspok)

The shares of questions, if-clauses, and relative clauses, found in the progressive datasets from BNC_spoken and BoE_brspok are visualised in Figure 28.

Of all progressives from BNC_spoken and BoE_brspok, 11.30 respectively 10.23 per cent function as questions. These results are rather different from Mindt’s. Mindt (2000:263) refers to the rarity of interrogative contexts (c. 5%) and states that of the progressives he analysed, “[a]bout 95% ... have declarative contexts.” Again, the deviating figures are probably due to register differences. Questions are generally a more common feature of speech than of writing (cf. Biber et al. 1999:203, 211), which may indicate that Mindt’s percentages go back to a corpus of mainly written material.

For if-clauses and relative clauses the percentages in the progressive datasets are much lower than for questions. Only 4.44% (BNC_spoken) and 4.53% (BoE_brspok) of the examined progressives fall in the first-mentioned category. In the great majority of the cases (roughly 90% in both corpora), the progressive form occurs in the if-part of the clause (as in (16)), and only very rarely in the main part (see (17)). These findings are largely in accordance with Mindt’s results (cf. 2000:265).

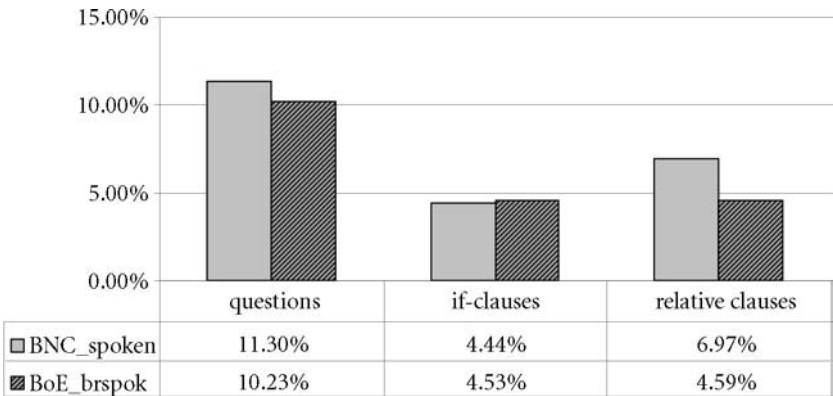


Figure 28. Shares of questions, if-clauses, and relative clauses in the BNC_spoken and BoE_brspok datasets

- (16) But if I *am picking* it up incorrectly there is # only one other option for me to explore (BoE_brspok)
- (17) if I don't do it then I'm not *providing* # a quality service (BoE_brspok)

The shares of relative clauses in which progressives feature are significantly lower in my spoken BrNSE data (4.59% and 6.97%) than in Smith's (2002: 355) written BrNSE corpus material. Smith found percentages of 14.5 and 11.4 per cent in LOB and FLOB – yet another important difference between speech and writing (cf. also Biber et al.'s 1999: 609–611 findings on relativisers across registers). Typical patterns in the co-selection of questions, if-clauses, relative clauses and progressive forms of particular verbs will be investigated in Section 4.5.7.

4.3.8 Adverbial specification

Degree of specification. As has been discussed in Section 3.5.3, most of the previous empirical studies on the progressive deal with the collocation of progressives and adverbials. However, different researchers found rather deviating shares of adverbial specification, based on not easily comparable types of spoken and/or written data. To briefly review some important findings, Ota (1963: 14) in an early empirical study gives a number of 13.4% for the share of temporal adverbial modification in his mainly spoken American English data, while Scheffer's (1975: 55) figure, based on a British English novel corpus, is with 21.2% much higher. Other authors who looked at (mainly temporal) adverbials in the context of progressive forms do not explicitly determine the actual degree of adverbial specification, but only refer, more or less vaguely (usually without giving percentages), to significant co-occurrences between progressives and individual adverbials of time (cf. Section 3.5.3).

Compared to Ota's and even Scheffer's figures, the numbers which go back to my BNC_spoken and BoE_brspok data are considerably higher. In the set of concordance lines from BNC_spoken, I found a share of 23.63% of temporal adverbial specification. If we also take into account other adverbials, e.g. place adverbials like *here* and *there*, the percentage amounts to 29.96%. Of the 3,157 examples from BoE_brspok, 27.87 per cent contain a time adverbial and altogether 39.53% of the progressives are adverbially specified. Figure 29 displays the shares of different types of adverbial specification, including occurring combinations (e.g. of time and place adverbials), in BNC_spoken and BoE_brspok.

Time adverbials, such as *still*, *now*, or *at the moment*, represent the most frequent type of adverbial in the immediate lexical-grammatical context of progressives from both corpora (23.63% and 27.87%) while adverbials of place, e.g. *here*, or *in Nottingham*, are relatively rare (1.41% and 6.30%).

In addition to these two types, a number of other adverbials, such as *actually*, or *whether*, tend to collocate with progressive forms. The percentages determined for this "mixed bag" of adverbials are 6.61% (BNC_spoken) and 9.82% (BoE_brspok). As the diagram shows, combinations of adverbials are not very common. Sometimes,

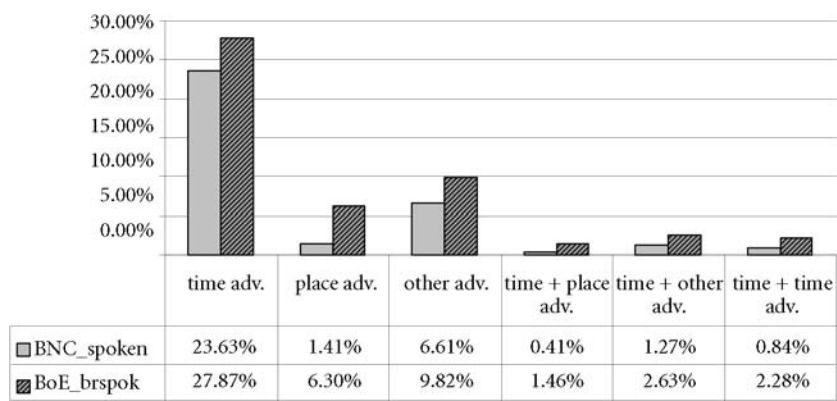


Figure 29. Shares of adverbial specification of progressives in the BNC_spoken and BoE_brspok datasets

however, concordance lines contain a place or other adverbial in addition to a time adverbial (cf. examples (18) and (19)).

- (18) And er he said it had been standing *there for quite a while*. (BNC_spoken)
- (19) I wonder *if* FX's *still* having lessons (BoE_brspok)

For all analysed types of adverbial modification of progressives, higher shares were found in BoE_brspok than in BNC_spoken data. These differences between the two datasets are surprising, given that all other context features under analysis yielded largely comparable and often even strikingly similar results. The lower number of adverbials in BNC_spoken examples appears to go hand in hand with the already noted fragmentary nature of many concordance lines from this corpus. BoE_brspok examples contain fewer instances of hesitation and fewer repetitions and are thus more lexically-dense, which, from a purely statistical perspective, may mean more adverbials per concordance line of 200 characters.

In any case, the corpus results indicate that other researchers’s findings concerning the adverbial modification of progressives are probably not fully representative. My larger scale study shows that progressives are more frequently modified by adverbials than some early small-scale empirical analyses convey. In the next section it will be investigated which individual adverbials typically occur together with progressive forms in spoken English.

Most frequent adverbials. I will now take a closer look at the different types of adverbials that collocate with progressives and at the degree of lexical variation within the three groups (time/place/other adverbials).

The distribution of the most frequent time adverbials in BNC_spoken and BoE_brspok is illustrated in Figure 30. Some typical examples of temporally modified progressives from the two corpora are given in (20) to (23) below.

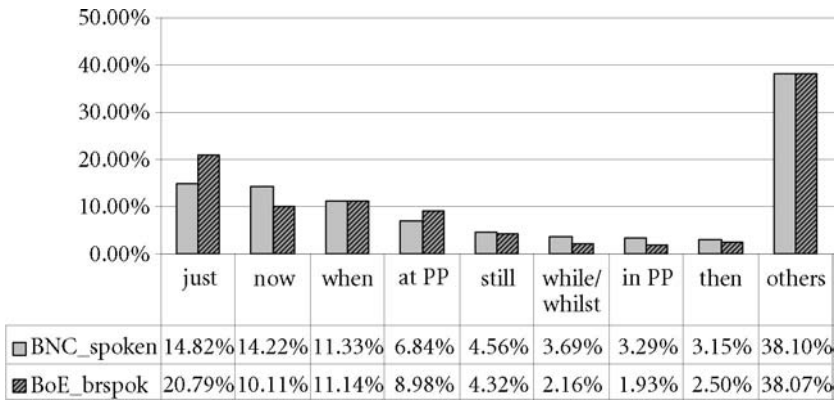


Figure 30. The distribution of frequent time adverbials in the BNC_spoken and BoE_brspok datasets

- (20) I think I can do better that I 'm *now* seeing detail that I never saw before with the naked eye. (BNC_spoken)
- (21) I'm seeing a member *at the moment* female member head of music (BoE_brspok)
- (22) Yes people were *still* using these sort of light fittings? (BNC_spoken)
- (23) Well what can you do while you're watching *while* # you're listening to the radio. (BoE_brspok)

As Figure 30 displays, the eight most common time adverbials in the progressive datasets from the two corpora make up about 62% (61.90% in BNC_spoken and 61.93% in BoE_brspok) of all specified cases, i.e. of all progressives that contain a time adverbial. This hints at some significant collocations between progressive forms and the most frequently used types *just*, *now*, *when*, *at* in a prepositional phrase (*at* PP, e.g. *at the moment*), *still*, *while/whilst*, *in* PP (e.g. *in the morning*), and *then*, but also leaves some room for repeated occurrences of other time adverbials in the context of progressives.⁶⁵ Though, if we go on calculating cumulative frequencies of common time adverbials, we find that the top 20 items in our list account for 81.69% (BNC_spoken) and 81.70% (BoE_brspok) of all temporally specified concordance examples in the respective datasets. Knowledge of a relatively small group of 20 time adverbials hence suffices to produce and understand a large share of progressives in spoken English.

On a similar note, Mindt (2000:265) states that *still*, *now*, *also*, *already*, *just*, and *always* cover about 65% of all cases of progressive/time-adverbial combinations in his corpora. However, his list of the six most common adverbials is somewhat deviant from my top-six ranking (*just*, *now*, *when*, *at* PP, *still*, *while/whilst*), probably because of different types of underlying corpus data. On the whole, my findings and analy-

ses on progressive-adverbial collocations do not contradict earlier empirical studies (e.g. by Ota 1963; Scheffer 1975; Virtanen 1997; Mindt 2000, cf. Section 3.5.3) that also list significant “time-indicators” in progressive constructions, but go one step further in providing detailed numerical evidence from one large corpus which are then supported by figures from a second corpus of a similar type.

Significant patterns could also be found in the co-occurrence of progressive forms and adverbials in the “place adverbials” and “other adverbials” group. Figures 31 and 32 display the distribution of typical place and other adverbials in the BNC_spoken and BoE_brspok datasets. Some examples from the two corpora are given below in (24) to (27).

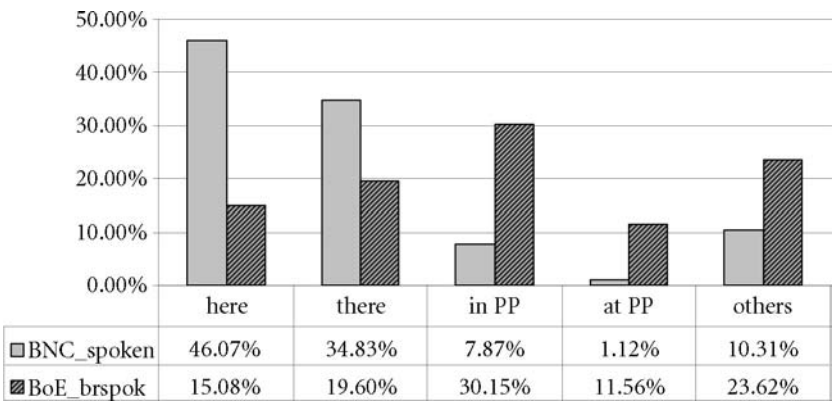


Figure 31. The distribution of place adverbials in the BNC_spoken and BoE_brspok datasets

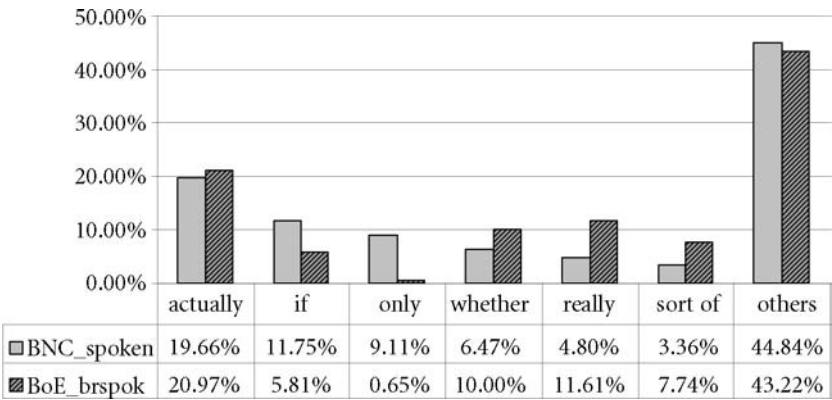


Figure 32. The distribution of other adverbials in the BNC_spoken and BoE_brspok datasets

- (24) *in America* now they're selling a new coffee did # anybody hear about that (BoE_brspok)
- (25) I meant in fact to put them in my bag when I was coming *here* and I forgot. (BNC_spoken)
- (26) We are *actually* putting the clock forward not back. (BNC_spoken)
- (27) it # wasn't *really* bothering me I was quite happy to leave it. (BoE_brspok)

As Figure 31 shows, there is more intra-group variation concerning place adverbials in the BoE_brspok than in the BNC_spoken datasets. Only 10.13% of the specified BNC_spoken examples fall into the “others” group. In both corpora, though, *here* and *there* count among the most frequent adverbials of place in progressive constructions. Even more common place adverbials (only in BoE_brspok concordance lines) are prepositional phrases introduced by *in*, such as *in the street* or *in Birmingham*. Prepositional phrases with *at* (e.g. *at home*, *at Wimbledon*) occur repeatedly in BoE_brspok but not in BNC_spoken. In this case, however, the absolute numbers of occurrence are rather small and do not allow statements about significant differences.

Among the items in the “other adverbials” group, *actually*, non-conditional *if* (in the “whether” sense), *only*, *whether*, *sort of*, and *really* are most frequent. Together, these six types make up 55.16% (BNC_spoken) and 56.78% (BoE_brspok) of all “other adverbial” tokens. This means that, although some of the individual shares of progressive/adverbial collocation in the two datasets differ (cf. Figure 32), the overall findings are certainly comparable and a small core set of typical adverbials could be determined. The relationship between progressive forms of particular verbs and selected adverbials will be further explored in Section 4.5.8 below.

4.3.9 Summary of the findings [spoken English – contexts]

A number of interesting observations with respect to significant progressive collocations could be made in the careful analysis of the lexical-grammatical contexts of more than nine thousand concordance examples from spoken BrNSE.

To briefly sum up some key findings, by far the most common progressive form in our data is the PresProg, followed by the PastProg (cf. 4.3.1). Perfect tense forms are very rare. Within the tense form groups, highest frequency shares go to the patterns *'m V-ing*, *are V-ing*, *'re V-ing*, *'s V-ing*, *was V-ing*, and *were V-ing*. Contracted realisations of progressives, e.g. *'re selling*, are more common than non-contracted forms, such as *are selling* (cf. 4.3.2). Among the subjects of the analysed corpus examples, personal pronouns (particularly *I* and *you*) are most frequent by far (cf. 4.3.3). Other typical subjects of progressive verb forms in spoken BrNSE include noun phrases introduced by *the* (e.g. *the government*), *people*, and names of people (e.g. *Anne*, *Mr Laycock*). There is quite some variation concerning progressive objects in the two datasets (cf. 4.3.4). Some of the more common items in object position of progressives are *the/a+ noun* group (e.g. *a young lady*, *the point*), *it*, *them*, *you*, and *me*. Approximately one

third of the progressive verb form tokens in the two datasets are directly followed by a preposition (cf. 4.3.5). Most frequent are the prepositions *up*, *about*, *with*, *out*, *for*, *in*, *to*, and *on*. These eight items make up about 66% of all modified examples. The progressive form is negated in roughly 8% of all examined corpus examples (cf. 4.3.6). A bit more than 10% of our BNC_spoken and BoE_brspok concordance lines are (or contain) questions. The shares of if-clauses and relative clauses lie between 4.4 and 6.9 per cent (cf. 4.3.7). A modification with time adverbials was found in roughly one fourth of the BNC_spoken and in a bit less than one third of the BoE_brspok examples (cf. 4.3.8). Particularly typical are collocations of progressive forms and the adverbials *just*, *now*, *when*, *at PP* (e.g. *at the moment*), and *still*. Other types of adverbials occur much less frequently in the context of progressives.

With the exception of some frequencies of individual adverbials and shares of less common perfect progressive forms, the contextual examination of progressives has on the whole lead to very similar results for the two analysed datasets of concordance lines from BNC_spoken and BoE_brspok. I shall get back to further examining selected context features and their relationships with individual progressive forms in parts of Section 4.5. As previously stated, the findings obtained for progressives and their contexts in spoken English will later on (in Ch. 6) be compared with the contextual information retrieved from our EFL textbook corpus (cf. 5.5 and 5.7).

4.4 The use of progressives in spoken English (II) – functions

The analysis of the functions of progressives in spoken English was probably the most demanding part of the whole corpus-driven investigation. The aim was to find out, firstly, what was expressed by each of the 9,468 progressive concordance lines from BNC_spoken and BoE_brspok, i.e. which functions the progressive forms served in particular utterances and, secondly, how these functions could best be described in a systematic way.

This difficult process of function detection and description was not guided by any particular theory or previous analyses of progressives and progressive aspect, but solely driven by the collected corpus data, i.e. by copious instances of progressives in context. Of course, when I approached the data, my mind was not a *tabula rasa* and I was naturally familiar with what some grammars and theoretical studies say about the progressive, including the meaning labels used by several authors. The reader will hence find some familiar concepts in my account, which are, however, combined and partly defined in new ways. I generally believe that there is no need to reinvent the wheel, but there might be a need to develop new and better versions of it, in this case versions which respond to new types of data and which incorporate new developments and advances in linguistic analysis or, to stick to the wheel-metaphor, which adjust to new types of roads. Maybe, after all, some kinds of wheels have to be reinvented time and again before we eventually get them entirely right. Energetic corpus work can perhaps prevent wheel-development from a standstill.

The results in this chapter are based on all BNC_spoken and BoE_brspok examples taken together and treated as two representative groups of progressive form tokens. Section 4.5 will explore whether it makes sense to treat the progressive tokens of all different types (in this case 100) as a group or whether the use of progressive forms, including their typical contexts and functions, is strongly lexically determined. In the latter case, the existence of a “purely grammatical progressive”, i.e. a progressive in the sense of an empty slot-and-filler model in which each verb would behave in exactly the same way, might be seriously questioned.

4.4.1 Time reference

The functional analysis of progressive forms in spoken BrNSE was carried out step-by-step in the presence of large amounts of corpus data. A first step in the examination of the functions of progressives was the analysis of their time reference. The aim was to decide for each of the 9,468 concordance examples from BNC_spoken and BoE_brspok whether they referred to something happening in the past, present, or future.

This time reference attribution was relatively straightforward in the case of past actions or events (cf. examples (28) and (29)).

- (28) and I probably make more money today than I ever did when I *was selling* a lot of records as a teenage idol. (BNC_spoken)
- (29) as you’ve heard I think er J K’s *been playing* it on his show in their remix version of Whole Lot of Love by Led Zeppelin (BoE_brspok)

Also, there was a large number of progressives with clear future or present time reference, such as (30) and (31).

- (30) the football match is next week when England *are playing* (BoE_brspok)
- (31) How *are you feeling* now? (BNC_spoken)

However, in many cases, I found it hard, if not impossible, to decide whether the verb construction in question related to the present or the future, as for instance in examples (32) and (33) below. Similar problems with the future/present time reference attribution and meaning overlaps between these two categories were also noted by Mair and Hundt (1995:116) and by Smith (2002:324). It was often the case that such “indeterminate” progressive concordance lines expressed something that was valid not only at the moment of speaking but also in general for situations yet to come. This meaning feature of “general validity” will be further dealt with in Section 4.4.5.

- (32) Because that ’s what it ’s all about. Yes. And when you *’re listening* you ’re watching for body language as well. (BNC_spoken)
- (33) *Are* people not *bringing* their cars in for servicing and repairs anyway (BoE_brspok)

Hence, the first meaning feature box “time reference” contained four feature values to choose from: “past”, “present”, “future”, and “present/future (indeterminate)”. I will now deal with the overall distribution of the BNC_spoken and BoE_brspok progressives across these four categories and, in a second analytic step, discuss the form-function relationship between different types of progressives (PresProg, PastProg, PresPerfProg, PastPerfProg) and time references.

Distribution. To find out which shares of progressives refer to actions or events in the past, present, or future, the BNC_spoken and BoE_brspok concordance lines in the database were filtered according to the possible values in the “time reference” feature box. Figure 33 illustrates the relative frequencies of our four values (“past”, “present”, “future”, “present/future (indeterminate)”) in the two datasets.

As we can see in Figure 33, the highest share of progressives expresses a present time reference. The relative numbers are almost identical for BNC_spoken (38.09%) and BoE_brspok (38.07%). A bit less than one third of the concordance lines from the BoE (31.86%) and 27.36% of the analysed BNC examples refer to an event or action in the past. Similar but altogether lower percentages go to the future reference group (18.55% in BNC_spoken and 15.68% in BoE_brspok) and to the present/future group (15.99% in BNC_spoken and 14.38% in BoE_brspok). If we now assume that of the indeterminate “present/future time reference” cases, most of which could be valid for both present and future actions and events or are in fact timeless statements (cf. examples (32) and (33) above), half can be assigned to the “present time reference” set and the other half to future time reference, we come to a time reference distribution as displayed in Figure 34 below. The large majority of progressives in BNC_spoken and BoE_brspok (46.09 and 45.26 per cent) would then refer to the present, while roughly the same share of BNC_spoken instances express past and future time reference re-

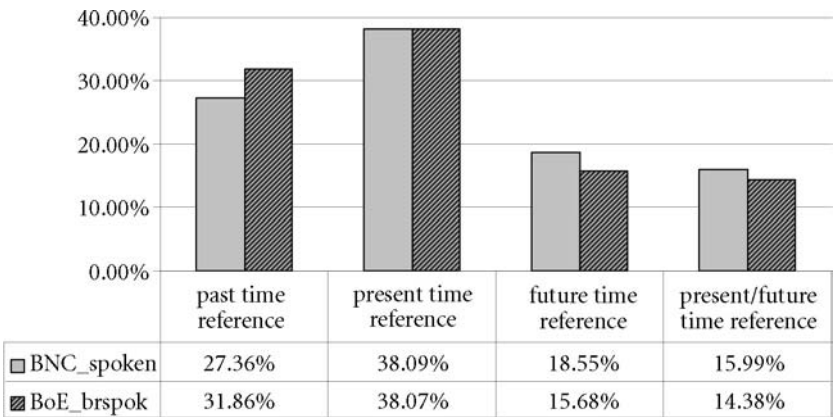


Figure 33. The distribution of time references of progressives in BNC_spoken and BoE_brspok

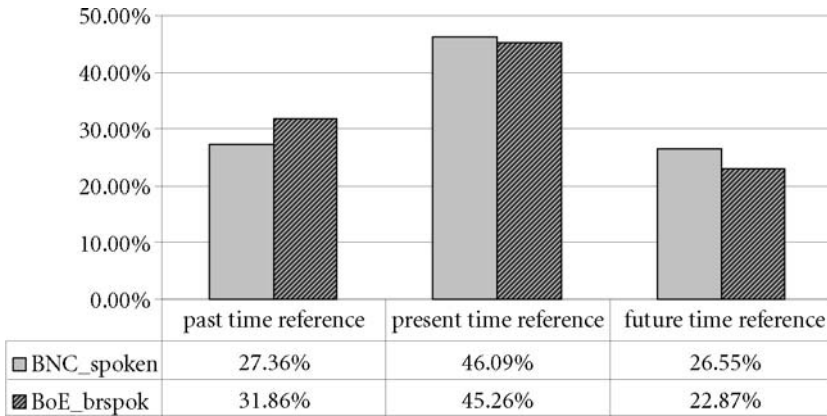


Figure 34. The distribution of time references of progressives in BNC_spoken and BoE_brspok, with indeterminate cases equally distributed (50/50) to “present time reference” and “future time reference”

spectively. For BoE_brspok the percentage of future time expressions is lower and the share of past time references higher than for BNC_spoken, but the frequency order (present → past → future) is the same for both datasets.

If we now compare the values illustrated in Figure 34 with the findings of other scholars on progressives and time reference, we notice a couple of differences. First of all, the shares of future examples are much higher in our datasets than in Scheffer’s, Smith’s, or Mindt’s corpora. While Scheffer (1975:94) assigns future meaning to 18.1% of the progressives in his small corpus of British novels, Smith (2002:324), in his study based on the written British English FLOB and LOB corpora, notes that just 4.5% of the present progressives in LOB and 3.9% in FLOB refer to future events. Mindt’s (2000:253) findings are very similar to Smith’s. He states that in his meta-corpus of spoken and written English “[f]uture time orientation (c. 5%)” is rare. What is puzzling in his account, though, is the relation between this figure (c. 5%) and the percentages he gives for the meanings “prediction” (c. 7%) and “volition/intention” (c. 5%) (Mindt 2000:253).⁶⁶ I consider it difficult to imagine that a prediction or a person’s intention does *not* refer to the future. In part this mismatch between the numbers can be explained by the combinations of prediction and volition/intention in many of Mindt’s (2000:260) examples, but apparently, there are a number of instances in his data which express one of these two meanings and have either past or present time orientation, or are timeless.

Form and function relationship. In addition to the actual time reference distribution, a second point of interest in connection with the temporal orientation of progressives is the relationship between progressive tense forms (i.e. PresProg, PastProg, PresPerfProg, PastPerfProg) and the time span they refer to (past, present, future, present/future). It was to be determined whether, in spoken BrNSE, there is a clear

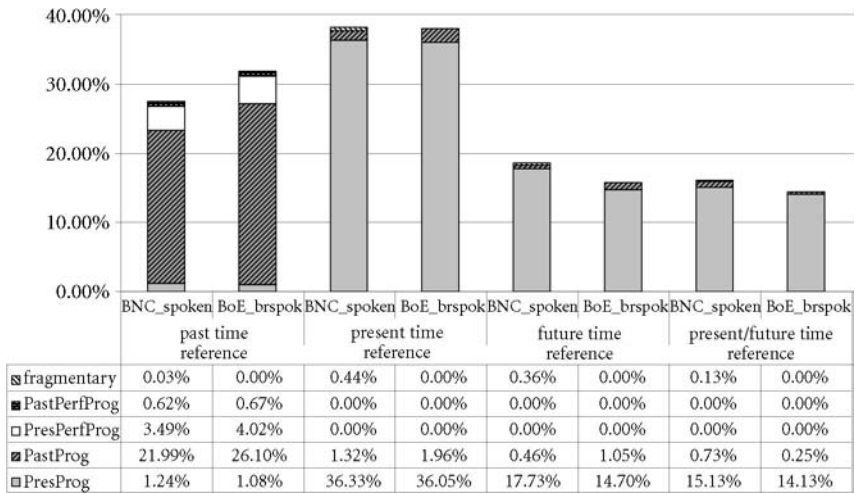


Figure 35. The relationship between time reference and progressive tense form in BNC_spoken and BoE_brspok

connection between these forms and functions. To work out this relation, the progressives in the database were first filtered according to time references and sorted for forms of TO BE, which were then grouped into the respective tense forms (e.g. *was*, *were* → PastProg). Figure 35 serves to illustrate the results of this analytic procedure.

If we look at the two columns representing past time reference of progressives in BNC_spoken and BoE_brspok, we notice that, unsurprisingly, the great majority of examples contains PastProg forms. Only just over one per cent of the BNC_spoken and BoE_brspok concordance lines contain PresProgs with past time orientation as, for instance, shown in examples (34) and (35). Most of these PresProgs with past time orientation have the function of reporting or storytelling. Some examples, however, are difficult to interpret; they sound as if they were taken from colloquial exchanges, maybe among teenagers (cf. (36)), or they are possibly mistranscriptions from the recordings. In fast speech, “we were” may be misunderstood as “we’re” and “there were/are” may get wrongly transcribed as “they’re” (cf. examples (37) and (38)).

- (34) Last night this these, as I said, he ’s *playing* this rabbi Yeah. and he ’s having a punch up with his brother and he keeps letting his brother (BNC_spoken)
- (35) Robinson Crusoe arrived on a desert island with his shipwrecked boat okay. He’s not # *expecting* anything else he’s got to manage with what he’s got today. (BoE_brspok)
- (36) Well I says to him, I says if you were that interested I says while I ’m *watching* I says why do n’t you come and see, I says (BNC_spoken)
- (37) On the board this morning we ’re *going* to have a bash at thinking about some targets (BNC_spoken)

- (38) Was it boring? It was n't boring but it was, just, I mean you know they 're *sitting* five hundred people all at one go and it 's Boring. (BNC_spoken)

The rest of the two “past time reference” columns in Figure 35 is made up of PresPerfProg and PastPerfProg forms (see examples (39) and (40)).

- (39) and erm made them all laugh by by saying that for # years he's *been selling* what he described as 'total crap' to the general public (BoE_brspok)
- (40) And Yes. er it was sold out, you know, and everybody, and people *had been ringing* up thanking them and everything. Oh that 's lovely is n't it? (BNC_spoken)

The relation between form and function is even more significant in the cases of present, future, and “indeterminate” present/future orientation. The six right-hand columns in Figure 35 nicely show that PresProg forms clearly dominate in these three time reference groups. There are only a few PastProg examples (between 0.25% and 1.96%) which have the following functions: A first pattern that emerges is that most PastProg lines with present time reference are from *wondering*-concordances and contain the subject “I”. The function of the progressive form in these examples (see e.g. (41) and (42)) is clearly a softening one, expressing politeness by moving the actual act of wondering further away to the past although it actually happens at the time of speaking. Some other examples with *ringing* (*up*) and *thinking* basically express the same meaning; the PastProg forms make the utterances sound more polite (see e.g. (43) and (44)). Significant also in this context is the frequent collocation with *just* in a large number of the *wondering/ringing/thinking* concordance lines. In these examples, *just* as a hedging or softening device further enhances the politeness effect.

- (41) Sorry, I *was just wondering* if you actually asleep or not. Want a cup of tea? No. (BNC_spoken)
- (42) I *was just wondering* if you want to share it with the rest of # us (BoE_brspok)
- (43) I *was just thinking* <ZF1> is there <ZF0> is there going to be # some kind of report on <ZGY> performance (BoE_brspok)
- (44) What can we do for you # <F09> Well <ZF1> I *was* <ZF0> I *was just ringing* you up <ZF1> to <ZF0> to see if you could tell where I came from in fact. (BoE_brspok)

Past progressive forms with future time reference occur in reporting contexts and in embedded (that-) clauses (cf. (45) to (48)).

- (45) I 'm forty seven. I thought you said she, I thought you *were stopping* having birthdays. (BNC_spoken)
- (46) I asked FX where she *was moving* to (BoE_brspok)
- (47) I thought you *were n't staying* the night though? Well I 'm not. (BNC_spoken)

Table 9. Relations between time references and progressive forms in spoken BrNSE

Time reference	Forms of progressives
past time reference	predominantly PastProg, some Pres/PastPerfProgs, few PresProgs
present time reference	predominantly PresProg, few PastProgs
future time reference	predominantly PresProg, very few PastProgs
present/future time reference	predominantly PresProg, very few PastProgs

(48) at the back of my mind I knew she *was* # *leaving* my disabled grandmother (BoE_brspok)

Another function of such “future-in-the-past” examples is the expression of a hypothetical meaning (cf. e.g. (49) and (50)). Hypotheticalness is also expressed in PastProgs with indeterminate time orientation. In these cases we find a significant collocation with conditional *if* (cf. e.g. (51) and (52)).

- (49) that would set me into a situation where I *was* # *saving* money (BoE_brspok)
- (50) I did n’t know they *were sending* out such a letter, otherwise I would have er, told them in advance (BNC_spoken)
- (51) I would still be me even if I *were* n’t *holding* a book (BNC_spoken)
- (52) Like I do if I *was showing* off. <tc text=pause> <F0X> <ZGY> <F0X> <tc text=laughs> <F0X> All the time. (BoE_brspok)

On the whole we certainly find a clear connection between time reference and forms of progressives in spoken British English, as summarised in Table 9. Some specific “deviating” patterns are, as we could see, strongly lexically determined. Section 4.5 will further look into the functional and contextual patterns around specific lexical items from the list of 100 *-ing* forms.

4.4.2 Two central function features: Continuousness and repeatedness

Having determined the distribution of time references and connections between tense forms and temporal functions in the data, I will now turn to examining the central function or functions of my progressives. The central question is “What do progressive forms in general, independent of their time orientation, mainly serve to express?”

The discovery process of central function features of progressives was entirely data-driven, meaning that, instead of existing descriptions in grammars and theoretical studies, concordances were consulted with the aim of deriving functions directly from the corpus evidence. An apparently very common function feature of progressives that arose from an initial analysis of several hundred randomly selected concordance lines from the two corpora was the feature “continuousness”. Of course, this finding did not come as a big surprise; after all, the term “continuous” is often used as an alternative expression to “progressive” in grammatical reference works (cf. e.g.

Cobuild 1990; Eastwood 1994; Greenbaum 1996; Parrot 2000). The analysed corpus examples described actions or events that happened without interruption or without an intervening time interval.

At first it seemed that almost all instances referred to single, continuous events or actions (as in examples (53) and (54)). However, the more concordances and individual examples I looked at, the more deviations from this initially spotted functional pattern (“single, continuous event”) emerged. A large number of concordance lines was found to refer to repeatedly occurring actions or events. These repeated actions were mostly continuous, as in examples (55) and (56), but sometimes they were also non-continuous, as in example (57).

- (53) the strange black things that some of you *are holding* in your hands are called riders, and these are end leaves for the storage binder (BNC_spoken)
- (54) Now the sequence I’m *showing* you is er Chris Mullen’s # first interview with MX. (BoE_brspok)
- (55) but it also is much easier for the form teacher to know who ’s not *bringing* in homework (BNC_spoken)
- (56) I’ve lost half # a stone but it was only like er because I *was eating* too much before anyway. (BoE_brspok)
- (57) Well I’ve *been buying* all the mountain bike magazines lately # just to get to know what’s going on. (BoE_brspok)

As “continuousness” and “repeatedness” were felt to be very appropriate labels to account for the data, possibly more appropriate than some other terms, such as “limited duration”, “incompletion”, or “imperfectivity”, which are traditionally used to describe progressive functions, these features were chosen and defined as central features to help capture the basic function(s) of progressives. The two central features and their realisations in the corpus data (continuous vs. non-continuous and repeated vs. non-repeated) are dealt with below in this section. All possible combinations of the features and their frequencies of occurrence in spoken BrNSE are the subject of discussion in 4.4.3.

Continuousness. If we look at a BNC concordance of the lexical item “continuous”, we note that the typical collocates, e.g. *activity, development, flow, growth, improvement, movement, process*, refer to something “procedural” or to ongoing actions. The collocates create a context of un-interruption and extension over a certain period of time. In my assessment of the BNC_spoken and BoE_brspok corpus data, I hence used the criteria of “break/interruption” and “extension over a certain time span” to decide whether a progressive construction expressed continuousness or not.

The reader may be confused at this point and wonder how it can be that progressives do *not* express continuousness, when this very construction is sometimes labelled “continuous” instead of “progressive”. I am sorry for this confusion, but according to my definition of continuousness, i.e. extending over a certain time span without in-

interruption, a number of progressives in the database refer to non-continuous actions or events, as the action contained in the verb form is a very short-termed, punctiform one, rather than one that extends over a certain time span. In examples (58) to (60), for instance, I perceive the acts of *asking*, *picking (up)*, and *suggesting* as non-continuous, as they do not meet the criterion of temporal extension but refer to short-term actions. What the progressive forms convey in these examples is rather a sense of softening or downtoning, or, in the case of *picking*, a special emphasis on the problematic fact that no one is willing to pick up the phone (cf. also the sections on politeness/softening and emphasis/attitude in 4.4.5).

- (58) I say real progress has been made but today I *am asking* you to think about the next step (BNC_spoken)
- (59) even when # it's hard and no-one's *picking* up the phone or it's engaged (BoE_brspok)
- (60) I'm *suggesting* we report him for being drunk # on duty. (BoE_brspok)

Figure 36 demonstrates that, although the great majority of corpus examples express continuousness, about 18 per cent of the progressives in the two datasets refer to non-continuous actions or events. Table 10 gives a list of verbs that often occur in such non-

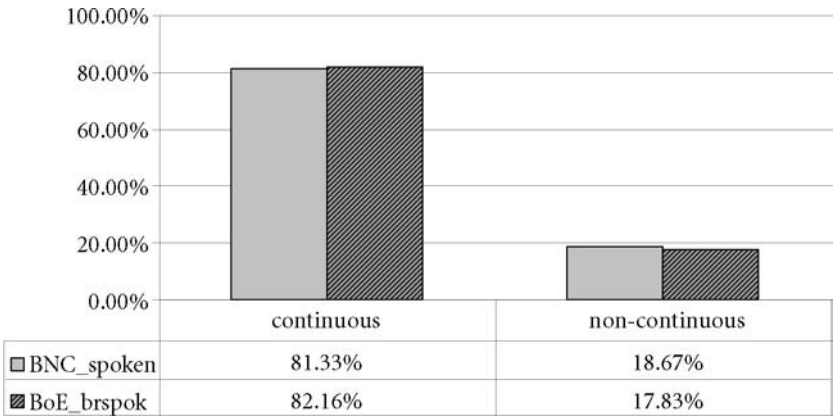


Figure 36. The distribution of progressives referring to continuous and non-continuous actions or events in BNC_spoken and BoE_brspok

Table 10. Verb forms frequently occurring in “non-continuous” contexts in BNC_spoken and BoE_brspok

asking	costing	picking	sending	stopping
buying	giving	putting	showing	suggesting
calling	leaving	ringing	spending	taking
checking	paying	saving	starting	winning

continuous contexts. On the back of these findings, the use of the terms “continuous aspect” or “the continuous” might appear to be not entirely appropriate to account for the full range of TO BE + V-ing constructions in spoken English.

Repeatedness. The second function feature that occurred as central in the semantic annotation of the BNC_spoken and BoE_brspok progressives was the feature “repeatedness”. Depending on whether it referred to a single event or described an action occurring more than once, each of the 9,468 analysed concordance examples from spoken English was assigned the feature value “non-repeated” or “repeated”, respectively.

Repeatedness is, for instance, expressed in (61) and (62), whereas (63) and (64) refer to non-repeated single actions.

- (61) she does n’t eat that much, but what she *is eating* i everything ’s sweet. Mm. Chocolates and Is she eating them? Pardon? (BNC_spoken)
- (62) So we were # concerned with things they *were* currently *doing* and things # that they erm could really conceive of doing. (BoE_brspok)
- (63) Oh I see. Pork ’s very nice. What ’s Geoff *eating*? Sausage roll. Oh Geoff, you ’ve only just had your tea! (BNC_spoken)
- (64) One of the Governors a very nice lady came in one morning and # they *were doing* needlework (BoE_brspok)

Among the 100 verb forms in the database those listed in Table 11 are particularly common in contexts in which they express repeatedness (cf. 4.5.10 for further details on connections between central functions and individual lexical items).

The shares of corpus examples with repeated and non-repeated content in the two selected spoken BrNSE corpora are displayed in Figure 37. As the figure shows, repeatedness is expressed in roughly 35 per cent of the 6,311 BNC_spoken examples and in more than 38 per cent of the 3,157 concordance lines from BoE_brspok. The remaining shares of progressives (65.95% in BNC_spoken and 61.36% in BoE_brspok) refer to single/non-repeated events.

In the light of many traditional, and particularly non-empirical, accounts of the progressive and its functions, these findings may appear surprising. In these studies the emphasis is usually on progressive constructions which relate to single, continued events. Comrie (1976: 37–38), for example, regards the indication of a particular contingent situation as the basic meaning of the progressive and treats iterative, i.e. repeated, or habitual uses as specific cases. In several other publications the stress is

Table 11. Verb forms frequently occurring in “repeated” contexts in BNC_spoken and BoE_brspok

becoming	dealing	giving	picking	trying
buying	doing	happening	selling	using
calling	eating	helping	spending	wearing
costing	getting	paying	taking	working

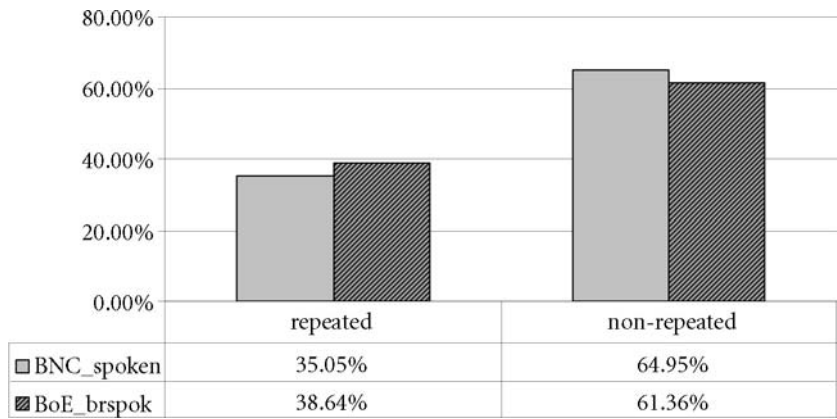


Figure 37. The distribution of progressives referring to repeated and non-repeated actions or events in BNC_spoken and BoE_brspok

also explicitly on duration (cf. Lucko 1995:181; Palmer ²1988:54; Potter ²1975:118; Quirk et al. 1985:197f.; Sammon 2002:29; Scheffer 1975:21; or Virtanen 1997:306, who describes duration as the “main function” of the progressive). Repetition is rarely mentioned as an important feature. Mindt (2000: 256), however, lists “iteration/habit” as one of the four prototypical meanings of progressive forms, but strangely enough this function is assigned to only 12% of the examples in his corpora. This discrepancy between Mindt’s and my own figures may again be due to the different types of data which our research is based on (predominantly written material in Mindt’s case vs. exclusively spoken material in my case) or on deviating definitions of the features “repeatedness” and “iteration”.

Even though repeatedness is usually not treated as an important characteristic of the progressive, my analysis shows that a large proportion of progressives in spoken English does in fact refer to repeatedly occurring actions or events.

4.4.3 One central function or several central functions?

Having specified two central function features of progressives and their distributions in spoken BrNSE, we can now move on to the identification of the basic use or uses of the forms.

To find out whether it is advisable to take a “one-central-function” or a “several-central-functions” approach, massive corpus evidence will again be considered and the possible combinations of the essential features “continuousness” and “repeatedness” will be examined. Altogether, there are four possible feature value combinations which represent four different functions of progressive forms:

- continuous + repeated
- continuous + non-repeated

- non-continuous + repeated
- non-continuous + non-repeated.

Examples (65) to (72) serve to illustrate these functions.

Continuous and repeated action or event:

- (65) you know I mean it 's funny when I 've *been giving* talks on communications
erm one of the things I say to people is erm (BNC_spoken)
- (66) he *was spending* nearly forty hours a week on governors' business you know
(BoE_brspok)

Continuous and non-repeated action or event:

- (67) Yes. We 're, we 're currently *running* about fifteen minutes late. (BNC_spoken)
- (68) And we *are not talking* here about the # needs and desires of a son or a mother
or a business associate (BoE_brspok)

Non-continuous and repeated action or event:

- (69) We 're not *giving* out mortgages at the moment I 'm sorry (BNC_spoken)
- (70) They're just # erm *taking* any staff on really. (BoE_brspok)

Non-continuous and non-repeated action or event:

- (71) Yeah. So you 're not *looking* at the answer. Do n't look at the answer.
(BNC_spoken)
- (72) I'm just *ringing* up to say congratulations to Birmingham City # on making
the final at Wembley <000> at last. (BoE_brspok)

The frequencies of occurrence of the four feature combinations in the BNC_spoken and BoE_brspok datasets are displayed in Figure 38. The figure clearly shows that the two "continuous" functions (continuous + repeated and continuous + non-repeated) occur significantly more often in the corpora than the "non-continuous" meanings. As has been true for most of the corpus findings so far, the individual figures for BNC_spoken and BoE_brspok are very similar. It can thus be said that the illustrated distribution of progressive functions (cf. Figure 38) is likely to be significant of spoken BrNSE in general.

While the non-continuous functions show shares of about 10 per cent or less, the relative frequencies of continuous + repeated and continuous + non-repeated actions or events amount to 26.64%/29.05% and 54.70%/53.12% respectively. This distribution indicates that there are two central functions (CFs) of progressives in spoken English:

- CF1: the expression of **continuous and non-repeated** actions and events (cf. examples (67) and (68)), and
- CF2: the expression of **continuous and repeated** actions and events (cf. (65) and (66)).

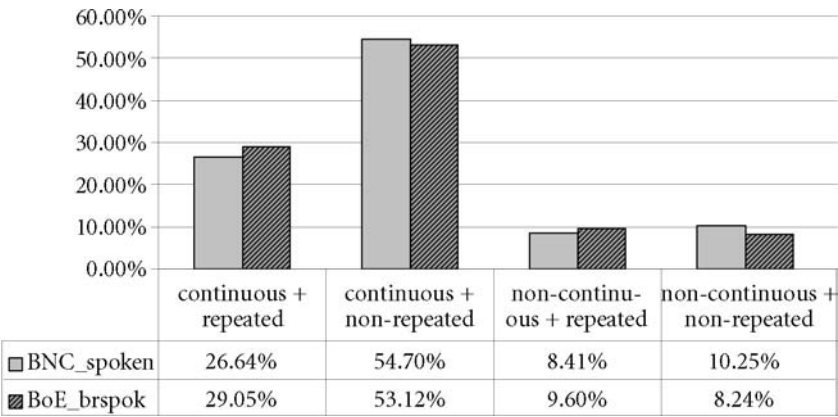


Figure 38. The distribution of central function feature combinations in BNC_spoken and BoE_brspok

Table 12. Verb forms with tokens that express “non-continuousness + repeatedness” in BNC_spoken and BoE_brspok

asking	giving	picking	sending	suggesting
buying	leaving	providing	showing	taking
calling	looking	putting	spending	turning
checking	making	ringing	starting	winning
costing	paying	saving	stopping	

Of the 100 selected verb forms in the database, 97 types showed instances of CF1 (continuous + non-repeated)⁶⁷ and 90 forms showed instances of CF2 (continuous + repeated).⁶⁸ The two functions (CF1 and CF2) can hence be described as basic functions that apply to the great majority of progressive forms, or, in other words, as functions of the progressive in general. On the other hand, the expression of the two less prevalent progressive functions (expression of non-continuous + repeated and non-continuous + non-repeated actions or events) is restricted to a much smaller set of lexical items. Table 12 lists all verb forms (24 out of 100) that occur to express “non-continuousness + repeatedness”. The same 24 verb forms and two additional items from the list of 100 (*meeting, sitting*) serve to refer to non-continuous + non-repeated actions or events. This finding allows us to conclude that, unlike CF1 and CF2, both non-continuous functions (\pm repeated) are apparently strongly lexically determined. They cannot be called functions of the progressive as such, but are rather functions of certain progressive constructions of particular verbs. Section 4.5 will discuss this relationship between individual verb forms and progressive functions and contexts in more detail.

Let us now briefly consider what other scholars have found out about central functions of the progressive. As already discussed in Section 3.5.2, Mindt distinguishes four prototypical meanings of the progressive (plus five other less common ones), the first

three of which are what he calls “progressive meanings” (incompletion, 60%; temporariness, 36%; iteration/habit, 12%). The fourth one, highlighting/prominence (9%), is a “non-progressive meaning” (cf. Mindt 2000:252–253).⁶⁹ It was stated above (cf. 4.4.2) that most researchers stress duration as a or as *the* central meaning of the progressive. Many authors further specify this duration as being temporally limited, i.e. restricted to a certain limited stretch of time (cf. e.g. Quirk et al. 1985:197f.). All this does not represent a contradiction to my empirical findings. An important difference I see, however, lies in the reference to repeated actions or events. In the literature, the expression of repeatedness (or iteration) tends to be treated as one possible function of progressives, though certainly not as a common one. My findings show that a considerable amount of all examined progressives from spoken English in fact express repeatedness. This might suggest that this function feature should move further to the centre of descriptions of English progressives. The following section of this chapter will look into possibly existing connections between central functions and time references.

4.4.4 Central functions and time reference

I will now take a look at the correlation between the two specified central functions (CF1 and CF2) and different types of time reference. The question is “Are there any tendencies for progressives that occur in either CF1 or CF2 to predominantly refer to the past, present, or future?” Answering this question will also tell us something about the status of individual types of time reference. Put differently, if we find that both central functions are found in all possible time references, we may conclude that “future meaning” is not just a separate specific function of the progressive but a temporal dimension, just like the past and present, that can combine with different functions.

Figures 39 and 40 illustrate the connections between central functions and time reference. Both distributions differ to some extent from the general distribution of time references in the whole dataset (cf. Figure 33). The first thing we notice about Figure 39 is the very low share of present/future “indeterminate” cases (3.53% in BNC_spoken and 1.49% in BoE_brspok). Apparently, most of the “continuousness + non-repeatedness” examples could be easily identified as either referring to the past, the present, or the future. Obviously not at all easy to assign to one of these three time reference categories were progressives that express a continuous and repeated action (CF2). Here we find much higher percentages of indeterminate instances (41.52% in BNC_spoken and 33.04% in BoE_brspok), including those referring to generally valid situations or habitual actions (cf. the respective Sections in 4.4.5).

Having significantly higher shares of indeterminate (present/future) examples in the case of CF2 of course implies in turn that the clear past, present, and future reference columns have to be lower than average (cf. the six left-hand columns in Figure 40). As we can see, the percentages of these three types of time reference are altogether lower than for CF1-progressives (Figure 39) and for progressives in general (Figure 33). With respect to the expression of future meaning, we find considerable shares of future concordance lines in both central functions, especially when we take into account that

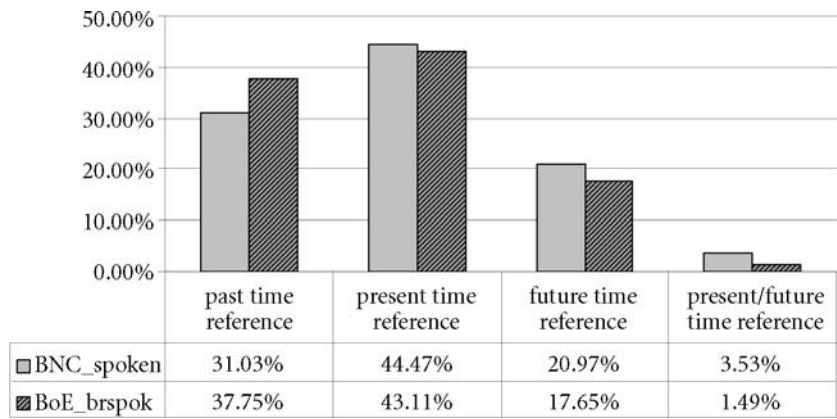


Figure 39. Relations between the central function “continuousness + non-repeatedness” (CF1) and time references in BNC_spoken and BoE_brspok

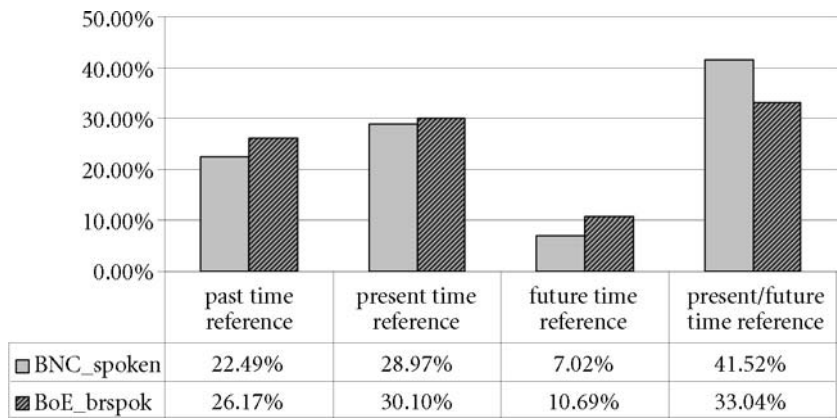


Figure 40. Relations between the central function “continuousness + repeatedness” (CF2) and time references in BNC_spoken and BoE_brspok

most of the indeterminate cases refer to present *and* future actions or events, i.e. to things that are generally valid at the time of speaking and at a later point in time. I would hence say that “future” as such is not a separate function of the progressive, as many scholars state (cf. e.g. Greenbaum 2000; Hirtle & Curat 1986; Palmer ²1988; Scheffer 1975; Smith 2002), but an important temporal feature which combines with both identified central functions. Section 4.4.6 will discuss in how far this is also true for more or less common additional functions of progressives.

4.4.5 Additional functions of the progressive

In addition to helping us find the basic functions of progressives, the examination of several thousand progressive form tokens in context showed that progressives often express more than just “continuousness + repeatedness” or “continuousness + non-repeatedness”. The two above-specified central functions thus cannot suffice to fully capture what progressives actually denote and for which purposes they are typically used in spoken English.

The following seven functions were found to be expressed repeatedly in the analysed concordance lines from BNC_spoken and BoE_brspok:

- general validity
- politeness or softening
- emphasis or attitude
- shock or disbelief
- gradual change and development
- old and new habits
- framing

Figure 41 displays the frequencies of occurrence of the seven additional functions in BNC_spoken and BoE_brspok. The first thing worth noting about Figure 41 is, again, the strong similarity of BNC and BoE results. Both datasets were independently assessed, one after the other, and the frequencies of the additional functions were only retrieved after the annotation of the BNC_spoken and BoE_brspok data had been completed. The fact that the results obtained from the BoE_brspok analysis fully support those derived from the BNC_spoken examples, certainly implies the significance and a generalisability of the findings.

With shares of 22.04 and 22.96 per cent, “general validity” is the most common of the additional functions in both corpora. Also rather frequent are progressives which

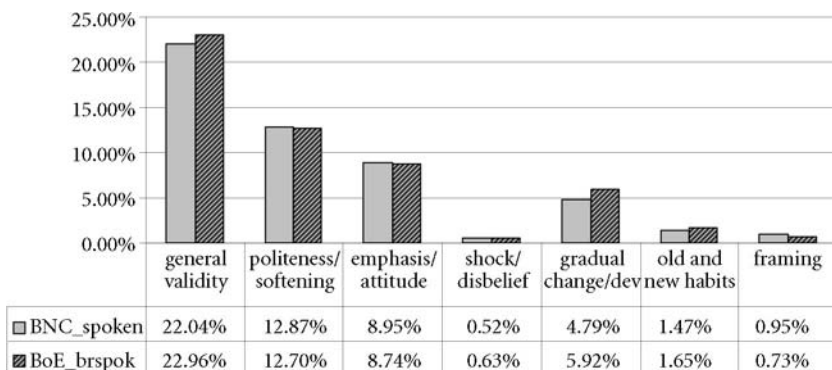


Figure 41. The distribution of additional progressive functions in BNC_spoken and BoE_brspok

convey the meanings of either politeness or softening (12.87 and 12.70 per cent), or of emphasis or a particular, usually negative, attitude (8.95 and 8.74 per cent). A gradual change or development is expressed in 4.79% of the BNC_spoken concordance lines and in 5.92% of the BoE_brspok examples. Comparatively rare are the three functions “old and new habits”, “framing”, and “shock or disbelief”. They amount to shares between 0.52% and 1.65% in the two datasets. The next subsections will in turn discuss and exemplify each of the additional functions of progressives. Relations between individual verbs and additional functions will be touched upon briefly in each section and dealt with more thoroughly in 4.5.11 below.

General validity. The common use of progressives to refer to situations that are generally valid was already mentioned above in connection with progressive time reference (cf. 4.4.1). We observed that, in many cases, it was impossible to assign concordance examples to either the group of progressives with *present* time reference or to the *future* reference group, because these examples expressed something which was valid not only at the moment of speaking but also for situations to come. As the name indicates, the “general validity” function refers to exactly that kind of situations. Something is true or valid in general, or at least for some time, not just for the moment. Examples (73) to (76) illustrate this function.

- (73) Ca n’t help it I I ’m such a give away when things like that *are happening*. (BNC_spoken)
- (74) all these travellers are # paying Poll Tax # <M01> Well I don’t see why they’re not *paying* Poll Tax. (BoE_brspok)
- (75) But in general the people who *are doing* some of those things, a lot of it ’s probably escapism (BNC_spoken)
- (76) if the # man and a woman *is working* it’s still the woman who does most of the # housework. (BoE_brspok)

The situations described here do not refer to a specific present or future event but to repeatedly occurring, general actions, such as paying taxes or doing a job. Significant is the strong collocability of general validity progressives with *if* and *when*. Of the concordance lines which express this function, more than 20% contain conditional *if* or *when* (the average share of *if*-clauses in the whole analysed set of progressives only makes up 4.5%; cf. Section 4.3.7). Verbs which most often express “general validity” are listed in Table 13. However, there do not seem to be any very strong lexical restrictions concerning the use of progressive forms in this function (cf. also Section 4.5.11). Out of 100 forms, 94 were found to refer to something that is generally valid. This means that, like the central functions CF1 and CF2, “general validity” can be called a function of the whole group of progressives or of the progressive as such.

Even though it seems that this additional function is not very much lexically triggered (with respect to verb forms), there are some interesting lexical relations between “general validity” progressives and their subjects and adverbials. Most common in sub-

Table 13. Verb forms frequently occurring in “general validity” contexts in BNC_spoken and BoE_brspok

buying	eating	happening	providing	using
costing	getting	helping	selling	wanting
dealing	giving	paying	spending	wearing
doing	going	picking	trying	working

ject position, apart from the generally frequent personal pronouns (in particular *you*, *they*, *we*, and *I*), is the item *people*. Talking about people (in general) fits, of course, nicely in the context of generally valid situations (cf. examples (77) and (78) below).

(77) but *people are drawing* their pensions and they get their lump sums at fifty (BNC_spoken)

(78) But then there's too much # saving and *people are not spending* are not buying and things are being produced (BoE_brspok)

Among the time adverbials in the concordance lines, *when*, *always*, and *all the time* are most frequent. All three also help to support the sense of general validity, as examples (79) to (81) illustrate.

(79) Kids are all right *when you're buying* things for them. (BoE_brspok)

(80) You know like they 're *always picking* their feathers. (BNC_spoken)

(81) certain people in certain areas in universities *are # carrying* out all these this research and finding out all # different things *all the time* (BoE_brspok)

Despite its frequency in spoken BrNSE, the function of “general validity” is not explicitly referred to in any of the grammars or studies on progressive aspect I have come across. Mindt (2000: 250), however, covers progressives that commonly collocate with *if* (cf. example (76) in this section) under the meaning label “matter-of-course”, which he defines as “something which is required by a specific situation or something that is expected to happen in the normal course of events or something that follows from a rule or agreement”, but this infrequent function, or “meaning” in Mindt's terms, only constitutes a small fraction of my general validity progressives.

Politeness or softening. The next common additional function of progressives in spoken BrNSE is “politeness or softening”. In many concordance lines from the two spoken corpora the progressive form has the function of rendering what is said more polite or less direct. This effect is, for instance, conveyed in examples (82) to (85). Without the progressive forms the utterances would probably sound somewhat more aggressive or face-threatening to the addressees.

(82) I 'm sorry I 'm not clear as to whether you *are suggesting* that there should be policy upper case criteria and some non policy lower case criteria. (BNC_spoken)

- (83) So I *was* just *wondering* how you'd be paid. (BoE_brspok)
- (84) but today I *am asking* you to think about the next step a step that I am sure we all feel is at the heart of the mat (BNC_spoken)
- (85) what I *was* # *meaning* there was <ZF1> not <ZF0> not whether you'd got # extra money that you could spend (BoE_brspok)

The effect of politeness or softening is aided by some typical collocates in the respective concordance lines. Most frequent among the adverbials is *just* (cf. (83)). It occurs in more than 18 per cent of the examined examples. Common also are non-conditional *if*, *whether*, and *actually*. These adverbials contribute to making an utterance sound more polite. Even more significant patterns can be found with respect to subjects and forms of TO BE in the progressive constructions. By far the most common subject, with 56.56% of all politeness cases, is the personal pronoun *I* (cf. (83) to (85)). Also quite often used is the pronoun *you* (cf. (82)) with 16.74%. All other subjects are rather infrequent. Like (83) and (85), a large number of examples contain PastProg forms. With roughly 40 per cent the share of past progressives lies about 10–15% above average (cf. Section 4.3.1). This frequent use of *was* and *were* in progressive constructions also helps to enhance the politeness or distancing effect.

To find out whether “politeness or softening” is a function of progressives in general or only of a small group of verb forms, all BNC_spoken and BoE_brspok tokens were filtered for “politeness” and sorted according to verb types. 62 out of 100 types appeared in the resulting list, but 43 of these 62 forms had only between one and nine occurrences in the whole dataset. This means that only 19 out of 100 progressive forms are reasonably common in contexts of politeness or softening. In contrast to our observations on the most common additional progressive function “general validity”, we cannot say that politeness is a function of the progressive in general. Its use in spoken English is to a large extent lexically determined, meaning that only a restricted set of forms can express this function. Table 14 provides a list of these items. Altogether, the 15 verb forms listed in the table make up 81.14% of all “politeness or softening” examples.

The use of progressive forms to render an utterance more polite or less direct is also mentioned by some other researchers. Scheffer (1975:36), for instance, notes

if someone is telling you something which you know already, it is politer for you to say “George was telling me about it” than “George told me about it”.

Quirk et al. (1985:210) talk about the possible use of progressives “to refer tentatively to a present wish or attitude”. As an illustrative example they give the sentence *I’m*

Table 14. Verb forms frequently occurring in “politeness or softening” contexts in BNC_spoken and BoE_brspok

asking	hoping	ringing	suggesting	trying
checking	meaning	saving	telling	wanting
forgetting	paying	saying	thinking	wondering

hoping to borrow some money (see also Greenbaum 2000:268). Williams (2002:204f.) also reports on the “diplomatic” use of the form, though with reference to the future progressive which is not covered in the present study. An example he uses is *Will you be using the car this evening?* Unfortunately, neither of these scholars provides information about how common the politeness function of the progressive actually is. The only empirical study that deals with politeness or softening as a function of progressives is Mindt’s 2000 monograph. The author gives three examples of “politeness/downtoning”, one of his nine meanings of the progressive, containing typically used verb forms (*wondering, thinking, suggesting*; cf. Table 14) and notes that this meaning is very infrequent (below 5%; cf. Mindt 2000:249, 256). This discrepancy between Mindt’s and my own findings may again be explained on the basis of the different corpora we used. Apparently, the “politeness or softening” function of progressives is a function of spoken rather than of written English.

Emphasis or attitude, shock or disbelief. As Figure 41 above has shown, almost ten per cent of the analysed progressives express an emphatic or evaluative meaning, subsumed under the labels “emphasis/attitude” and “shock/disbelief”. Speakers often use progressive forms to put stress on something, to convey their (mostly negative) attitude to something, or to express strong surprise or severe doubt about something.⁷⁰ In (86) and (87), *are hoping* and *’re wanting* clearly serve to emphasise the event in question, while (88) and (89) convey an attitude of annoyance or irritation on part of the speaker.

- (86) various people here are er, *are hoping* that that that smoking will become a thing of the past (BNC_spoken)
- (87) Now faith means that you’re *wanting* # this and you’re hungry for this (BoE_brspok)
- (88) You never drank this coffee! You ’re always *buying* things and then wanting Oh! to buy something better! (BNC_spoken)
- (89) Cos like when he’s around his friends <ZGY> he’s always *calling* me names and saying I’m stupid. But when we’re at # home he’s really nice and that. (BoE_brspok)

Emphatic progressives show a comparatively high percentage of first person subjects (*I* and *we*) and often collocate with the time adverbials *always*, *now*, and *all the time*. Unsurprisingly, *always* also typically occurs in utterances which convey a certain annoyance about the fact that somebody keeps doing something the speaker does not really appreciate (cf. (88) and (89)). In such instances, we find a large number of second and third person pronouns (mainly *you* and *he*).

As in the case of “politeness/softening” progressives, 62 out of 100 forms show occurrences in the “emphasis/attitude” function, but only 16 of them occur at least ten times. The 15 most frequent verb forms that express emphasis or a negative attitude are presented in Table 15. They account for roughly 60% of all “emphasis/attitude” verb

Table 15. Verb forms frequently occurring in “emphasis/attitude” contexts in BNC_spoken and BoE_brspok

asking	giving	listening	ringing	suggesting
bothering	hoping	meaning	seeing	telling
costing	letting	needing	showing	wanting

Table 16. Verb forms which occur in “shock/disbelief” contexts in BNC_spoken and BoE_brspok

agreeing	calling	letting	suggesting	telling
asking	leaving	saying	supporting	wearing

tokens. This may lead us to conclude that the expression of criticism or of an emphatic meaning is not a general function of progressive forms but just a function of a select group of lexical items.

Negative evaluation or a critical attitude is also expressed in (90), (91), and (92). It becomes clear that the speakers are very much surprised, shocked even, by what they hear and are not quite prepared to believe it.

- (90) You ’re not *suggesting* pregnancy ’s a disease there are you? (BNC_spoken)
- (91) So what, *are* you *telling* me you ’re gon na be a nun, and never have sex. I ’m not saying that (BNC_spoken)
- (92) What? *Am* I *hearing* right FX dear (BoE_brspok)

This rather infrequent function of “shock/disbelief” is only expressed by a very small number of verbs in my database. Ten out of 100 types (listed in Table 16) were found to feature in contexts comparable to those in examples (90) to (92). Three of them, the semantically related forms *suggesting*, *saying*, and *telling*, make up most of the occurrences which means that this additional function is clearly lexically determined and not a general function of the progressive.

As for significant contextual patterns in the small set of “shock/disbelief” concordance lines, a particularly common collocation of these progressives is with the personal pronoun subject *you*. More than 75% of the analysed examples feature the combination “you are” or “you’re”. Quite clearly, speakers are upset about something that another person, the person they are addressing (“you”), has said. Significant also is the typical occurrence of “shock/disbelief” progressives in interrogatives. While the average percentage of questions in the entire BNC/BoE dataset was below 11% (cf. 4.3.7), the share amounts to almost 87 per cent in the case of this particular subset. This is certainly a colligational peculiarity that enforces the effect of expressing doubt.

Other studies on the progressive do not usually deal with the expression of shock or disbelief. The emphatic or attitudinal use is however touched upon by several scholars. In an article published in 1931, Jespersen for instance refers to the collocation of progressives and the adverbial *always* (or synonyms) and notes that “the expressions often have an emotional colouring: You are always finding fault with me.” (1931:526)

For Comrie (1976:37) and Scheffer (1975:91) the “emotive”, or in Scheffer’s terms “emotional”, use of the progressive is a rather specific one and can be regarded as a deviation from the normal use of progressives to refer to contingent situations or to express duration. Pürschel on the other hand describes the expression of emphasis as a preferred use of the present progressive, however only with reference to his small corpus of BFBS radio programmes (cf. Section 3.5.2).

Among Mindt’s nine meanings of the progressive, “highlighting/prominence” and “emotion” are closest to my “emphasis or attitude” function, “emphasis” corresponding to the former and “attitude” to the latter. The author defines “emotion” as a “[t]erm for the use of the progressive to express an emotional involvement like agreement/consent, sympathy, *annoyance*” (Mindt 2000:249; my emphasis). The label “highlighting/prominence” is used for progressives which “draw attention to something or [...] stress something which is felt to be of special importance” (2000:249). With c. 9%, Mindt’s percentage for “highlighting/prominence” progressives is comparable to the shares I found for “emphasis or attitude” in BNC_spoken (8.95%) and BoE_brspok (8.74%). According to the frequency diagram Mindt (2000:256) provides, “emotion” progressives are very rare.

Gradual change and development. A number of progressives in the database (4.79% of the BNC_spoken and 5.92% of the BoE_brspok examples) indicate the gradual change of a situation or refer to some kind of development. Examples (93) to (96) nicely illustrate this function.

- (93) And you ’ll say, well it ’s *getting* closer, as it gets closer and closer to zero it ’s getting bigger and bigger and bigger (BNC_spoken)
- (94) Oh dear I ’m *forgetting* my engineering terms (BNC_spoken)
- (95) there # is a er er a discernable trend that people *are becoming* more selective they # have more choice now with the advent of cable and satellite (BoE_brspok)
- (96) I mean you know it’s *been* erm *changing* all the time since Sealink # <ZGY> took it over (BoE_brspok)

In (93) something is *gradually* getting closer (and closer), and the speaker in (94) is *gradually* forgetting his/her engineering terms. The changes indicated in (95) and (96) also do not happen suddenly but rather step by step, in small stages over an extended period of time.

There are clear preferences with respect to the verbs that can express this function. Of our set of 100, 22 different types are found in “gradual change/development” progressives, but only thirteen of them are more or less frequent (between 7 and 162 occurrences). These thirteen verb forms, listed in Table 17, account for almost 90% of all tokens. Particularly frequent are the four types *becoming*, *getting*, *starting*, and *changing*. Their tokens make up 67.21% of the total. This demonstrates that, like the additional functions “politeness/softening”, “emphasis/attitude”, and “shock/disbelief”,

Table 17. Verb forms frequently occurring in “gradual change/development” contexts in BNC_spoken and BoE_brspok

becoming	forgetting	making	setting	turning
changing	getting	moving	starting	
coming	learning	picking	taking	

the “gradual change/development” function is clearly lexically determined and not a function of progressives as such.

Most common collocates in subject position are *it*, *people*, *I*, and *we*. Also common are combinations of *the* or a demonstrative pronoun (*that*, *this*) and a noun or noun group as subjects, as illustrated in (97) and (98).

- (97) I was in South Africa when Chris was assassinated and *this anger* was turning into rage and the country was on a knife edge it could have blown up (BNC_spoken)
- (98) you # know *the grammar* is more is becoming more demanding. <M01> Yes. <F01> (BoE_brspok)

Frequent time adverbials in the examined contexts are *now* and *just*. More significant, however, is the collocation of “gradual change/development” progressives with the items *more* (*and more*), *increasingly*, and comparatives such as *bigger*, *better*, or *closer*. If we look at the semantic prosody in the respective concordances (particularly those of *becoming*, and *getting*), we observe a tendency for negative expressions. Even though some things are getting *better*, people more often talk about things that are becoming *worse* or (*increasingly*) *difficult*. Typical examples from the *becoming* concordances are given in (99) and (100).

- (99) Er, it is a very difficult climate, it ’s becoming *increasingly difficult*, and indeed, it ’s affecting the work that we do (BNC_spoken)
- (100) I’m getting a bit of a headache # from Birmingham City at the moment erm it really is becoming *very frustrating* that place and # er I’d appreciate some calls from Albion and Villa fans (BoE_brspok)

If we think about the “gradual change/development” function and what is said in other studies on the progressive, Williams’s (2002) concept of “susceptibility to change” comes to mind. This concept, which forms the centre of Williams’s theory of the progressive, is of course a much broader one than the additional function described here. The particular examples discussed in this section, however, might be regarded as prototypical cases of and as supportive evidence for the “susceptibility to change” approach.

Using authentic examples such as “*The village is changing but it is still undisturbed*” and “*His handwriting is improving*” (1990: 248, emphasis in original), the COBUILD grammar refers to the function of progressive forms “to indicate changes, trends, development, and progress” (1990: 248). On a similar note, Jørgensen (1991: 178) com-

ments on the use of *forgetting* “to express the gradual fading away from the memory of facts which ought to be remembered”. Unfortunately though, there is no information telling us anything about the status of this function with respect to its frequency of use.

Old and new habits. Habitual actions or events are referred to in only about 1.5 per cent of the progressive forms from the two spoken BrNSE corpora (1.47% in BNC_spoken and 1.65% in BoE_brspok). I classified some of the habits, like those expressed in (101) and (102), as “old habits” and others, like those in (103) and (104) as “new habits”.

- (101) How long *were* you *seeing* your boyfriend then before you married him? Er, about six years. (BNC_spoken)
- (102) Fortunately I *was* always *living* in North London and # North London was always pretty quiet. (BoE_brspok)
- (103) I think now that they, they *are* really erm *accepting* us for the fact that we have brought things that they would never have had (BNC_spoken)
- (104) I just heard on the radio # the other day <tc text=coughs> that in America now they’re *selling* a new coffee did # anybody hear about that (BoE_brspok)

One thing we notice immediately about these examples is the co-occurrence of the “old habit” sense with PastProgs and of “new habits” with PresProg forms. And indeed, 61.54% of the former group contain either *was* or *were*, while 90.24% of the latter set have present tense forms of TO BE. As for common subjects with habitual progressives, *I* is most frequent in both sets, followed by *you* in old habit contexts (cf. (101)) and by *they* in new habit examples (cf. (103) and (104)). A significant pattern in the concordance lines referring to new habits, such as (103) and (104), is the collocation of progressive forms and the time adverbial *now*. *Now* occurs in roughly 90 per cent of all adverbially specified examples in this group. The two most frequent adverbials in the old habit set are *when* and *always*. While it is often not entirely clear what time span these habitual progressives with *always* cover and whether there is any time limit, *when*-examples usually refer to temporarily restricted situations (compare examples (105) and (106) below).

- (105) I’m *always* y # *checking* dates when bills got to be # paid (BoE_brspok)
- (106) I probably make more money today than I ever did *when* I *was* *selling* a lot of records as a teenage idol. (BNC_spoken)

The variety of different verb forms in habitual functions is again very limited, which may in part be due to rather low absolute frequencies of occurrence. The “old habit” function is expressed by altogether nine different tokens in the datasets (cf. Table 18). All verb forms occurring to refer to new habits are listed in Table 19. Of the nine types in Table 18, only four (*living*, *staying*, *seeing*, *calling*) are relatively frequent. *Living* and *staying* are by far most common in this function. On top of the “new habit” verb frequency list we find *accepting* and *buying*. Of all other forms there are only between

Table 18. Verb forms which occur in “old habit” contexts in BNC_spoken and BoE_brspok

adding	calling	changing	looking	staying
bringing	carrying	living	seeing	

Table 19. Verb forms which occur in “new habit” contexts in BNC_spoken and BoE_brspok

accepting	buying	finding	making	selling
bothering	calling	living	seeing	understanding

one and three tokens in the BNC_spoken and BoE_brspok datasets. It is hence obvious that this not so common additional function is a rather special one of a small number of lexical items and not very central with respect to the use of the progressive in general.

Back in 1975 Scheffer remarked that “[t]here is some disagreement among grammarians whether habit can be expressed by the progressive or not.” (1975: 89; emphasis in original). The expression of habits is usually described as a central function of simple forms (cf. for instance Eastwood 1994: 84; Hirtle 1967: 50; Leech 1971: 5; Sammon 2002: 30; Tesch 1999: 504), not of their progressive counterparts. Linguists who agree that there is such a habitual function of progressives (e.g. Comrie 1976: 37 or Bache & Davidsen-Nielsen 1997: 303), stress that the habit in question is a temporal one, one which is only valid for a certain period of time. I have already noted above that it is often difficult to determine the actual time span the expressed habits cover and found that habitual (non-attitudinal) progressives with *always*, such as (105), apparently refer to real temporally unlimited habits. These cases are, however, very rare in spoken English, which means that it appears indeed unwise to describe “the reference to non-temporary habits” as a general function of the progressive. It is thus advisable to stress, like COBUILD (1990: 248) do in their entirely corpus-driven grammar, that reference is made especially to those habitual actions that are “new or temporary”. Worth taking into account in any case are the inherent meanings of the individual verb forms used to express this special function. Obviously, a form like *staying* has a stronger notion of temporariness than *living*.

Framing. A final rather infrequent additional function of progressives in spoken English (0.95% in BNC_spoken and 0.73% in BoE_brspok) is the function usually referred to as “framing” (cf. e.g. Jespersen 1931: 524; König 1995: 161; Williams 2002: 36). In such “framing” contexts, something new (mostly set in the past) happens while something else is in progress. The ongoing action, in the progressive, frames the new interrupting event or action. Examples (107) to (110) illustrate this function.

- (107) So anyway yesterday afternoon I *was checking* through it when the phone went again (BNC_spoken)
- (108) Mummy dropped it when she *was giving* it to me. (BNC_spoken)

- (109) erm the phone went just as I *was leaving* the house this # morning and the result was that I forgot to bring <ZF1> my <ZF0> my file. (BoE_brspok)
- (110) MX had a tendency # <F0X> <tc text=coughs> <F0X> to talk while he *was writing* <ZF1> an # <ZF0> and he used to go really # fast and we found it really difficult to keep up (BoE_brspok)

Worth noting about these and other “framing” corpus examples are a few contextual patterns. First of all we find above average shares of PastProgs in the respective sets of concordance lines (roughly 56%). This finding nicely fits in the general story-telling mood conveyed in most of the examples. Next, there is a strong collocational connection between framing progressives and the time adverbials *while* (sometimes *whilst*) and *when*. Over 80% of the examined adverbially specified concordance lines contain either *while/whilst* or *when* (cf. (107), (108), and (110)).

When we look at the verbs that can express this function, we observe a surprisingly high number of 42 (out of 100) types in the small total set of 82 tokens. None of these 42 verb forms stands out as being particularly frequent in framing contexts. All forms occur between one and five times in the BNC_spoken and BoE_brspok datasets. In this particular case and in contrast to most of the above-discussed additional functions, it is thus not possible to talk about a largely lexically determined function of some progressive forms. Instead, “framing” appears to be a general function of the progressive, at least concerning its distribution across verb forms, though not at all with respect to its overall frequency of occurrence. Noteworthy also in this context is the significant relation between this additional function and the central function “continuousness + non-repeatedness” (CF1). The fact that all framing examples express this most frequent central function certainly hints at the prototypical quality of the “framing” sense. Still, one important characteristic of a prototype must be frequency of use, and that does certainly not apply to the framing examples in our datasets.

This finding contradicts what is said about progressive meanings in some linguistic studies. Back in 1931, Otto Jespersen noted that “[t]he *chief use* of the expanded tenses is to serve as a frame round something else, which may or may not be expressly indicated” (1931:524; my emphasis). As we could see, our empirical findings deviate considerably from Jespersen’s observation. His “which may or may not be expressly indicated” expresses, however, that Jespersen used this function in a very general way to merely indicate that somebody is “in the middle of [doing] something” (1931:524). Although he first notes that “the frame idea is not the primary or essential thing about the EXF” [expanded form] in general, Bodelsen (1936:222, 226) goes on stating that “perhaps the most frequent use of the expanded preterite is the one represented by *he was writing when I came in*”, and thus also attributes a significant value to framing contexts.

Several decades later, some scholars are still convinced about the typicality of this particular progressive function. Greenbaum (2000:268), for instance states that “[t]he progressive is *often* used to indicate that one event is in progress when another event occurs” and “to indicate the simultaneity of an event with a state or another event depicted with the simple present or the simple past” (my emphasis; cf. also Greenbaum

1996). Leech (1971:17), in *Meaning and the English Verb*, subscribes to the importance of this function too and describes the constructions of a “temporal frame” around an action as a general effect of the progressive, logically following “from the notion of ‘limited duration’”. König (1995:161), Sammon (2000:32), and Scheffer (1975:40) also focus on the framing function. While Sammon claims that this is a frequent use, Scheffer regards it as one of the secondary meanings of progressives. I would claim that my empirical findings certainly question the centrality of “framing” in descriptions of the use of the progressive. Framing contexts do exist in spoken BrNSE, but they are very infrequent.

4.4.6 Additional functions and time reference

Similar to Section 4.4.4 which dealt with some relations between the two identified central progressive functions (CF1 and CF2) and time references, the present sub-chapter will discuss whether any of the additional functions of progressives shows a significant preference for situations in the past, present, or future. For each of the above-described additional progressive functions (cf. 4.4.5) the connections to the four possible time reference values (past, present, future, present/future “indeterminate”; cf. 4.4.1) will be illustrated by diagrams. Comparisons will be made to the time reference distribution of the full set of spoken BrNSE progressives in the database (cf. Figure 33).

As Figure 42 shows, “general validity” progressives have a strong tendency for present/future “indeterminate” examples. This is of course not very surprising; after all we are here dealing with constructions which refer to something that can apply to the present moment but also to situations to come. The shares of the other three “clear” time reference types are rather low, also if we compare them with the general

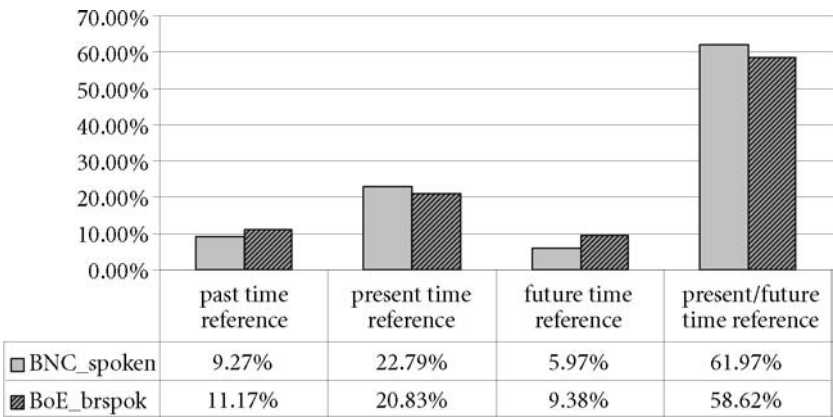


Figure 42. Connections between “general validity” progressives and time reference in BNC_spoken and BoE_brspok

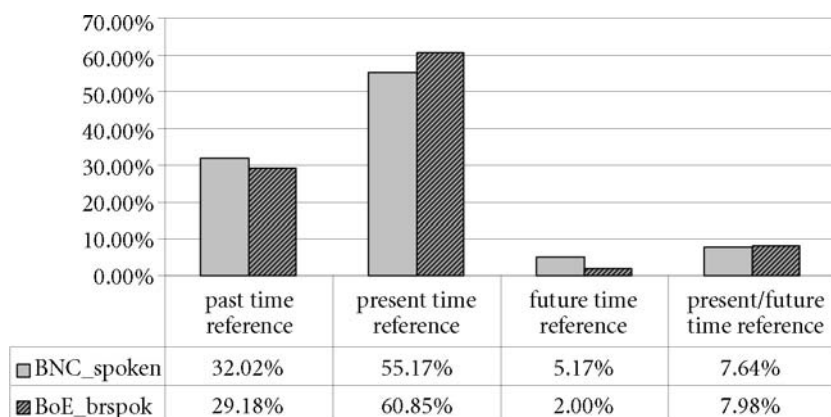


Figure 43. Connections between “politeness/softening” progressives and time reference in BNC_spoken and BoE_brspok

distribution of time references as displayed in Figure 33. We can thus say that there is a significant relation between the function “general validity” and “indeterminate” time reference.

If we now turn to the group of progressives which conveys politeness or softening, we see that the time reference distribution differs quite a lot from the one found for “general validity” examples. Figure 43 indicates, for instance, that there are very few cases of indeterminate and future time reference. Instead, the shares of present time examples in both corpora are significantly high. Politeness is obviously an effect that applies to the here-and-now of a present communicative situation. The percentages for past time reference are roughly average. As observed in Section 4.4.5, PastProg forms often serve to enhance the politeness or softening effect of an utterance.

The connections between the function “emphasis or attitude” and time reference (cf. Figure 44) are quite similar to those just described for politeness and softening. Apparently, the two functions have a lot in common. Like politeness/softening progressives, emphasis/attitude examples mainly refer to the here-and-now. The stress is on something which is relevant at the time of speaking. Interesting in this similarity context is the fact that many verb forms that frequently express politeness are also common in emphatic or attitudinal contexts. Depending on their collocates, progressives around *suggesting*, for instance, can have a softening or downtoning effect, as in the phrase “What I am suggesting...”, or, usually in negative contexts, they can convey emphasis, e.g. in utterances containing “I am not (for a moment/minute) suggesting...”.

Progressives which express shock or disbelief are almost always related to the moment of speaking (cf. Figure 45). Moving them away from the present would probably weaken the effect they convey. A past time reference version of example (90) above for instance, *What? Was I hearing right FX dear*, would not have the same effect as the “real” version with present time reference.

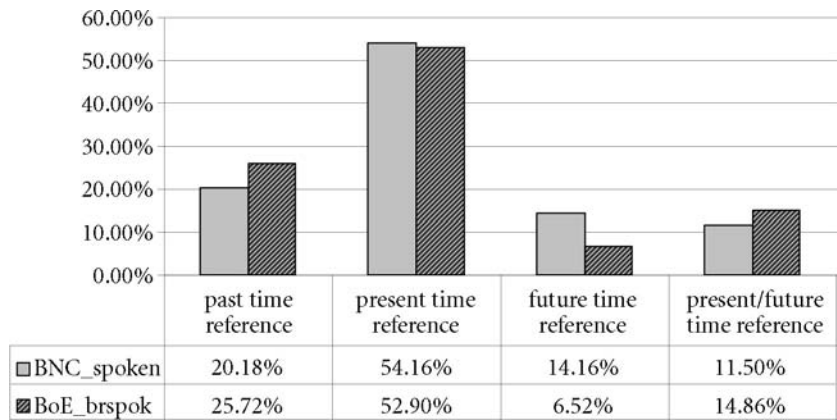


Figure 44. Connections between “emphasis/attitude” progressives and time reference in BNC_spoken and BoE_brspok

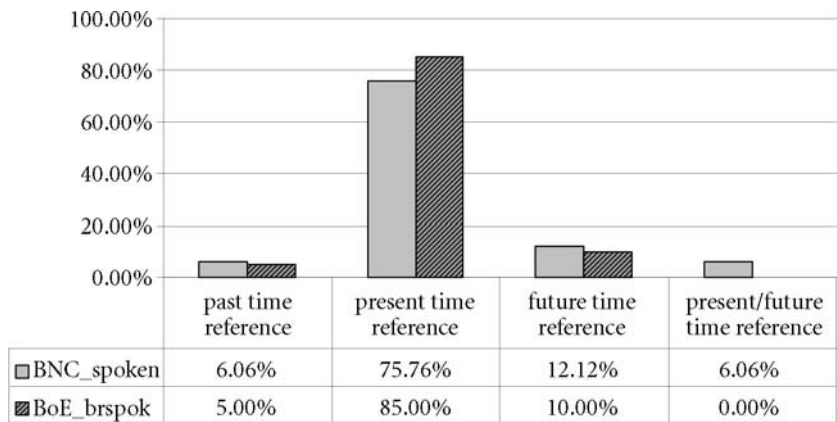


Figure 45. Connections between “shock/disbelief” progressives and time reference in BNC_spoken and BoE_brspok

“Present” is also the main time reference of concordance lines that express a gradual change or development (cf. Figure 46). Speakers normally focus on the fact that something *is changing* at the moment of speaking, even though this change or development goes on in the future and ranges over a longer stretch of time. Considerable shares of gradual change/development examples also refer to things that happened, or were happening, in the past. The percentages, however, are significantly below average (cf. Figure 33).

The last two functions discussed in this section, “old and new habits” and “framing”, show clearly above-average shares of past time reference tokens. Figures 47 and 48 display the distributions of concordance examples which express these two additional

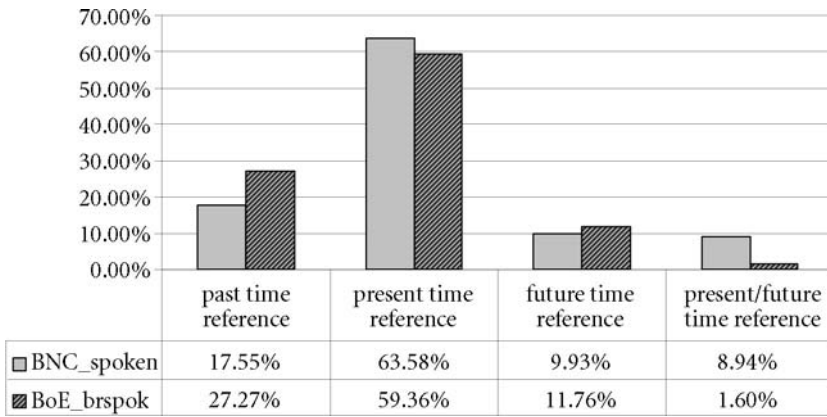


Figure 46. Connections between “gradual change/development” progressives and time reference in BNC_spoken and BoE_brspok

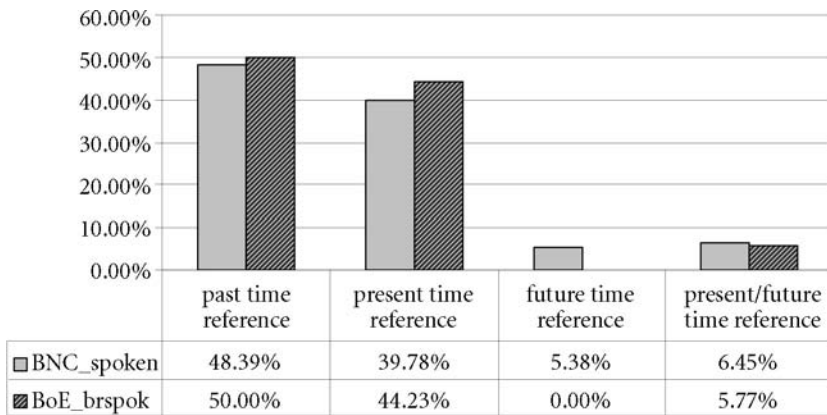


Figure 47. Connections between “old and new habit” progressives and time reference in BNC_spoken and BoE_brspok

functions across the four defined time reference categories. The still relatively high percentage of present time reference in Figure 47 goes back to examples referring to new habits (most frequent collocate: *now*). Of the small group of framing instances, only very few show non-past temporal relations. If we consider that most of these concordance lines are taken from “storytelling” contexts, this significant connection is hardly surprising.

To sum up, significant patterns concerning additional functions and time references are the following:

- there is a clear tendency for “general validity” progressives to refer to present/future (indeterminate) events (cf. Figure 42)

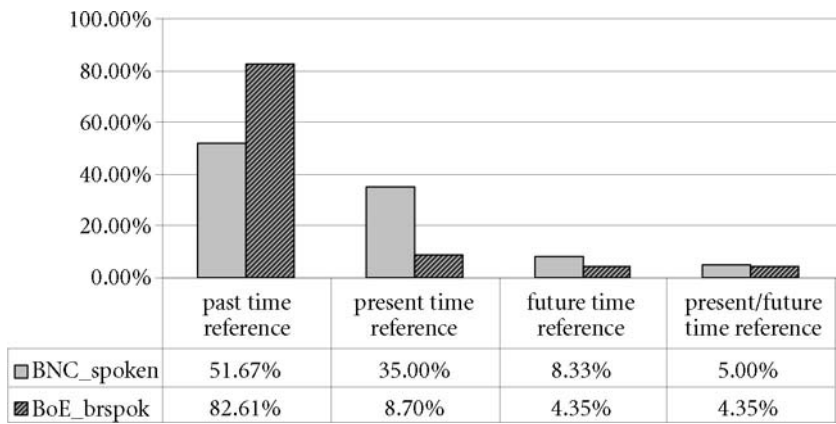


Figure 48. Connections between “framing” progressives and time reference in BNC_spoken and BoE_brspok

- “politeness or softening” and “emphasis or attitude” examples show particularly high shares of present and average shares of past time reference (cf. Figures 43 and 44)
- “shock or disbelief” progressives almost exclusively refer to present time actions or events (cf. Figure 45)
- “gradual change or development” cases also show an above average percentage of present time reference (cf. Figure 46)
- “old and new habits” and “framing” progressives are significantly frequent in past time reference contexts; examples in the “new habit” subgroup refer to the present (cf. Figures 47 and 48).

4.4.7 Summary of the findings [spoken English – functions]

The main aim of this Section (4.4) has been to throw some light on the actual use of progressives in spoken English, on the types of things they commonly refer to, and on the functions they can serve in communication.

In a first analytic step, I determined the time reference expressed in each of the 9,468 corpus examples from BNC_spoken and BoE_brspok. Highest shares (38.09 and 38.07 per cent) were attributed to the present time reference group. Less than one third of the examples referred to actions or events in the past, and between 14.38% and 18.55% of the concordance lines expressed a future or an “indeterminate” (present and/or future) reference (cf. Figure 33). I noted that the shares of future reference examples are significantly higher than those identified in previous empirical studies.

I then looked at the relationship between time reference and progressive verb form and, unsurprisingly, found clear connections between past time reference and PastProg, PresPerfProg, and PastPerfProg and between the other three time reference

values (present, future, indeterminate) and PresProg, but also some typical exceptions from these “normal” relations.

Next was the determination of central function features and the analysis of their distributions and most frequent combinations. Section 4.4.3 then led to the conclusion that, based on the corpus findings, it makes sense to distinguish two central functions, CF1 and CF2. CF1 serves to refer to continuous and non-repeated actions or events while CF2 expresses continuousness and repeatedness. It was found and empirically proven that these two functions can be called basic functions of progressives in general. With respect to existing descriptions of the progressive we observed that the feature “repeatedness” often does not receive enough attention and that it should be more clearly emphasised as an important function feature of progressives in use. Examining the relations between central functions and time reference led us to conclude, among other things, that “future” as such is not a separate function of the progressive, but a feature that combines with both central functions. It is hence not significantly different from past or present time reference categories and it appears unjustified to treat “future”, like it is often done by a number of scholars, as a separate meaning of the progressive.

As the two specified central functions did not suffice to fully account for the use of all examined progressives, I then turned to looking at additional progressive functions. The seven additional functions of progressive forms that were found to occur repeatedly are (in order of frequency of occurrence in the data): “general validity”, “politeness or softening”, “emphasis or attitude”, “gradual change and development”, “old and new habits”, “framing”, and “shock or disbelief”. While “general validity” and “framing”, according to their wide distribution across different verb types, can be called functions of the progressive in general (though not necessarily common functions as we saw in the case of “framing”), the other five additional functions are strongly lexically determined and form certain typical patterns with restricted sets of progressive forms. We also noted a number of typical patterns concerning additional functions and subjects, adverbials, tense forms, and time references (cf. Subsections of 4.4.5 and 4.4.6).

In Chapter 6 of this book, the results of this detailed function analysis of progressives in spoken BrNSE will be compared with the findings obtained in a close examination of progressives in EFL coursebook English (cf. Chapter 5). However, before moving on to investigating “school” English, the following parts of the present chapter will look further into the relations between progressives and their functions and contexts on the one hand and individual verbs on the other.

4.5 Verbs and progressives – How lexical is grammar?

Throughout the previous chapters repeated reference has been made to the concept of “lexical grammar” and to “lexical-grammatical” phenomena in the context of progressives. As the name implies, “lexical grammar” stresses the integration of lexis and grammar in language analysis and description. Instead of keeping the two components

apart and treating them separately, the approach sees language as an integrated whole, built from both parts. In other words, lexis and grammar are considered to be closely linked and strongly dependent on each other (cf. Section 2.1).

It was mentioned earlier on that the present corpus-driven study has been strongly influenced by research in the lexical-grammatical tradition, in particular the work of Hunston, Sinclair, and Tognini-Bonelli. Empirical corpus-driven research has demonstrated that there are certain important restrictions concerning the co-selection of and connections between lexical items (cf. e.g. Hunston & Francis 1998, 2000; Hunston 2002b; Sinclair 1991, 2000b; Tognini-Bonelli 1996a, 1996b, 2001). In many cases, these connections are strong enough to prohibit a strict separation of lexis (or the lexicon) on the one hand and grammar (or syntax) on the other hand (cf. Hunston & Francis 2000; Sinclair 1991 for examples).

The aim of the following sections is to analyse the strength of the connections between progressive constructions and individual verb forms, and, at least with respect to the progressive, to answer the question “How lexical is grammar?” A detailed analysis of the behaviour of 100 -ing form types (9,468 tokens) in context, taken from BNC_spoken and BoE_brspok concordances, is supposed to help me find out whether there is a *purely grammatical* or rather a *lexical-grammatical* progressive. I will look at a number of co-selection types on the contextual and functional levels and examine, among other things, relations between verbs (i.e. my 100 selected lexical items treated individually) and progressive tense forms (cf. 4.5.2), verbs and personal pronouns in subject position (4.5.3), verbs and negation (4.5.6), verbs and time reference (4.5.9), and verbs and progressive functions (4.5.10 and 4.5.11). The existence or inexistence of significant patterns for each context or function phenomenon will provide me with further insights about the nature of progressives. It will be interesting to see whether there are strong dependencies between the progressives of individual verbs and their functions and contexts, or whether the average distributions determined in Sections 4.3 and 4.4 are roughly valid for the majority of the examined progressive types, which would mean that it is legitimate to talk about “the progressive” as such, independent of its individual lexical realisations.

Whereas, up to now (in the previous sections of Chapter 4), I have been looking at a *group of 100 verbs* and 9,468 progressive form tokens taken together, I will now look separately at *100 individual lexical items*, i.e. progressive forms of 100 verbs and their relationships with contexts and functions. After they showed largely similar distributions in almost all context and function categories (see Sections 4.3 and 4.4), the concordance datasets from the two included spoken BrNSE corpora, BNC_spoken and BoE_brspok, will from now on be treated together under the label “BNC/BoE”. The strong similarities between the empirical results independently retrieved from BNC_spoken and BoE_brspok can be said to indicate that my findings are highly representative of spoken British English in the 1990s. All results presented in the following sections are based on progressives from BNC_spoken and BoE_brspok treated as one dataset.

With the exception of 4.5.1, references to other scholars's findings on verbs and progressives will be rarely made in this chapter. This is because mostly linguists deal with the use of progressives in general, especially when they include a quantitative dimension (cf. e.g. Mindt 2000 or Biber et al. 1999). Little research has been done on the behaviour of individual progressive verb forms. The present approach will try to contribute to this topic and provide new insights into the lexical grammar of progressive forms.

The identification of central co-selectional patterns of individual progressives and context and function features in spoken English will also be of major importance for the pedagogical part of the study. All identified typical patterns of co-selection will later on be used as a basis for comparison with the most common patterns in the selected teaching materials.

4.5.1 Distribution and restrictions: 100 verbs and 9,468 concordance lines [BNC/BoE]

The present section serves to give a brief account of the internal composition of the BNC/BoE dataset and comments on some obviously existing restrictions concerning the progressive and verb selection. We will see whether our corpus-driven findings support or contradict what has previously been said about verb-progressive relations.

Table 20 presents the type-token distribution of the included verb forms and shows how the 9,468 progressive form tokens from BNC_spoken and BoE_brspok are split up across 99 types.⁷¹ The verb forms are sorted according to progressive token frequency and displayed in decreasing order, which means that verb forms with high numbers of occurrence in the progressive top the list and items which are uncommon in progressive constructions can be found towards the bottom of the list. The token numbers, however, have to be treated with some caution as they are determined by different factors. For one thing, depending on the item's frequency of occurrence in BNC_spoken and BoE_brspok, the size of the originally saved concordance sets could differ considerably. For most forms, sets of 300 concordance lines were downloaded for further analysis, but a couple of items occurred less frequently in the selected corpora (cf. Table 5 in 4.2.3 for individual figures).

Besides, in the data filtering process described in Section 4.2.3, differing shares of "non-progressive" instances were deleted from the concordances. In some cases the number of concurring non-progressive form tokens was particularly high, for example in the two concordances of *meaning*, where I found a lot of instances of *meaning* used as a noun, which of course had to be eliminated. However, this was a rather extreme and exceptional case. With respect to most of the other forms, the numbers of the remaining tokens in fact reveal something about the forms's status as "progressive" verbs, i.e. verbs that like to occur in the progressive. Taking into account the just-mentioned caveats, we can say that types with comparatively high token numbers are typical "progressive" verbs, whereas those with low token numbers either clearly favour non-progressive tense forms or, like *meaning*, have strong homograph competitors.

Table 20. The distribution of 9,468 progressive form tokens across 99 verb types* (in order of frequency), based on BNC/BoE data

verb form	number of tokens	verb form	number of tokens	verb form	number of tokens
1. wondering	239	34. leaving	110	67. speaking	70
2. happening	235	35. playing	110	68. supporting	69
3. hoping	229	36. starting	110	69. hearing	67
4. expecting	228	37. stopping	108	70. accepting	66
5. suggesting	223	38. carrying	107	71. learning	66
6. going	208	39. making	104	72. cutting	65
7. talking	188	40. selling	104	73. drawing	63
8. trying	184	41. sending	102	74. adding	61
9. doing	165	42. putting	101	75. writing	61
10. becoming	164	43. coming	100	76. worrying	60
11. telling	160	44. pulling	99	77. providing	59
12. listening	156	45. showing	99	78. saving	57
13. saying	153	46. holding	97	79. bothering	56
14. asking	151	47. walking	97	80. setting	54
15. paying	150	48. letting	96	81. agreeing	52
16. wearing	149	49. running	96	82. winning	50
17. looking	146	50. standing	93	83. explaining	48
18. getting	144	51. finding	91	84. needing	42
19. thinking	141	52. moving	91	85. following	41
20. staying	135	53. using	91	86. forgetting	37
21. dealing	128	54. bringing	90	87. meaning	37
22. costing	127	55. changing	89	88. sorting	33
23. giving	126	56. checking	85	89. remembering	30
24. ringing	126	57. helping	85	90. imagining	22
25. sitting	125	58. finishing	83	91. meeting	13
26. calling	124	59. keeping	81	92. understanding	9
27. taking	120	60. turning	80	93. being	8
28. buying	116	61. living	77	94. liking	8
29. working	116	62. spending	77	95. betting	7
30. watching	114	63. feeling	75	96. believing	6
31. picking	111	64. seeing	75	97. seeming	2
32. wanting	111	65. reading	72	98. knowing	1
33. eating	110	66. having	70	99. supposing	1

* The form *matter* never occurs in progressive constructions, neither in the BNC_spoken nor in the BoE_brspok dataset.

As Table 20 shows, there is quite some variation in progressive token numbers among the 99 listed verb forms. Absolute numbers range from 1 to 239. Particularly frequent in the dataset of 9,468 concordance examples are progressives with *wondering*, *happening*, *hoping*, *expecting*, *suggesting*, and *going*. These and other items in the first column of the list are obviously typical progressive-favouring verbs. Semantically, many of these progressive-favouring verbs in column one can be grouped into two classes, mental activity verbs (e.g. WONDER, HOPE, EXPECT) and communication verbs

(e.g. TALK, TELL, SAY, ASK). Other verb types displayed towards the bottom of Table 20 only occur very rarely in the progressive. Most infrequent are progressives with the items *knowing*, *supposing*, and *seeming*. The only occurrences of these verb forms in our spoken English datasets are given in examples (111) to (114).

- (111) we *were* not *knowing* it we didn't # know where to go. <M01> Okay. <F01> I think we got to about there somewhere (BoE_brspok)
- (112) I think he *was* expecting and *supposing* that there was far # more there than there actually is (BoE_brspok)
- (113) And how old would he be? Well he *was seeming* to six when he died. You see he used to get over the stick pretty often, if you understand what I mean by over the stick (BNC_spoken)
- (114) I # say MX *was* always *seeming* to be more naughty so he was always getting # told off more. (BoE_brspok)

Interestingly, in the only *knowing* example the speaker rephrases her utterance, first using a PastProg but then a past simple tense form. In (112) it is not clear whether *supposing* is really part of a progressive construction ("he was expecting and supposing that..."). It might be that in the original conversation there was a pause after "expecting" and that what follows is a new utterance, starting off with "supposing". Example (113) is somewhat obscure too. It is taken from a story told about an Irishman who was found dead by the police and who, apparently, used to "get over the stick" (get drunk) quite often. In this context "was seeming to six" sounds odd. I think something like "was seeming to be *sixty*" would be the more likely version of this phrase. In any case, the PastProgs in both (113) and (114) have an emphatic effect and serve to stress the content of the utterance.

Also not very commonly used in the progressive are the forms *believing*, *betting*, *liking*, and *being*. One corpus example for each of these items is presented in (115) to (118) below. In examples (115) and (117) the progressives emphasise the respective situations, whereas (116) indicates a certain temporary behaviour or non-permanent characteristic of the addressee. The form *liking* in (118), postmodified by "more and more", clearly expresses the function of gradual change.

- (115) And he's genuinely *believing* what he's telling us (BoE_brspok)
- (116) you 're *being* very intimidating in the way that your [you're] talking to people. (BNC_spoken)
- (117) But Richard Branson is so determined to win he 's now *betting* with his own fortune to put more planes like this one in the sky, (BNC_spoken)
- (118) ooh I 'm *liking* the idea of this more and more! Yes! Oh no! Ooh! Oh definitely! (BNC_spoken)

We will now see what other scholars have said about verbs which favour or do not usually occur in progressive constructions and discuss some relations to our findings.

According to Recktenwald (1975:21), the fact that “some verbs do not normally appear in the progressive form [...] was observed as early as 1789 by the philologist J. Pickbourn.” Two centuries later a number of linguists make largely similar statements. In their argumentations they usually distinguish between “stative” and “non-stative” or “dynamic” verbs, i.e. verbs that refer to states (e.g. BE, KNOW, OWN) and verbs that refer to actions, processes, or events (e.g. PLAY, RUN, WORK), and claim that stative verbs are generally incompatible with the progressive.

Comrie (1976:35), for instance, notes that “verbs tend to divide into two disjoint (nonoverlapping) classes, those that can appear in the progressive forms, and those that cannot” and that “this distinction corresponds to that between stative and nonstative verbs”. He points out, however, that it is possible to use lexically stative verbs nonstatively in the progressive, as in “*Fred is being silly*” (Comrie 1976:36; cf. Section 3.3.1; see also Williams 2002:28). A case in point in this group of “stative progressive verbs” are the verbs of perception, e.g. SEE OR HEAR. Although they are generally stative, these verbs can express continuousness, as the corpus examples in (119) to (122) show. Obviously, it is very much context-dependent whether verbs classified as “stative” can appear in progressives. Also, as Quirk et al. (1985:202) note, in such cases “some change of interpretation” may be required. In examples (119) to (122), I would argue, the main function of the progressive is an emphasis of the proposition.

- (119) And you still *were seeing* your mum and dad and # your brothers and sisters?
(BoE_brspok)
- (120) You *'re not seeing* the real me! I am not just a teacher, I 'm so much more than that! (BNC_spoken)
- (121) and I *'m not hearing* too good. (BNC_spoken)
- (122) Is this something you're *hearing* a lot about at the moment (BoE_brspok)

The numbers of occurrence for *hearing*, *hoping*, *seeing*, and *thinking* contradict the observation made by Joos and others that certain verbs, such as HEAR, HOPE, SEE, and THINK, are rarely or “practically never” used in the progressive (Joos 1964:115; cf. also Jespersen 1931:529; Jørgensen 1991:173ff.; Leisi 1960:221; Palmer ²1988:71ff.; Scheffer 1975:65ff.; Zydatis 1976a:199). It can be argued that an increasing number of progressive form tokens of so-called stative verbs occur in contemporary spoken English. On the other hand, our findings largely confirm what Bache and Davidsen-Nielsen (1997:298) say about the verbs BELIEVE, KNOW, and MEAN, namely that they “do not normally take kindly to the progressive form” (cf. low absolute numbers in Table 20).

Biber et al. (1999:472) also discuss several “[l]exical associations of progressive aspect”. In their first large-scale empirical account on verb-progressive relations, the authors find, in contrast to previous non-empirical studies, that

both dynamic and stative verbs are included among the most common verbs in the progressive – and that both dynamic and stative verbs are included among the verbs that very rarely take the progressive. (Biber et al. 1999:472)

The results of my corpus-driven analysis clearly support Biber et al.'s findings. Whereas the progressive forms of the stative verbs *LISTEN*, *LOOK*, *STAY*, and *COST* count among the common progressive types, some dynamic verbs, such as *FOLLOW* or *SORT*, rarely occur in progressives. The authors's observation that those mental verbs which normally take a human subject are frequent in progressive constructions (cf. Biber et al. 1999: 472–473) was also confirmed with reference to large amounts of spoken English corpus data in my study. Verbs like *WONDER*, *HOPE*, *EXPECT*, and *SUGGEST* top the frequency list displayed in Table 20. We noted above that many of the typical progressive verbs refer to mental activities or communication processes – an observation which is also in accordance with Biber et al.'s results. Concluding it can be said that while some traditional beliefs on verb-progressive relations are certainly still valid, other statements concerning non-progressive verbs do not meet corpus evidence.

4.5.2 Verbs and tense form distributions [BNC/BoE]

The first type of co-selection I am going to look at in order to tackle the above-formulated problem of a purely grammatical vs. a lexical-grammatical progressive is the one between verbs and progressive tense forms. What I wish to analyse is whether the average tense form distribution determined in Section 4.3.1 can be considered representative of progressives in general or at least of the majority of progressives, or whether the individual verbs differ considerably concerning their frequencies of use in the past progressive (PastProg), the present progressive (PresProg), the present perfect progressive (PresPerfProg), and the past perfect progressive (PastPerfProg), which would then be an indicator of a strong lexical orientation of a grammatical construction.

Table 21 displays the shares of the four included progressive tense forms for each of the 99 verbs (BNC_spoken and BoE_brspok taken together). Let us first deal with the individual percentages of the past progressive. When we look at the graphic illustration of PastProg shares in the right-hand column of Table 21, we notice immediately that the values for each individual verb form differ considerably. Percentages range from below 10% (e.g. *accepting*: 7.58%, *changing*: 8.99%, *remembering*: 3.33%) to over 40% (e.g. *expecting*: 43.86%, *living*: 44.16%, *standing*: 44.09%, *thinking*: 43.97%, *walking*: 49.48%) or even 54.39% in the case of *wondering*.⁷² There is obviously a lot of variation in PastProg distributions if we compare these values with the average share of 26.00% determined for the entire BNC/BoE dataset (cf. Figure 8 in 4.3.1 for distributions in BNC_spoken and BoE_brspok). Some -ing forms clearly favour PastProgs, while others very infrequently combine with *was* or *were*.

The two extracts from BNC_spoken concordances in Figures 49 and 50 serve to illustrate this contrast between “PastProg-ophile” and “PastProg-ophobe” verbs. While there are only three PastProg form tokens in the concordance of *changing*, the *walking* concordance displays a large number of instances of “was walking” and “were walking”. Apparently, “change” is something clearly related to the present moment and/or to the future, whereas “walking” is one of the actions that often feature in a reporting

Table 21. The distribution of progressive tense forms for each analysed verb (verbs in alphabetical order)

verb form	PastProg ø26.00%	PresProg ø69.69%	PresPerf Prog ø3.66%	PastPerf Prog ø0.64%	distribution graphically illustrated ■PastProg ■PresProg □PastPerfProg ■PastPerfProg
accepting	7.58%	90.91%	1.52%	0.00%	
adding	11.48%	81.97%	6.56%	0.00%	
agreeing	26.92%	73.08%	0.00%	0.00%	
asking	33.11%	61.59%	3.97%	1.32%	
being	25.00%	75.00%	0.00%	0.00%	
becoming	20.12%	79.27%	0.00%	0.61%	
believing	33.33%	66.67%	0.00%	0.00%	
betting	14.29%	71.43%	0.00%	14.29%	
bothering	16.07%	80.36%	3.57%	0.00%	
bringing	18.89%	77.78%	3.33%	0.00%	
buying	25.00%	71.55%	2.59%	0.86%	
calling	22.58%	75.00%	2.42%	0.00%	
carrying	33.64%	59.81%	4.67%	1.87%	
changing	8.99%	85.39%	5.62%	0.00%	
checking	29.41%	68.24%	2.35%	0.00%	
coming	23.00%	72.00%	3.00%	2.00%	
costing	22.05%	77.95%	0.00%	0.00%	
cutting	29.23%	69.23%	1.54%	0.00%	
dealing	16.41%	78.91%	4.69%	0.00%	
doing	15.76%	78.18%	6.06%	0.00%	
drawing	11.11%	80.95%	7.94%	0.00%	
eating	24.55%	67.27%	5.45%	2.73%	
expecting	43.86%	54.39%	1.32%	0.44%	
explaining	41.67%	39.58%	16.67%	2.08%	
feeling	30.67%	65.33%	4.00%	0.00%	
finding	18.68%	76.92%	4.40%	0.00%	
finishing	24.10%	74.70%	1.20%	0.00%	
following	17.07%	80.49%	0.00%	2.44%	
forgetting	18.92%	81.08%	0.00%	0.00%	
getting	25.69%	71.53%	2.78%	0.00%	
giving	15.87%	79.37%	3.17%	1.59%	
going	15.38%	84.13%	0.48%	0.00%	
happening	20.09%	76.50%	3.42%	0.00%	
having	25.71%	71.43%	1.43%	1.43%	
hearing	23.88%	52.24%	23.88%	0.00%	
helping	30.59%	60.00%	9.41%	0.00%	
holding	21.65%	75.26%	3.09%	0.00%	
hoping	24.89%	73.36%	0.44%	1.31%	
imagining	36.36%	63.64%	0.00%	0.00%	
keeping	16.05%	81.48%	2.47%	0.00%	

Table 21. (continued)

knowing	100.00%	0.00%	0.00%	0.00%	
learning	19.70%	75.76%	4.55%	0.00%	
leaving	21.82%	78.18%	0.00%	0.00%	
letting	22.92%	73.96%	3.13%	0.00%	
liking	0.00%	100.00%	0.00%	0.00%	
listening	26.28%	67.31%	6.41%	0.00%	
living	44.16%	48.05%	5.19%	2.60%	
looking	26.03%	71.92%	2.05%	0.00%	
making	25.00%	71.15%	2.88%	0.96%	
meaning	45.95%	32.43%	21.62%	0.00%	
meeting	15.38%	76.92%	7.69%	0.00%	
moving	15.38%	83.52%	1.10%	0.00%	
needing	23.81%	76.19%	0.00%	0.00%	
paying	19.33%	74.67%	4.67%	1.33%	
picking	12.61%	84.68%	2.70%	0.00%	
playing	19.09%	70.00%	7.27%	3.64%	
providing	16.95%	81.36%	1.69%	0.00%	
pulling	29.29%	67.68%	3.03%	0.00%	
putting	12.87%	83.17%	1.98%	1.98%	
reading	36.11%	51.39%	11.11%	1.39%	
remembering	3.33%	96.67%	0.00%	0.00%	
ringing	30.95%	61.90%	6.35%	0.79%	
running	27.08%	56.25%	15.63%	1.04%	
saving	21.05%	59.65%	17.54%	1.75%	
saying	31.37%	68.63%	0.00%	0.00%	
seeing	20.00%	76.00%	1.33%	2.67%	
seeming	100.00%	0.00%	0.00%	0.00%	
selling	29.81%	66.35%	3.85%	0.00%	
sending	17.65%	80.39%	1.96%	0.00%	
setting	24.07%	72.22%	1.85%	1.85%	
showing	25.25%	72.73%	2.02%	0.00%	
sitting	36.00%	61.60%	0.80%	1.60%	
sorting	36.36%	60.61%	3.03%	0.00%	
speaking	27.14%	64.29%	7.14%	1.43%	
spending	15.58%	75.32%	9.09%	0.00%	
standing	44.09%	52.69%	2.15%	1.08%	
starting	24.55%	75.45%	0.00%	0.00%	
staying	31.85%	65.93%	1.48%	0.74%	
stopping	18.52%	77.78%	2.78%	0.93%	
suggesting	14.80%	83.86%	1.35%	0.00%	
supporting	27.54%	71.01%	1.45%	0.00%	
supposing	100.00%	0.00%	0.00%	0.00%	
taking	20.00%	75.00%	4.17%	0.83%	
talking	23.94%	67.02%	8.51%	0.53%	
telling	45.00%	51.25%	3.13%	0.63%	
thinking	43.97%	55.32%	0.71%	0.00%	

Table 21. (continued)

trying	16.85%	78.26%	4.35%	0.54%	
turning	16.25%	81.25%	2.50%	0.00%	
understanding	22.22%	77.78%	0.00%	0.00%	
using	25.27%	65.93%	6.59%	2.20%	
walking	49.48%	46.39%	4.12%	0.00%	
wanting	27.93%	67.57%	1.80%	2.70%	
watching	40.35%	49.12%	8.77%	1.75%	
wearing	33.56%	64.43%	2.01%	0.00%	
winning	34.00%	66.00%	0.00%	0.00%	
wondering	54.39%	43.93%	1.26%	0.42%	
working	39.66%	51.72%	6.03%	2.59%	
worrying	16.67%	73.33%	8.33%	1.67%	
writing	32.26%	64.52%	3.23%	0.00%	

or story-telling context. Significant also is the high share of the first person pronoun subjects *I* and *we*. “I was walking” and “we were walking” are obviously important clusters in spoken English, much more than “I was changing” or “we were changing”. Among the PastProg-ophile forms we find some of the central verbs that are often used in reporting contexts. The PastProg shares of *asking*, *explaining*, *saying*, *telling*, and *thinking* lie clearly above average and amount to values between 31 and 45 per cent. Biber et al. (1999:475) have also observed a “particular frequency of the verbs *saying* and *thinking* with the past progressive” and point out that this use of the PastProg with reporting verbs “conveys a more vivid imagery and a greater sense of involvement than the simple past tense”. The corpus examples presented in (123) and (124) are supposed to illustrate this emphatic function of the PastProg in reporting contexts. I would say that the use of *was saying* in (123) serves to highlight the message that we need to help young girls get self-confidence, while the progressive in (124) puts some stress on the thinking activity of the speaker and thus nicely enhances the ironic effect of the utterance.

- (123) we need to er, go back to what Kathleen *was saying* about education, we need to help young girls get self-confidence, (BNC_spoken)
- (124) I was on, on top of [Mount] Everest, er, I *was thinking* about you, because I’ve heard you’re the best thing since sliced wholemeal bread, (BNC_spoken)

Quite some variation can also be found with respect to the PresProg shares of the 99 different verb forms listed in Table 21. The average share is 69.69% but individual values range from 32.43% (*meaning*) to 96.67% (*remembering*). As their above-average percentages show, the forms *accepting* (90.91%), *changing* (85.39%), *going* (84.13%), *moving* (83.52%), *picking* (84.68%), and *putting* (89.17%) clearly count among the PresProg-favouring types and either refer to the present or the future. On the other hand, comparatively infrequent in PresProg constructions are the forms *expecting* (54.39%), *hearing* (52.24%), *living* (48.05%), *reading* (51.39%), *standing* (52.69%),

and of technology now. Technology is changing our world so much that small groups
 E do we need to know that the trend is changing or do we need to know the particular
 it goes good again! So Weird! I 'm changing paste anyway. I 'm gon na try a dif
 s right, yeah. Yeah. And that, I 'm changing that way, but, if our kids have to
 ? Er, I can say for instance that I 'm changing because I 'm er getting older, I 'm
 ded quite soon soon that 's why I 'm changing the disc to put in something else th
 wrongly, that actually things are n't changing very much, and they may think they
 it again. So the voltage is n't changing, you do n't You need to worry about
 g Go and change it then No, I 'm not changing it now That 's alright but you 've g
 okeless for. Yeah. And they 're not changing anything. And you get twice as much
 okeless for. Yeah. And they 're not changing theirs then. And you get twice as m
 ng what 's in the memory. You 're not changing anything that 's on the computer at
 at 'll be velocity, so where it 's not changing, there, the slope is zero, no acce
 ending of the Cold War things are now changing, there is not a demand. I do n't kno
 arish members etcetera. If we are now changing the area to ten miles then the b
 that er we have er mentioned is er now changing. Right, now you you 've already ma
 not very long is it? Because they 're changing so rapidly. Yes, yeah. I mean my,
 metimes if it would . . . yo if they 're changing key Yeah. they might put a double
 confusing is n't it. Perhaps they 're changing the language round back to front. Ri
 ch is what Right. So they 're changing the way structures, their possibilit
 t 's probably why get into. They 're changing places. Mhm. Well she 's quite enj
 ou the rules. Tell daddy what they 're changing. Oh they 're going to change the ru
 ve got a big file right? And you 're changing all your currencies of a certain var
 pting this but I think it 's You 're changing battle ground there Mr Chairman
 was stopped. Okay, well yeah, you 're changing it as you go now. Yeah. The rain of
 r at the top there where the colour 's changing. Is it? And there 's another wee b
 m? The guy 's always playing! He 's changing, changing some every bloody minute
 now, it 's an area, it 's er, it 's changing, you know, all the time. Erm, we a
 ucation at schools and look like it 's changing a lot cos there 's a lot more to do
 sed are you? The way the weather it 's changing No I t 's very changey Well it 's
 that she 's from Hill And she 's changing address And she want 's to change o
 een working together. I think that 's changing very much. We 've seen in a number
 red and forty pounds per year? That 's changing the question slightly, so, let's g
 ell it was yesterday I had her. I was changing her nappy for her. Well she started
 This tells us water temperatures were changing and we believe this represents a maj
 when you shouted earlier when we were changing the room round? Soon as you said yo

Figure 49. Extract from a BNC_spoken concordance of *changing*, illustrating a comparatively low share of PastProgs

've done it, I think someone er who 's walking along the road, see you standing t
 and I 'd better get moving. Was she walking in? No, they went on the bus this i
 . Where were you then when Carole was walking up the street? I did n't see you. I
 the last one people out. Everyone was walking, you just holding on your like a w
 ession they were passing notes and I was walking past dropping notes in his bag and s
 matter of fact I the other day, I was walking up and he called me, so I took me f
 ctually break into the flat? No, I was walking up the ramps and it was ripped out o
 oody It was quite funny cos I, I was walking down the road with my, my mate Robb
 sergeant majors in the Signals, I was walking through shed four, one day, I was d
 or people you know, came loose, I was walking through the shed and these two, I sa
 were n't going to Tramps I was, I was walking around going oh boo hoo boo hoo I
 The monster from the marsh. When I was walking past the marsh one, one night, when
 up a bit. I saw him once. Who? I was walking and he was with his mum and I looked
 reby been born! One day, Mildred was walking along in the park it was very, very
 yesterday afternoon as the teenager was walking along a footpath through Hospital
 yesterday afternoon as the teenager was walking along a footpath through Hospital
 e, that was in the other day, that was walking around, he seemed to of had his tr
 hat they were doing, the old vixen was walking them higher up into the mountain for
 she 's had a, a some old chap who was walking around here, she just ran past him
 y! deaf! It finishes at nine, are we walking back or are you picking us up? Du n
 on there. It finishes at nine are we walking back or are you picking us up? Oh,
 r what? Sorry? What? Are we we are walking back or are you giving us a lift? W
 there were all ready to go and were walking with the press into the massive grea
 it 's racist. If I 'd a, if I were walking down the street going Nigger! sorry!
 efore the two lovely colds? And we were walking up and down on the beach at Aldeburg
 at direction. Yes. Fine. If if we were walking down there now I think I could proba
 ng the daytime, so much so that we were walking about in just T-shirts er and all th
 ed the back door? I thought we were walking? No, I thought we were walking?
 we not? Ok. Pardon? I thought we were walking Well, do you want to walk or do yo
 , I was thinking about the way you were walking Are you talking to yourself? So you
 did actually aye. Ah, when you were walking about and said I 'm trying this out
 not falling off you see. When you were walking the rope, there 's so many of you w
 at. Oh my God! Come here are you walking that way. Oh baby it 's Too much w
 for like half an hour. Where are you walking though? I du n no. Just I need

Figure 50. Extract from a BNC_spoken concordance of *walking*, illustrating a comparatively high share of PastProgs

telling (51.25%), wondering (43.93%), and working (51.72%). These types favour PastProg and PresPerfProg forms more than other verbs and are hence typical past time reference verbs (cf. also Section 4.5.9).

One of the verb forms that often occur in the present perfect progressive is *meaning* with a share of 21.62%. The concordance extract in Figure 51 displays a number of PresPerfProg tokens, in particular occurrences of the phrase “I’ve been meaning” followed by a to-infinitive. What the phrase conveys is a meaning of politeness or tentativeness (cf. Section 4.5.11). An affinity for PresPerfProg constructions was also found with *explaining* (16.67%), *hearing* (23.88%), *running* (15.63%), *saving* (17.54%), and *spending* (9.09%; cf. Table 21, fourth column). If we take into account that the average share of PresPerfProg lies at only 3.66%, these values are certainly considerable and clearly hint at a dependency between individual lexical items and progressive tense forms. These findings differ from those of Schlüter (2002a: 278) who lists *waiting*, *doing*, *looking*, and *having* as the most frequent PresPerfProg verbs. However, his results are based on a relatively small set of spoken and written subcorpora (samples from the CEC, LOB, and Brown corpora⁷³), and absolute numbers of occurrence are altogether rather low (between 3 and 8 tokens).

got in here screw hole Oh I 've been meaning to ask and I keep forgetting. I
me I pass I fume inwardly. I 've been meaning to write to The Post letter page a
weeks. Yeah. Well done. I 've been meaning to offload those on Give us a five
him! We let him out. No, I 've been meaning to. Come on Mark! See you later!
I can get. Yeah. Cos I 've been meaning to get myself some. does it round
er that 's something else we 'd been meaning to do. If you had made a will whc
Have you? All the jobs you 've been meaning to do Yeah, yeah Yes I know Hell
ng around since Christmas. What? Been meaning to give them to me. Ever so slimm
up ready to go. Ah but I was n't just meaning on the concrete bit. I was meani
f you use the word conc you 're just meaning your nose. Conical. Ah now we 'r
Our suggestion box , yes. You just meaning that someone else is gon na sort t
it was it. It 's more the stooks I 'm meaning, you know with the Oh yes you had
ke out in the summer. Mm. I was n't meaning that motorbike in any case. That
l No. No, no, no, not I was, I was n't meaning you to go up in single numbers I w
the way you were saying it I was n't meaning to suggest it was either good or
if we can . No I was n't, I was n't meaning that. No, well except that it wou
he God created the world but he 's not meaning it in that respect is he? It just
ow they make wreaths just like you 're meaning in Germany f at Easter. Aha. A wc
ght do something. I know what you 're meaning. Cos you du n no what Alan might
ut to see people, is this what you 're meaning? Yeah. Once you 've go we what w
I did n't really mean yes I was really meaning we might get some answers that are
nety erm that the community charge was meaning wa was bringing about a change whe
st meaning on the concrete bit. I was meaning on the soil. There 's loads and l
s an old sounding name . Mhm. I was meaning to ask about the would you you wou
C E D G I thing. Well. I mean I was meaning I want to, to get your standard d
g you to go up in single numbers I was meaning, go up in hundreds. Start off wi
ee than ninety two. That 's what I was meaning. So that does n't include any red
s. Yes. Yeah. Yeah. That 's what I was meaning. Er er ca can can I digress a lit

Figure 51. Extract from a BNC_spoken concordance of *meaning*, illustrating a comparatively high share of PresPerfProgs

Paul on one occasion, and er he 'd been eating with Gentiles and er then some Jews
and get for themselves that they 'd been eating of late, but here every thing is th
: of mind No He came in and he 'd been playing with this bloody lighter yeah, wel
on extra work. Yeah. Cos it had been playing on her mind for weeks. Mm. So she
roduced it, it 's something you 'd been wanting to do was n't it? something we cons
en wanting to do, not we, you had been wanting to yes I, I wanted to introduce a c
at er no, no, something you had been wanting to do, not we, you had been wantin

Figure 52. Examples of PastPerfProgs from BNC_spoken concordances of *eating*, *playing*, and *wanting*

Percentages of PastPerfProgs are generally very low. However, while the average value is only 0.64%, some verb forms, such as *eating*, *living*, *playing*, *seeing*, and *wanting*, show above-average shares of around 3%. Some PastPerfProg examples are given in Figure 52.

On the whole the distributions illustrated in Table 21 unquestionably point towards a certain lexis-grammar interrelatedness. The next section will discuss whether this interrelatedness is also valid for different typical subjects of progressives and individual verb forms.

4.5.3 Verbs and subjects [BNC/BoE]

In Section 4.3.3 above, I dealt with the co-selection of progressives and subjects and identified a set of commonly occurring items in subject position of progressive constructions. I noted that there is a significant progressive-personal pronoun colligation and determined the frequencies of occurrence for each occurring pronoun in the complete datasets from BNC_spoken and BoE_brspek (cf. Figure 23). It will now be discussed whether the distribution specified for the group of 9,468 progressive tokens is roughly representative of the majority of progressive types or whether many individual verbs show significant preferences for one particular type of subject. The discussion will centre on the seven personal pronouns *I*, *you*, *he*, *she*, *it*, *we*, and *they* as other subjects were found to be much less frequent.

Table 22 lists the shares of different personal pronoun subjects for each of the 99 analysed -ing forms. Particularly high ($\geq 3\%$ above average) and significantly low ($\geq 3\%$ below average) values have been shaded grey.

Table 22. The distribution of personal pronoun subjects across verb forms (verbs in alphabetical order)

verb form	subject <i>I</i> ø24.28%	subject <i>you</i> ø17.13%	subject <i>he</i> ø6.23%	subject <i>she</i> ø3.41%	subject <i>it</i> ø4.06%	subject <i>we</i> ø12.37%	subject <i>they</i> ø9.63%
accepting	7.58%	10.61%	1.52%	1.52%	0.00%	18.18%	16.67%
adding	14.75%	19.67%	4.92%	1.64%	4.92%	29.51%	8.20%
agreeing	30.77%	17.31%	7.69%	1.92%	0.00%	11.54%	3.85%
asking	29.80%	21.85%	5.96%	3.31%	0.00%	9.93%	8.61%
being	50.00%	12.50%	25.00%	0.00%	0.00%	0.00%	0.00%
becoming	3.66%	2.44%	1.83%	0.00%	18.29%	5.49%	5.49%
believing	0.00%	0.00%	33.33%	16.67%	0.00%	16.67%	0.00%
betting	0.00%	28.57%	28.57%	0.00%	0.00%	28.57%	0.00%
bothering	26.79%	8.93%	3.57%	1.79%	21.43%	1.79%	14.29%
bringing	12.22%	11.11%	7.78%	10.00%	1.11%	7.78%	22.22%
buying	23.28%	33.62%	4.31%	4.31%	0.00%	8.62%	11.21%
calling	21.77%	13.71%	4.84%	1.61%	1.61%	10.48%	8.87%
carrying	16.82%	12.15%	19.63%	4.67%	5.61%	8.41%	13.08%

Table 22. (continued)

changing	10.11%	5.62%	4.49%	1.12%	11.24%	8.99%	21.35%
checking	35.29%	16.47%	5.88%	1.18%	1.18%	11.76%	2.35%
coming	9.00%	9.00%	8.00%	6.00%	15.00%	5.00%	7.00%
costing	0.00%	1.57%	1.57%	0.00%	69.29%	0.79%	1.57%
cutting	15.38%	15.38%	3.08%	3.08%	6.15%	7.69%	29.23%
dealing	7.81%	20.31%	3.13%	0.00%	2.34%	28.91%	10.94%
doing	14.55%	22.42%	7.27%	0.61%	1.82%	18.18%	11.52%
drawing	20.63%	25.40%	7.94%	1.59%	6.35%	9.52%	3.17%
eating	19.09%	24.55%	12.73%	11.82%	0.00%	8.18%	9.09%
expecting	30.26%	15.35%	6.58%	8.33%	0.00%	17.54%	13.60%
explaining	39.58%	16.67%	4.17%	4.17%	0.00%	4.17%	2.08%
feeling	25.33%	37.33%	13.33%	5.33%	0.00%	2.67%	8.00%
finding	15.38%	27.47%	6.59%	1.10%	0.00%	20.88%	12.09%
finishing	27.71%	20.48%	7.23%	8.43%	3.61%	7.23%	10.84%
following	26.83%	26.83%	2.44%	4.88%	2.44%	2.44%	14.63%
forgetting	21.62%	24.32%	8.11%	5.41%	0.00%	21.62%	13.51%
getting	20.83%	17.36%	4.86%	3.47%	13.89%	13.19%	11.11%
giving	22.22%	19.05%	5.56%	3.17%	3.17%	13.49%	15.08%
going	22.12%	14.42%	4.81%	2.40%	7.69%	16.35%	7.21%
happening	0.00%	0.00%	0.00%	0.00%	6.41%	0.00%	0.00%
having	11.43%	20.00%	5.71%	4.29%	1.43%	25.71%	11.43%
hearing	35.82%	31.34%	1.49%	0.00%	0.00%	23.88%	1.49%
helping	11.76%	8.24%	3.53%	2.35%	10.59%	5.88%	18.82%
holding	20.62%	15.46%	11.34%	0.00%	2.06%	14.43%	8.25%
hoping	52.40%	8.73%	2.18%	1.75%	0.00%	24.89%	3.06%
imagining	45.45%	18.18%	22.73%	0.00%	0.00%	9.09%	4.55%
keeping	11.11%	22.22%	3.70%	8.64%	3.70%	13.58%	19.75%
knowing	0.00%	0.00%	0.00%	0.00%	0.00%	100.0%	0.00%
learning	30.30%	24.24%	4.55%	4.55%	0.00%	10.61%	9.09%
leaving	34.55%	13.64%	3.64%	8.18%	0.91%	9.09%	5.45%
letting	22.92%	15.63%	9.38%	7.29%	0.00%	6.25%	19.79%
liking	62.50%	0.00%	12.50%	12.50%	0.00%	0.00%	12.50%
listening	35.90%	34.62%	5.77%	3.21%	0.00%	3.21%	3.21%
living	12.99%	24.68%	6.49%	9.09%	0.00%	9.09%	14.29%
looking	21.92%	28.77%	4.11%	0.68%	1.37%	28.08%	7.53%
making	21.15%	16.35%	2.88%	2.88%	6.73%	8.65%	16.35%
meaning	62.16%	18.92%	2.70%	5.41%	0.00%	2.70%	0.00%
meeting	15.38%	7.69%	0.00%	0.00%	7.69%	38.46%	15.38%
moving	10.99%	13.19%	4.40%	2.20%	7.69%	16.48%	19.78%
needing	14.29%	33.33%	0.00%	0.00%	2.38%	26.19%	2.38%
paying	18.67%	21.33%	10.67%	3.33%	0.00%	15.33%	15.33%
picking	20.72%	16.22%	6.31%	2.70%	9.91%	9.01%	10.81%
playing	17.27%	22.73%	13.64%	2.73%	1.82%	10.91%	9.09%
providing	5.08%	10.17%	0.00%	1.69%	0.00%	35.59%	10.17%
pulling	12.12%	8.08%	17.17%	1.01%	6.06%	6.06%	13.13%
putting	24.75%	13.86%	9.90%	1.98%	0.00%	12.87%	17.82%

Table 22. (continued)

reading	54.17%	19.44%	2.78%	2.78%	0.00%	5.56%	4.17%
remembering	40.00%	26.67%	6.67%	0.00%	0.00%	3.33%	10.00%
ringing	44.44%	11.90%	6.35%	2.38%	4.76%	2.38%	6.35%
running	10.42%	7.29%	5.21%	7.29%	7.29%	17.71%	7.29%
saving	14.04%	26.32%	12.28%	7.02%	3.51%	12.28%	10.53%
saying	32.03%	25.49%	4.58%	1.96%	2.61%	7.84%	6.54%
seeing	29.33%	25.33%	4.00%	2.67%	0.00%	24.00%	8.00%
seeming	0.00%	0.00%	50.00%	0.00%	0.00%	0.00%	0.00%
selling	9.62%	12.50%	12.50%	1.92%	0.96%	6.73%	30.77%
sending	17.65%	15.69%	4.90%	3.92%	3.92%	16.67%	18.63%
setting	12.96%	24.07%	9.26%	1.85%	0.00%	20.37%	14.81%
showing	13.13%	7.07%	10.10%	6.06%	13.13%	4.04%	13.13%
sitting	22.40%	14.40%	10.40%	6.40%	0.80%	8.80%	10.40%
sorting	33.33%	3.03%	15.15%	0.00%	0.00%	24.24%	9.09%
speaking	30.00%	21.43%	12.86%	4.29%	1.43%	1.43%	7.14%
spending	6.49%	22.08%	3.90%	1.30%	0.00%	32.47%	14.29%
standing	19.35%	13.98%	8.60%	3.23%	6.45%	7.53%	5.38%
starting	18.18%	9.09%	5.45%	3.64%	4.55%	18.18%	10.91%
staying	25.19%	17.78%	7.41%	8.15%	1.48%	12.59%	6.67%
stopping	19.44%	14.81%	4.63%	5.56%	10.19%	12.96%	5.56%
suggesting	42.60%	22.87%	2.69%	0.45%	1.35%	7.62%	2.24%
supporting	10.14%	11.59%	1.45%	0.00%	2.90%	17.39%	8.70%
supposing	0.00%	0.00%	100.0%	0.00%	0.00%	0.00%	0.00%
taking	20.00%	13.33%	8.33%	4.17%	3.33%	13.33%	11.67%
talking	19.15%	20.21%	4.26%	2.13%	0.00%	36.70%	8.51%
telling	33.13%	18.75%	10.63%	6.25%	1.25%	2.50%	6.88%
thinking	59.57%	17.02%	2.13%	0.71%	0.00%	9.22%	2.84%
trying	29.35%	13.04%	3.80%	2.17%	0.54%	21.20%	14.13%
turning	20.00%	15.00%	7.50%	5.00%	11.25%	2.50%	8.75%
understanding	22.22%	55.56%	0.00%	0.00%	0.00%	0.00%	11.11%
using	12.09%	21.98%	4.40%	1.10%	0.00%	23.08%	16.48%
walking	23.71%	18.56%	10.31%	5.15%	0.00%	19.59%	4.12%
wanting	20.72%	28.83%	8.11%	2.70%	0.00%	5.41%	18.02%
watching	38.60%	22.81%	5.26%	3.51%	0.00%	11.40%	6.14%
wearing	29.53%	20.81%	14.09%	10.07%	0.00%	2.01%	11.41%
winning	10.00%	20.00%	4.00%	0.00%	0.00%	18.00%	8.00%
wondering	83.68%	7.11%	0.42%	1.26%	0.00%	4.60%	0.84%
working	15.52%	19.83%	6.90%	6.90%	5.17%	12.07%	11.21%
worrying	16.67%	23.33%	6.67%	5.00%	1.67%	10.00%	1.67%
writing	24.19%	30.65%	11.29%	6.45%	0.00%	8.06%	8.06%

If we now look at the column listing percentages for the subject *I*, we instantly note a large number of highlighted figures which differ significantly from the average of 24.28%. Only 16 out of 99 types show about-average shares. Particularly common is the co-selection of the subject *I* and the verb forms *hoping* (52.40%), *liking* (62.50%), *meaning* (62.16%), *reading* (54.17%), and *thinking* (59.57%). By far the highest share

I heard what you said! I 'm just wondering why why Hazel, Hazel thinks you sh
 used to have . I know, but I 'm just wondering how it 's got on there. Probably w
 Yeah I no it do n't hurt. I 'm just wondering what waxing does though. They 've
 's a vole? It could be, or I 'm just wondering if it 's actually squirrels, beca
 women to have a child and, I 'm just wondering how many of us are thinking about
 ae us going a full circle and I 'm just wondering when he 's gon if I 'll be alive w
 erm of of this erm session, I 'm just wondering if we could have ten minutes to r
 at nineteen ninety one. And I 'm just wondering whether er the er point at ninetee
 erm Madam Deputy Speaker. I 'm just wondering er the statutory instruments we 'r
 phrases, erm, not an asset, I 'm just wondering what the title below that, that o
 y Divisional Manager West. I 'm just wondering whether one or two things that 's
 are again does n't it ? Yeah I 'm just wondering if there that 's related to that s
 produce as many conkers. So I 'm just wondering whether this is a different variet
 me all night Doctor, Aye. I 'm just wondering if it 's my teeth or that it 's ju
 ay all tend to go now. Erm. I 'm just wondering th I mean how how how do you regar
 ed to contact, to find out I was just wondering whether it 's something, you know
 to leave that alone. Right I was just wondering if they would i if they would like
 top entrance. Well and er I was just wondering we we ought to do something, I me
 .. has gone up. Oh Sorry, I was just wondering if you actually asleep or not. Wan
 own entertainment you know. I was just wondering about well about that. I was also
 'e did you try? Foxes. Mm I was just wondering erm Is there anywhere else? Windb
 .. It 's just erm I 'm I w I was just wondering if you wanted to write them a lett
 'm going to show her! But I was just wondering where you was gon na put it. And I
 : they? So that is a possi I was just wondering which one it was cos I say I know
 ; well Yes and I think its I was just wondering what you very much the same its, t
 :rse twelve as I mentioned, I was just wondering if some, if that 's got something
 Oh yes I appreciate but I were just wondering what branches they were in, you
 laws, purposes and principals and I 'm wondering whether erm, that, that is why I
 ant to have a working party, but I 'm wondering if we should lay down some sort of
 Ken 's on Thursday. Fine. But I 'm wondering how to go on it being election day
 ig? Eh? What you doing? Oh crash I 'm wondering if, if that 's what David needs
 ish I would n't, you know, Yeah. I 'm wondering if I 'll be a bit short. Well, wel
 le are away on holiday then, er I 'm wondering if th anyone anticipates any parti
 set it out one way or the other. I 'm wondering whether this might be an exception
 em they have to be very careful, I 'm wondering if they put the right size in, exc
 Steven who 's Debbie 's husband, I 'm wondering if they knew that they had differe
 a bit too minor to report. Erm, I 'm wondering maybe due to the small amount of i
 I take a colossal amount of time. I 'm wondering how you 'll manage to fit it in wi
 : surprised. Probably. Mm. Mm. I 'm wondering what made them change their mind o
 that 's what I think it, it 's, I 'm wondering now if the nurses today get the sa
 h! Richard. What you got? And I 'm wondering if I should ring them actually, f
 I do, but it never mentions that. I 'm wondering whether to go the A one way, or Lo
 in connection with the cross and I was wondering what the reasoning was which took
 tile ago, Brusdeximit Mhm. And I was wondering maybe if I 'd be better going back
 's twin city in France, and as I was wondering through near to the cathedral in S
 's quite self explanatory because I was wondering where you got the business and th
 : care of the offspring, because I was wondering if erm if probability like the bas
 er whether you would like to come I was wondering whether you would like to come and
 out, you know. Yes. Erm and er I was wondering if Freud did that or if that 's wh
 le Mill blah di blah di blah, er I was wondering if it was possib oh yeah that 's c
 ad males parental investment erm I was wondering how they, how can a male when the
 of ways of doing it. Yes. Erm I was wondering if introducing that might make a y

Figure 53. Extracts from a BNC_spoken concordance of *wondering*, illustrating a high share of the personal pronoun *I* in subject position

of first person singular pronoun subjects, however, was found with *wondering*. Of 239 progressive concordance lines 200, i.e. 83.68%, contain *I* in subject position. The extracts from a BNC_spoken concordance of *wondering* in Figure 53 nicely display the most salient repeated clusters “I’m (just) wondering” and “I was (just) wondering”. The shade of meaning expressed here is clearly politeness or softening (cf. also 4.5.11 below). The adverbial “just” serves to enhance this function.

An example of a verb which shows a rather low percentage of *I* as subject is given in Figure 54. In the displayed concordance of *helping* we only find a very small number of progressives with a first person singular pronoun subject. Apparently, it is unusual to stress the fact that we “are helping” someone but comparatively common to refer to the help provided by other people.

The next personal pronoun, *you*, also shows a lot of variation in its distribution across verbs. Some verbs, such as *becoming* (2.44%), *changing* (5.62%), *showing* (7.07%), or *sorting* (3.03%), very rarely combine with a second person pronoun sub-

with the way the dividend stamps are 're working full time and they are come on, we 're helping you, you are is from Associates and Mike has been Trading Standards Officers have been I 've been busy today. David 's been you into a human being. I 've been have done? Well they, they were er The tractor and trailer is generally oing are n't they and they 're Mm. and education. Save The Children is ah. in effect what you 're doing is ys in which the British government is nd when the low level of inflation is to help him now. I mean Kevin is is o remain, but the other side of it is and look at the thousand which she is is one of the major factors which is ling fifteen thousand foxes a year is rdshire er doing surgery. I was just ok, you 're No, I 'm not, I 'm ost ten thousand pounds. That ain't ards. That fine but it 's not ds for getting the idea that they 're e them? Well the er they they 're r vigilantes because I think they 're atrolling them? Well I think they 're er to understand how much we 're twenty five per cent of that, we 'm saying to Graham, come on, we 're and deputy heads of schools, we 're re working in the shop er and you I 'm not gon na deny the fact you 're you have to do when you when you 're this drug saying it 's really really very long waiting list I was helping raty for the work this Committee 's end 's, come and watch them and he 's ve a very helpful husband, but he 's helping the hospitals, which is a, a very, helping with the children in the house, I n helping us obviously, I 'm not gon na deny helping the, the Q A team on a part time ba helping the police have made a large number helping me set up the I 've, I 've just boug helping everybody who 's took that flu, and helping the electricians and er but they we helping moving between cranes loading back a helping. Yeah. Smashing. Right and helping families to rebuild the ghettos, to helping us out with the cost that we incurre helping to meet the challenges of under deve helping them keep down the cost of providing helping him and he 's he 's checking a certa helping people to change and I have to say u helping and saying, no love is never wasted helping them to reduce their management turr helping to control foxes? Up there. I hunt helping out and erm learning the basics and helping you understand I 'm on about Lister helping to do things for children. All thes helping I think. Milvia now! We 're on the helping children. yes, then there 's the, helping me, Yeah. But they 're apt to make helping to protect society. Home Secretary helping us erm possibly if this system had t helping them through this. and appreciate t helping. The other thing of course, is leav helping you, you are helping us obviously, helping schools become more adept at self n helping perhaps with the slaughtering and t helping us with our research. Yeah? Yeah. E helping to run a club, you have to commit y helping them and you 're telling them it 's helping someone with an entrance exam for helping, but it should look in a wider issu helping me cleaning the house and everything helping, he sees him himself as helping me

Figure 54. Extract from a BNC_spoken concordance of *helping*, illustrating a low share of the personal pronoun *I* in subject position

ject (cf. Figure 55), whereas others are rather common in *you* BE *V-ing* clusters. Highest percentages of *you*-progressive collocations were found with *buying* (33.62%), *feeling* (37.33%), *listening* (34.62%), *looking* (28.77%), *needing* (33.33%), *understanding* (55.56%), and *writing* (30.65%). Figure 56 presents a number of examples of the typical “you’re looking” cluster. An additional pattern that stands out visually in the concordance extract is the collocation of “you’re looking” with *if* or *when*. Quite often, *looking*-progressives are used to state general conditions, as e.g. in lines 2 to 6 and lines 13 to 17 of the concordance in Figure 56.

Values for the third person singular pronouns *he*, *she*, and *it* range from 0.00% to 69.29%, although the average percentages lie between 3.41% (*she*) and 6.23% (*he*). Rather frequent with the male personal pronoun are progressives of the verb forms *carrying* (19.63%, cf. Figure 57), *pulling* (17.17%), *playing* (13.64%), and *wearing* (14.09%). *Being*, *believing*, *betting*, and *seeming* show even higher percentages (between 25% and 50%) but very low absolute numbers. A larger set of progressive concordance lines of these types would be required to see whether the particular co-selection of verb form and subject is significant.

Clearly above-average shares of *she* in subject position were found, for instance, with *eating* (11.82%), *expecting* (8.33%), *living* (9.09%), and *wearing* (10.07%). The concordance sample in Figure 58 illustrates that *expecting* occurs repeatedly in a pregnancy context, which of course normally requires a female subject. Figure 59 displays

carrying it. Here it is. I 'm always sorting out people 's computers for them. Ah' said what 's this here? Said I 've been sorting the notes out. I 've put them in on both, I do n't really know who was erm sorting that side of it out. You see I am, ' Liverpool. At the moment we 're er just sorting out the price of things, we think i' ose ranges. So that spreadsheet is just sorting them all way across the bottom and a ourself ready for I 'm just I 'm just sorting this out. What was it? This is what res round your neck? Because I 'm sorting them out Josephinel Oh yeah! Come o in she said I 'll treat it, we 're now sorting it out today what we 're doing. Mm. ' ou take out a box of stuff and they 're sorting through the stuff and you 're like, ' are, I are. Alright then, and we 're sorting this out when she 's gone out, when ep hold of it. Erm, we 're there we 're sorting that out the other thing is you see urrency and you 're sort the way we 're sorting it out is that we pay a quarter and n be there to confuse you while you 're sorting out the time. And when you 're sorti He was really happy and he, he, he 's sorting himself out. He has n't smoked, he, ut and ring you back tell him. He 's sorting it out. He 's getting it now. Cos I ' e 's meant to get eleven grand so he 's sorting out that. Is it? . Yeah, they 're, ' ou that letter yet? It 's Des what 's sorting it out. I forgot Oh! to ask him for then I got talking with gaffer I were sorting this job out Mm. the crafty sod is, e until I sort of, thought while I was sorting things out. Yeah, yeah. The bathroo s there was that er you see, they were sorting out Yeah. and I was blue with that

Figure 55. Extract from a BNC_spoken concordance of *sorting*, illustrating a low share of the personal pronoun *you* in subject position

; three or whatever. Yeah. And you 're looking and what, what, what relevance ' up with right angles first. If you 're looking at triangles you 're first thing y the imagination but it 's one if you 're looking desperately for an equalizer you ' it becomes more important, if you 're looking at sort of two hundred three hundr out your priorities. Erm say if you 're looking at erm say if and you think well I l group in this case. So And it you 're looking at this chain here. So, that 's in, and you say, oh from where you 're looking at it, it 's six rows with two ir useful. Erm Or to do the projects you 're looking up things. Right. What projects, again was? Yeah. Yeah. You 're you 're looking at the carbonate trying to work o 'ess people when you go back. So you 're looking for attention probably more than a / look after you. They do. Well you 're looking very well. Aye well I do n't whel ' problem was there. If that 's what you 're looking for, Yeah well that 's obviously ; obviously important and er when you 're looking at retirement I mean it 's always ' Basically. And therefore when you 're looking at the nutritional value you need But I would not see it as when you 're looking along the because that 's the first ink a golden rule must be, when you 're looking at financial er matters is er mak copy into the memory. So when you 're looking at a file on the screen, what you n. circumference. Good. So you you 're looking at It 's not a question of erm do where you 're going and what 's, you 're looking at those figures yourself anyway. you 're looking at it as two. You 're looking at it as one-to-one. If that 's a

Figure 56. Extract from a BNC_spoken concordance of *looking*, illustrating a high share of the personal pronoun *you* in subject position

roger the new He 's doing it now. he 's carrying he is in the post at the moment do ' rgerous amounts of it. Tests show he 's carrying ten times the acceptable level of m - And he do n't know what weight he 's carrying, the lad do n't. Got me? Mm mm. T you 've got ta declare what weight he 's carrying, whether he carries five, three, of this horse, and he 'd say oh, it 's carrying too much money. Yeah. Too much of n the mountains Mm. to the sea. It 's carrying with it a lot of soil, bits of rock do n't know what it is, but it oh it 's carrying me down, I know that much and I ha lengths, obviously depends on what it 's carrying but er how how does it compare with rou 're on Chapter eleven you see she 's carrying on, getting on with it. Perhaps if 's mum 's fridge? I do n't see what 's carrying or any of the colour Can I ride? C Second World War number two. Was she carrying anything at the time? Yeah she was still carrying, the arteries are still carrying oxygen charged blood, yes, becau lmonary artery, however, they are still carrying, the arteries are still carrying o re way you normally expect it, is still carrying the old rotten, you know, used up to monitor the situation we 're still carrying on inquiries at er at schools we 'r t 's slowed down a lot, it 's usually carrying very fine silt, Mm. which it drops behind ee he had the brains, and was carrying it out for him, cos erm this erm I t the first man into the flat and he was carrying the shield that has been mentioned it into that room first? P C as he was carrying the shield and afforded all the pro was still in the holster because he was carrying the shield and it was only as er I at weapon was he carrying? Er, he was carrying a shotgun and his revolver was in t ' reports, when he was seized, he was carrying an explosive device. Scotland Yard P C. Was he carrying a shield? He was carrying a shield as well. The fourth perso

Figure 57. Extract from a BNC_spoken concordance of *carrying*, illustrating a high share of the personal pronoun *he* in subject position

married, she 's got one son and she 's expecting another baby any day now. Er she 's
ate, it 's got to be on because she 's expecting it to be done straight away on the c
a lad. Well she said herself she 's expecting a lad though is n't she? Oh that 's
en to try for a natural birth. She 's expecting her second child in three weeks and
's oldest daughter. Oh! Well she 's expecting is n't she? Yeah. Yeah. Whatever
Charlie said he David oh well she 's expecting in August What What I come down
police were called. Mrs Bolger who 's expecting a baby next month has not been in co
eighteen month old son and his wife 's expecting another child next month. Erm, thri
ily in er two years time, or are they expecting not to be around in two years time? :

Figure 58. Extract from a BNC_spoken concordance of *expecting*, illustrating a high share of the personal pronoun *she* in subject position

strongly again there now. It 's not costing you money? It 's costing us a bit
y for a spirit. Mm. I mean it 's not costing two pound to send that! Well you h
or, two or three drinks to us is only costing like a quid, well it 's only a quid
ound the world ticket? Well it 's only costing Eddie 's brother a hundred erm one
d what good value it was, it was only costing them hundred and well I really ca
e 've paid for it. But it 's probably costing us about er, three quarters of a m
one, but lets recognise that they 're costing the police authority one point six
ell, I do n't suppose this contract 's costing us twenty-one million, but it 's c
d costing you millions? Yeah, he 's costing me fucking money! I 'd rather be w
any thing else or else this holiday 's costing me a fortune. I think you can m
ing from all over the county and it 's costing us a fortune. What But you say so
s costing millions of pounds and it 's costing thousands of pounds while he 's sat
us twenty-one million, but it 's costing us several million, and I would ha
one hundred pounds. Only! But it 's costing Eddie about nine hundred pound, an
ainst his duty as a trustee cos it 's costing the pension fund money this er
hat What what is that costing? It 's costing nothing, the day. It 's just the
bank would do because of course it 's costing us that from from the bank on our o
Oh right yeah. Yeah. So it 's, it 's costing us yes it is costing us, yeah. time
where she 's goes to school. It 's costing her a lot of money. Oh she does n
na say to him in future, it 's costing me four quid You know Linda? Linda?
on the National Health Service, it 's costing a fortune. H R T will avoid these
the petrol? I said, you know, it 's costing fifteen pound a week in a, a least
d he 's planning all this damn, it 's costing millions of pounds and it 's costin
ss is sweeping across industry. It 's costing the country more than seven billion
mployee per year to businesses. It 's costing them a thousand pounds per employee
d house. That 's some choice. It 's costing half a million pounds this year to
we 've approved and Long may it 's costing you. Er, but the question that I
. It 's not costing you money? It 's costing us a bit of money but er, you know,
It must have been. See how much it 's costing? Two hundred and fifty pounds a ni
of that have n't they? Otherwise it 's costing you lot over at Southwold a lot of
lagging circulation. Rivals say it 's costing the Murdoch empire millions, when
he vet of course it 's serious it 's costing me two and a half grand at least Eh
set out on this er thing. So it 's costing us approximately twenty pence for e
down to one Yeah and er, then it 's costing er three bags of coal every week Ye
stly spare time to go down there It 's costing you money is n't it? Well yeah And
, it it were twenty two Well it 's costing us it was costing us two pound fift
n't found out money- wise what it 's costing ya? No, not really I should imagi
, so we know what, we know what it 's costing them Yeah. that 's not a problem.
. How do you like to hear what it 's costing now. The old fashioned way was the

Figure 59. Extract from a BNC_spoken concordance of *costing*, illustrating a high share of the personal pronoun *it* in subject position

an example of a verb form with a strong affinity for *it* in subject position: *costing*. With a share of 69.29%, *costing* is certainly an exceptional case, but a few other verbs also show above-average percentages of *it*-progressive collocations. Frequent patterns are “it’s/is becoming”, “it’s (not) bothering”, “it’s getting”, and “it is showing”.

Like their singular counterparts, the two plural personal pronoun subjects *we* and *they* are not very evenly distributed across the 99 verb forms either. Shares of *we* in subject position vary from less than two per cent (e.g. with *bringing*, *costing*, *speaking*) to over 30% (e.g. with *providing*, *spending*, *talking*). Part of a BNC_spoken concordance of *talking* is displayed in Figure 60. It visualises the evident pattern “we’re talking about” which is very common in spoken British English. While there are also a few instances of the long form “we are talking” in the same concordance, the contracted form of TO BE is much more frequent in this particular context.

when you die. Of course although we 're talking about dying and nobody wants to thi
 d around the very big clubs, and we 're talking less here about what 's gon na hap
 lking about defensive driving and we 're talking about safety, which is the main ob
 ting the area of the architecture we 're talking about in the section. So right now
 an hour. Yeah. Is this copying we 're talking about? Yeah. This is, yeah. Yeah
 't we?? Yeah. You know? Er cos we 're talking about, you see we 're talking abou
 m. we all can get depressed if we 're talking about a mood state, that 's not an
 borah and I. In that we 've, if we 're talking division we 've got two full time c
 by telling me which pond it is we 're talking about? Yes. Because I just want
 ow is unbelievable we are. like we 're talking round it now. Hugh. like you kn
 of the mass of the Chinese people we 're talking about erm I do n't think I, I rathe
 na look at this graph the period we 're talking about is indicated with these dotte
 a idea that er all over the piece we 're talking about non-modernized railways simpl
 ask another question tha, since we 're talking about erm er, advances er which at
 ?? Yeah. Probably. Yes. So we 're talking about twenty Twenty four twenty fiv
 id three, at age seventy six, so we 're talking about the turn of the century. They
 the nature of the church. And so we 're talking not just about sacramental communio
 'm sorry to say but I think that we 're talking about this education day at the beg
 to build something better, then we 're talking about a project that takes place ov
 and use it for decision support. We 're talking about integrating both of those dis
 P points throughout the county. We 're talking about East Suffolk of course Yes. n
 thousand About two. Well. No. We 're talking about a thousand covers, cos the a
 are not taking the G M B over. We 're talking about an amalgamation. Yes, look
 v. Cos we 've always well, we 're talking about the choral society, and sayi
 re applications. And that 's what we 're talking about by transaction integrity log
 away the very principle of what we 're talking about, the block vote, especially
 of the mind, which is really what we 're talking about today, was erm a kind of hyd
 at I would, really a lot of what we 're talking about here is actually male chauvin
 jr needs are but that 's not what we 're talking about. That 's not what parish co
 c any privatization, in that year we 're talking about immediately, is it not? No i

Figure 60. Extract from a BNC_spoken concordance of *talking*, illustrating a high share of the personal pronoun *we* in subject position

enter the secondary schools and they 're selling it to these young kids that just
 're in all directions. Mm. and they 're selling them off, gold, gold prices hav
 That makes no sense, because they 're selling that product abroad, the farm pr
 ing them in Guildford Marks. If they 're selling those imitation trees, and if th
 die a handle on to the tree If they 're selling them in Guildford Marks. If they
 id they 're losing so much money they 're selling it. The thing is nobody wants to
 ds growing among the plants they 're selling. Messy. Yes, messy. And it was
 the earths it comes out of so they 're selling it by the bottle in Germany. Mm.
 yeah but it does n't mean that they 're selling out completely does it because it
 wn manufacturing operation where they 're selling to outside customers there 's r

Figure 61. Extract from a BNC_spoken concordance of *selling*, illustrating a high share of the personal pronoun *they* in subject position

They as subject was often found in the concordances of *bringing* (22.22%), *chang-*
ing (21.35%), *cutting* (29.23%), and *selling* (30.77%; cf. Figure 61 for some instances
 of “they’re selling” in context). Again, a large number of verbs show percentages
 which clearly deviate from the determined average of 9.63%. We can hence say that,
 with respect to personal pronoun subjects, the examined 9,468 progressives from
 BNC_spoken and BoE_brspok certainly show some typical connections between in-
 dividual pronoun and verb form type.

4.5.4 Verbs and objects [BNC/BoE]

In Section 4.3.4 we looked at typical co-occurrences of progressives and words and
 phrases in object position. For the two groups of BNC_spoken and BoE_brspok pro-
 gressives (i.e. the 99 verb forms taken together), the following frequent objects were
 identified: *the* + noun (group) (e.g. *the job*), *a/an*+ noun (group) (e.g. *a phone call*),
it, *them*, *you*, and *me*. I will now analyse whether these items are frequent objects of

progressives in general or whether, as was suspected earlier on, individual verbs trigger certain objects and form particular progressive-object patterns.

In order to determine the distribution of the above-mentioned common items in object position, the BNC/BoE datasets in *Access* were filtered according to individual objects and sorted for verb form type. For each verb the absolute numbers of co-occurrence with frequent objects could thus be retrieved. Table 23 displays the relative frequencies of occurrence of *the* + NP, *a/an* + NP, *it*, *them*, *you*, and *me* in object position for 99 individual verbs. Significant deviations from average values are shaded grey. As the table is largely self-explanatory, I will just comment on a couple of important co-selection patterns of objects and different verbs.

Table 23. The distribution of frequently used objects across verb forms (verbs in alphabetical order)

verb form	object <i>the</i> +NP ø10.38%	object <i>a/an</i> +NP ø5.99%	object <i>it</i> ø4.37%	object <i>them</i> ø2.21%	object <i>you</i> ø2.24%	object <i>me</i> ø2.15%
accepting	21.21%	4.55%	13.64%	0.00%	1.52%	0.00%
adding	16.39%	6.56%	1.64%	4.92%	0.00%	0.00%
agreeing	11.54%	3.85%	5.77%	0.00%	9.62%	3.85%
asking	9.93%	3.31%	0.00%	1.99%	14.57%	9.27%
being	0.00%	25.00%	0.00%	0.00%	0.00%	0.00%
becoming	4.27%	16.46%	0.00%	0.00%	0.00%	0.00%
believing	0.00%	0.00%	16.67%	0.00%	0.00%	0.00%
betting	0.00%	14.29%	0.00%	14.29%	0.00%	0.00%
bothering	3.57%	1.79%	0.00%	0.00%	16.07%	19.64%
bringing	14.44%	11.11%	11.11%	10.00%	2.22%	1.11%
buying	11.21%	14.66%	4.31%	2.59%	0.86%	0.00%
calling	8.06%	6.45%	4.03%	5.65%	5.65%	6.45%
carrying	14.95%	17.76%	7.48%	0.00%	0.00%	0.93%
changing	19.10%	4.49%	3.37%	2.25%	0.00%	0.00%
checking	15.29%	4.71%	8.24%	2.35%	1.18%	0.00%
coming	4.00%	1.00%	1.00%	0.00%	0.00%	1.00%
costing	5.51%	5.51%	0.79%	11.81%	5.51%	10.24%
cutting	10.77%	0.00%	7.69%	3.08%	0.00%	1.54%
dealing	18.75%	10.16%	6.25%	3.13%	0.78%	0.78%
doing	12.73%	6.06%	8.48%	0.00%	0.00%	0.00%
drawing	20.63%	20.63%	1.59%	0.00%	1.59%	0.00%
eating	5.45%	5.45%	7.27%	1.82%	0.00%	0.00%
expecting	6.14%	12.28%	7.46%	1.75%	8.33%	3.07%
explaining	12.50%	2.08%	12.50%	6.25%	2.08%	2.08%
feeling	5.33%	8.00%	1.33%	0.00%	0.00%	0.00%
finding	13.19%	4.40%	25.27%	1.10%	0.00%	0.00%
finishing	12.05%	2.41%	7.23%	0.00%	0.00%	0.00%
following	17.07%	7.32%	2.44%	0.00%	4.88%	12.20%
forgetting	18.92%	0.00%	8.11%	0.00%	0.00%	0.00%

Table 23. (continued)

getting	6.94%	9.72%	1.39%	0.69%	0.00%	0.00%
giving	14.29%	3.17%	5.56%	11.90%	11.90%	10.32%
going	0.96%	0.00%	0.00%	0.00%	0.00%	0.00%
happening	2.14%	0.00%	0.00%	0.43%	0.00%	0.00%
having	2.86%	35.71%	2.86%	1.43%	0.00%	0.00%
hearing	13.43%	4.48%	4.48%	0.00%	0.00%	1.49%
helping	15.29%	1.18%	1.18%	7.06%	4.71%	4.71%
holding	9.28%	12.37%	9.28%	2.06%	1.03%	1.03%
hoping	1.31%	0.00%	0.00%	0.00%	0.00%	0.00%
imagining	4.55%	4.55%	18.18%	4.55%	4.55%	0.00%
keeping	7.41%	8.64%	11.11%	7.41%	4.94%	2.47%
knowing	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%
learning	9.09%	1.52%	1.52%	3.03%	0.00%	0.00%
leaving	17.27%	1.82%	7.27%	1.82%	0.91%	1.82%
letting	10.42%	1.04%	2.08%	11.46%	4.17%	3.13%
liking	12.50%	0.00%	37.50%	0.00%	0.00%	0.00%
listening	18.59%	1.92%	1.28%	1.28%	3.85%	1.28%
living	10.39%	2.60%	2.60%	1.30%	0.00%	1.30%
looking	18.49%	3.42%	3.42%	0.68%	2.05%	0.68%
making	22.12%	18.27%	6.73%	1.92%	0.96%	0.96%
meaning	10.81%	0.00%	2.70%	0.00%	2.70%	0.00%
meeting	7.69%	0.00%	0.00%	0.00%	0.00%	0.00%
moving	8.79%	1.10%	2.20%	0.00%	1.10%	1.10%
needing	14.29%	4.76%	2.38%	0.00%	0.00%	2.38%
paying	15.33%	5.33%	4.67%	1.33%	1.33%	2.00%
picking	15.32%	4.50%	7.21%	2.70%	4.50%	3.60%
playing	10.00%	8.18%	5.45%	0.00%	1.82%	2.73%
providing	13.56%	32.20%	1.69%	3.39%	0.00%	3.39%
pulling	20.20%	6.06%	9.09%	4.04%	0.00%	3.03%
putting	8.91%	10.89%	14.85%	4.95%	1.98%	1.98%
reading	13.89%	6.94%	16.67%	4.17%	1.39%	0.00%
remembering	6.67%	0.00%	0.00%	10.00%	0.00%	0.00%
ringing	5.56%	0.79%	0.00%	3.17%	2.38%	2.38%
running	6.25%	13.54%	9.38%	0.00%	0.00%	2.08%
saving	12.28%	3.51%	10.53%	5.26%	1.75%	0.00%
saying	1.96%	0.65%	1.31%	0.65%	0.65%	0.00%
seeing	21.33%	8.00%	1.33%	2.67%	0.00%	0.00%
seeming	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
selling	11.54%	8.65%	11.54%	8.65%	0.00%	0.00%
sending	9.80%	11.76%	3.92%	3.92%	2.94%	1.96%
setting	31.48%	14.81%	7.41%	0.00%	0.00%	0.00%
showing	8.08%	9.09%	4.04%	2.02%	9.09%	6.06%
sitting	6.40%	1.60%	0.80%	0.80%	0.00%	0.80%
sorting	27.27%	3.03%	18.18%	9.09%	0.00%	0.00%
speaking	17.14%	4.29%	0.00%	1.43%	1.43%	2.86%
spending	12.99%	14.29%	3.90%	1.30%	0.00%	0.00%

Table 23. (continued)

standing	11.83%	2.15%	2.15%	1.08%	0.00%	1.08%
starting	4.55%	0.91%	0.91%	0.00%	0.00%	0.00%
staying	8.89%	2.96%	0.00%	0.74%	0.00%	1.48%
stopping	5.56%	1.85%	4.63%	2.78%	5.56%	4.63%
suggesting	3.59%	1.79%	0.90%	0.00%	1.35%	0.00%
supporting	23.19%	5.80%	4.35%	4.35%	4.35%	0.00%
supposing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
taking	14.17%	7.50%	3.33%	5.00%	0.83%	2.50%
talking	20.74%	7.45%	2.13%	0.00%	0.00%	1.60%
telling	3.13%	3.13%	0.63%	5.00%	21.25%	25.63%
thinking	15.60%	2.13%	4.96%	0.00%	1.42%	0.71%
trying	0.00%	0.54%	0.54%	0.00%	0.00%	0.00%
turning	11.25%	7.50%	8.75%	0.00%	0.00%	0.00%
understanding	0.00%	11.11%	22.22%	33.33%	0.00%	0.00%
using	28.57%	5.49%	6.59%	0.00%	1.10%	0.00%
walking	19.59%	3.09%	1.03%	1.03%	0.00%	0.00%
wanting	3.60%	3.60%	3.60%	1.80%	0.00%	0.90%
watching	11.40%	7.02%	11.40%	1.75%	4.39%	1.75%
wearing	11.41%	19.46%	4.70%	4.03%	0.00%	0.00%
winning	12.00%	0.00%	4.00%	0.00%	0.00%	0.00%
wondering	1.26%	0.00%	0.00%	0.00%	0.00%	0.00%
working	9.48%	3.45%	3.45%	1.72%	0.00%	0.86%
worrying	3.33%	0.00%	6.67%	1.67%	8.33%	18.33%
writing	8.06%	14.52%	8.06%	3.23%	1.61%	1.61%

If we look at the first common object type listed in Table 23, *the* + noun (group), some verbs show considerable above-average percentages. Among the verb forms that obviously have a preference for *the* + NP-objects are *accepting*, *making*, *seeing*, *setting*, *sorting*, *supporting*, and *using*. Two short BNC_spoken concordance extracts which illustrate this particular verb-object colligation have been combined in Figure 62. Other verb forms in our list of 99 clearly trigger *a/an*+ NP objects. *Becoming*, *buying*, *carrying*, *drawing*, *having*, *making*, *providing*, *setting*, and *wearing* belong to this group (cf. Figure 63 for concordance examples).

k you . Could you ask who, that is making the recording? Are we allowed to know w
top one, safety related admin. I'm making the point that man management procedures
teenth of sixty four. And you 're making the application for temporary authority
Because that 's you know he 's he 's making the right choice making the right judger
e got paid. No. No. Lawrence was making the point that if if wh there 's a hundr
I I would n't have in I when I was making the the analysis I w was n't including h
ng that we we need to know? Are we making the most of the films that we show? Are
ie, a united economy. We 're already seeing the strains on Eur European economic mo
inder of the pits because we 're now seeing the drift towards er longer working hou
getting very elaborate, and you 're seeing the stuff in the press for things like
i the right sort of guess and you 're seeing the sort of when it gets to the awkward
that whatever. And yet we were seeing seeing the film from the Americans where the

Figure 62. Extracts from BNC_spoken concordances of *making* and *seeing*, illustrating high shares of *the* + NP in object position

t me as an individual, if I was erm, buying a Covermaster Plan, for twenty pound
 ice when I was a boy and he was maybe buying a horse and that but he was very seldom
 saying to people come on you 're not buying a raffle ticket. let's talk serious
 id n't know this one. Erm if you 're buying an endowment policy is it possible to
 nly thing is, if you say, if you 're buying a house now, you 'll need a Mortgage
 life endowment, saving if you were buying a house, though, we do a Mortgage Ma
 year er, she said that Professor is having an extra lecture you know, telling wh
 maybe ten minutes on each topic, just having a little look at it and finding a prob
 t the break! So I said, okay, I 'm having a cigarette break now. And er this wa
 e meeting was n't that long ago. I 'm having a division meeting on. So, I 'll make
 she fell on on on Sunday again. I 'm having an awful problem to herself. She was
 a look. Yeah edit that. While I 'm having a look at this just a quick glance thr
 ls, unlike little Virginia she was n't having any nonsense from the proud house. Ivy
 the private sector. And they 're not having a go, they 're not having a go, they
 tually. I mean seriously. You 're not having a lot of trouble with these are you? I
 size now like I say once again we 're having erm problems with getting erm an econo
 ne in a fortnight 's time. And we 're having a baptism. Okay let's move on. I 'm s
 he four of us. Yeah. Erm, but we 're having a big party tomorrow night, and we 'r
 for ma, pa and the wains and so we 're having a family fun day at Chatlereau country
 y in the audience that we 're, we 're having a go at. But that 's not the problem I
 am? Yes. Go on then quickly. We 're having a lose a tooth week I think. Erm rig
 ou had in which. Yeah. is why we 're having a lot of problems with this. Alright ol

Figure 63. Extracts from BNC_spoken concordances of *buying* and *having*, illustrating high shares of *a/an+* NP in object position

we have a lot. It 's very hard, I am finding it hard to keep up with them at the
 ose people who supported it before are finding it impossible to support it now. Mr.
 Min Sherice and Gojas scientists are finding an unusually high incidence of mental
 hould n't be saying this the staff are finding it very difficult I mean what I 'm tr
 matches because the Midlands teams are finding it hard to get near the top. Oh that
 tablets for Mm. Mhm. Er I 've been finding that normally I can get up early in t
 er, but then it, I think you were just finding your words, that was good, well done
 I 'm not getting any younger and I 'm finding it a struggle to try and get everythi
 it with you. It 's just that I 'm finding it well I 'm not. My wife 's getting
 I said the sum assured, sorry, I 'm reading it off here. It 's the end of the da
 wn the bottom end ? I did. But I 'm reading it again. Mhm. And it says, It 's g
 eaking? Watching Absolutely they 're reading it with you so they 're with you so y
 get to cover that side up when you 're reading it. But if you go through there to a
 actly we mean and to help when you 're reading it. So if you 're reading something
 uter, even with his glasses on, he 's reading it like this. I am concerned about
 Isalah fifty three in it and he 's reading it, and Philip goes up to the man in
 ker than perhaps some would if we were reading them on a radio or something. Well I
 about the last one ! Think you were reading them upside down! Well they only,

Figure 64. Extracts from BNC_spoken concordances of *finding* and *reading*, illustrating high shares of *it* in object position

ow, these interruptions are actually costing us quite a lot of month. Next? Oh, that
 s government policy these days erm is costing them same amount of money, they 're ac
 ifty eight. They say a set of gear is costing them five hundred and ninety five pounc
 strongly again there now. It 's not costing you money? It 's costing us a bit of r
 what good value it was, it was only costing them hundred and Well I really ca n't r
 ll, I do n't suppose this contract 's costing us twenty-one million, but it 's costi
 ployee per year to businesses. It 's costing them a thousand pounds per employee per
 costing millions of pounds and it 's costing thousands of pounds while he 's sat in
 , so we know what, we know what it 's costing them Yeah. that 's not a problem. I dc
 sy brochures out in the post. It was costing them extra in postage Mhm. and they ke
 em so much I suppose and erm that was costing them so much on wages and that, so they
 y do n't understand them but you 're understanding them I think very well now so y
 y you like them do n't you cos you 're understanding them now. It 's all about under
 y problem, okay? Yeah. So you 're understanding it. You could explain that to m
 ed quite a bit of practice. You 're understanding them now very well erm if you do

Figure 65. Extracts from BNC_spoken concordances of *costing* and *understanding*, illustrating high shares of *them* in object position

Significant also are progressive verb form colligations with the personal pronouns *it*, *them*, *you*, and *me*. Some forms (esp. *accepting*, *finding*, *imagining*, *putting*, *reading*, and *watching*) have an affinity for *it*-objects, as the concordance extracts in Figure 64 demonstrate, while other verbs, e.g. *bringing*, *costing*, *giving*, *letting*, *selling*, and *understanding* often co-occur with *them* in object position (see Figure 65).

between national curriculum No, I 'm asking you there but I 'm asking you when ' it 's a constant. Mm. Right now I 'm asking you to integrate it which is work ba e important to you. The only thing I 'm asking you to do is put them in the diary r t know. Neither do I, that 's why I 'm asking you, cos I wo n't have time to shop tside this at all cos the ones they 're asking us are inside. So you 'd get that c ime signature in but you 're, they 're asking you to put in the bar lines and ther nglish Heritage are spending, they 're asking for us to fund the shortfall. There ump in because I 'll know what they 're asking me Yeah, I 've noticed it a wee bit n back here for the even later, we 're asking you for it to be taken at the end w the information is gathered in, we 're asking you to quite simply reject this. Th eople feel that vulnerable that you 're asking them to go Mm Yes out so you just, s well that 's the sort of thing you 're asking me to do. And, and you know how to . Yeah. Oh that 's not fair, you 're asking qu, no you ask questions I know fr , Danny come up over my socks no one 's asking you are they? Oh god Cor strewth Se to withdraw. Right. C E C are therefore asking you to oppose the motion. All those you come out Aye, aye. It 's not bothering me now. No. Just when when you y bothering you, leave it alone. Aye. Said tl . No I 'm not, it 's not, it 's not bothering me No it does n't bother me. Actua it 's so new to them that they 're bothering to cost it, but how did we go on . I said, I do n't know what you 're bothering for. I said by heck you 've some ted down with something that 's really bothering you. Because you do n't think th op and had to stop and that 's what 's bothering me. Wan na be a man, man cub. Go goodness! Oh! Did you get it was bothering me, that 's why I wanted her to go to no-one for that, and if anybody is telling me that I 'm tainted because I do that n who 's telling secrets? James is telling me a secret oh tell me a secret then : Mist and Smoke. So that one really is telling a story, commemorating something in i or about what what the witness is telling you. . Mrs, first of all can I invite you a number of holes. Yeah. Just telling you. 'Shit! Some of this stuff! So tl m off as you. No I 'm not I 'm just telling you Oh right so they, they, their phy: might are you Jo ? No. No, I 'm just telling the truth what I think of him. Er alti I 'm not complaining duck, I 'm just telling you now it 's extremely cold outside, about osmosis? We do n't, I 'm just telling dad, cos he How come you know about all it says is erm it just it 's just telling you what it is and saying that you do he? I That 's it, I said I was just telling the children well what happens if any do you get that from? Er I was just telling him, they 've got big ones like that Did he? Mm. He eat it all. And I 'm telling Jackie about this, I says, d 'ya kno ai n't heard from them. Yeah but I 'm telling you where you got ta If you sent a ch is thing here. It 's just that er I 'm telling everybody. Mm. How did you find out can go through that door. Now, I 'm telling you. Except for a three year secondm will do so and so, or you know, I 'm telling you you 'll do this. Now it 's how y lu n no what you do without it, I 'm telling you! . Muggy. Er er oh, cos muggy oi u? I have not been cutting it! I 'm telling you now, I have n't! I 'm sure you w the time at school when the teacher 's telling you, why do n't you do it this way? a good laugh like so it was Donna was telling me Jamie, Donna says and he says no nterviewed him er last year and he was telling me that he remembers going up into th the door and let him in and he was telling me I 've been here half an hour tryin lement Freud Mm. talking and he was telling about a visit he had paid to China fo t costs, I mean Jimmy next door he was telling us yesterday when he come to er, we t You just suit yourself! Oh, he was telling me, he said hey! ? Start. He 's stil rd anything, so Peter rang up, he was telling me the other day, er, he rang up yea said oh they 're really nice. He was telling me that they 'd died of the frost or

Figure 66. Extracts from BNC_spoken concordances of *asking*, *bothering*, and *telling*, illustrating high shares of *you* and *me* in object position

Typical patterns with *you* and *me* as objects include “asking you”, “bothering you/me”, “costing me”, “expecting you”, “giving you/me”, “showing you”, “telling you/me”, and “worrying me”. Figure 66 presents a few concordance samples to exemplify some of these special choices of objects. It will be interesting to see later on whether these common collocates also feature in teaching materials used in the EFL classroom.

On the whole, my observations concerning verbs and objects are pretty similar to those made in the case of subject-selection and tense form distribution: individual verbs show clear preferences for particular items and do not generally follow the “normal” or average distributions determined earlier on for the whole set of progressives. It hence appears somewhat problematic to talk about a general progressive profile, at least with respect to the context features examined so far.

4.5.5 Verbs and prepositions [BNC/BoE]

In order to identify further existing patterns in the lexical context of progressives, we will now turn to examining the co-occurrences of individual progressive verb forms and frequent prepositions. As noted above (cf. 4.3.5), preposition choice is largely lexically determined, meaning that certain verbs trigger certain prepositions. Hence we do not expect that the average distributions determined in Section 4.3.5 (cf. Figure 26) are generally valid for all or at least for the majority of our examined verbs.

The aim of the present section is simply to draw the reader's attention to some of the most salient progressive-preposition collocations in spoken English. These collocations will later on be used as a basis for comparison with typical co-occurrences of prepositions and progressive verb forms in the selected teaching materials. Table 24 lists those verbs that are frequently postmodified by a preposition together with the shares of the most commonly occurring items. Fields which indicate a significant collocation are shaded grey.

As we can see in Table 24, there are a number of verb-preposition clusters around the prepositions *about* (e.g. “talking about”, “thinking about”, “worrying about”), *for* (e.g. “asking for”, “looking for”, “paying for”), *in* (e.g. “believing in”, “bringing in”, “letting in”), *on* (e.g. “carrying on”, “pulling on”), *out* (e.g. “carrying out”, “sorting out”), *up* (e.g. “adding up”, “picking up”, “setting up”), and *with* (e.g. “agreeing with”, and especially “dealing with”). Some typical collocations are exemplified in Figures 67 to 73. Not included in Table 24 are the following frequent collocates:

“bringing back”, “coming back”, “coming down”, “cutting back”, “cutting down”, “letting down”, “listening to”, “looking at”, “pulling down”, “sitting down”, “staying at”, “talking to”, “thinking of”, “walking down”, “writing down”.

I shall refer back to these clusters and their frequencies of occurrence in spoken English when I compare the results based on the textbook corpus GEFL TC with those obtained here on the basis of BNC_spoken and BoE_brspok data.

4.5.6 Verbs and negation [BNC/BoE]

Part of the contextual analysis in Chapter 4.3 was a brief discussion of the relation between progressives and negation. We found that in 7.98% of the BNC_spoken examples and 8.59% of the concordance lines from BoE_brspok the verb form is negated (cf. Figure 27). It will now be examined whether these average values roughly apply to progressives in general or whether our 99 individual verb form types differ a lot with respect to their shares of negation.

Table 25 gives percentages of negated and non-negated verb tokens for each analysed verb and includes small diagrams to visualise the percentages. As we can see in the left-hand parts of these diagrams (the light grey bars), a number of verbs in the list of 99 are particularly frequently negated while others clearly show below-average shares of negation.

Table 24. The distribution of frequently used prepositions across selected verb forms (verbs in alphabetical order)

verb form	prepos. <i>about</i> ø3.12%	prepos. <i>for</i> ø1.81%	prepos. <i>in</i> ø1.82%	prepos. <i>on</i> ø1.77%	prepos. <i>out</i> ø2.04%	prepos. <i>up</i> ø3.24%	prepos. <i>with</i> ø2.59%
adding	0.00%	0.00%	3.28%	3.28%	0.00%	21.31%	0.00%
agreeing	1.92%	0.00%	0.00%	1.92%	0.00%	0.00%	42.31%
asking	6.62%	14.57%	0.00%	0.00%	0.00%	0.00%	0.00%
believing	0.00%	0.00%	16.67%	0.00%	0.00%	0.00%	0.00%
bothering	0.00%	1.79%	0.00%	0.00%	0.00%	0.00%	12.50%
bringing	0.00%	0.00%	24.44%	2.22%	6.67%	17.78%	2.22%
calling	0.00%	12.90%	2.42%	4.03%	3.23%	0.81%	0.00%
carrying	0.00%	0.00%	0.00%	15.89%	15.89%	0.00%	0.00%
checking	1.18%	1.18%	1.18%	1.18%	7.06%	7.06%	1.18%
coming	0.00%	1.00%	9.00%	2.00%	7.00%	11.00%	1.00%
cutting	0.00%	0.00%	0.00%	0.00%	12.31%	1.54%	0.00%
dealing	0.00%	0.00%	4.69%	0.00%	0.00%	0.00%	91.41%
forgetting	10.81%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
hearing	11.94%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
helping	0.00%	0.00%	0.00%	0.00%	7.06%	0.00%	4.71%
holding	0.00%	0.00%	1.03%	4.12%	4.12%	7.22%	0.00%
keeping	0.00%	0.00%	1.23%	1.23%	0.00%	7.41%	0.00%
letting	0.00%	0.00%	18.75%	0.00%	3.13%	0.00%	0.00%
looking	0.68%	38.36%	0.00%	0.00%	0.00%	0.68%	0.00%
moving	0.00%	0.00%	2.20%	2.20%	10.99%	3.30%	2.20%
paying	0.00%	16.00%	0.00%	0.00%	0.67%	0.00%	0.00%
picking	0.00%	0.00%	0.00%	8.11%	6.31%	67.57%	0.00%
playing	0.00%	1.82%	0.00%	0.91%	0.00%	4.55%	17.27%
pulling	0.00%	0.00%	2.02%	3.03%	11.11%	6.06%	0.00%
putting	0.00%	0.00%	6.93%	18.81%	2.97%	5.94%	0.00%
reading	8.33%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
ringing	2.38%	0.00%	0.79%	0.00%	0.00%	30.95%	0.00%
saving	0.00%	0.00%	0.00%	0.00%	0.00%	15.79%	0.00%
sending	0.00%	0.98%	2.94%	0.00%	10.78%	3.92%	0.00%
setting	0.00%	0.00%	0.00%	0.00%	11.11%	57.41%	0.00%
sitting	0.00%	0.00%	9.60%	8.80%	0.00%	0.80%	3.20%
sorting	0.00%	0.00%	0.00%	0.00%	81.82%	0.00%	0.00%
speaking	10.00%	0.00%	0.00%	2.86%	0.00%	0.00%	2.86%
standing	0.00%	3.23%	8.60%	11.83%	0.00%	3.23%	0.00%
staying	0.00%	0.74%	8.15%	0.74%	2.96%	0.74%	16.30%
talking	75.53%	0.00%	0.53%	0.53%	0.00%	0.00%	0.53%
thinking	21.99%	0.00%	0.00%	0.00%	0.00%	0.71%	0.00%
turning	0.00%	0.00%	1.25%	2.50%	3.75%	10.00%	0.00%
walking	5.15%	0.00%	2.06%	0.00%	2.06%	9.28%	0.00%
worrying	41.67%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

nd three, at age seventy six, so we 're talking about the turn of the century. The
 the nature of the church. And so we 're talking not just about sacramental communi
 'm sorry to say but I think that we 're talking about this education day at the be
 to build something better, then we 're talking about a project that takes place di
 and use it for decision support. We 're talking about integrating both of those di
 P points throughout the county. We 're talking about East Suffolk of course yes. I
 thousand About two. Well. No. We 're talking about a thousand covers, cos the
 are not taking the G M B over. We 're talking about an amalgamation. Yes, look
 W. Cos we 've always . well, we 're talking about the choral society, and say
 he applications. And that 's what we 're talking about by transaction integrity lo
 e away the very principle of what we 're talking about, the block vote, especiall
 of the mind, which is really what we 're talking about today, was erm a kind of hy
 at I would, really a lot of what we 're talking about here is actually male chauvini
 ur needs are but that 's not what we 're talking about. That 's not what parish c
 t any privatization, in that year we 're talking about immediately, is it not? No.
 ividuals that have come in here? You 're talking about individuals that have come in
 something. It 's you know, if you 're talking in milliseconds and then you go to
 two amps or erm five years if you 're talking about charging up a capacitor or s
 at 's your opinis opinion I mean you 're talking about erm the er marriage from er
 ently cos you can tell which one you 're talking about then. Whether you 're talkin
 ing. There 's only if the person you 're talking to thinks you do n't care, well I
 a squad and do light jobs that you 're talking about, would Mhm. it be known the
 really You know the thing that you 're talking about that Cath is doing? You know
 's wa that 's the overall thing you 're talking about is it? No, t the actual ma
 a different war from, from what you 're talking about. Yes. Several people were
 f those words I do n't know what you 're talking about, quite honestly. Erm I do

Figure 67. Extract from a BNC_spoken concordance of *talking*, illustrating the typical progressive-preposition collocation “talking about”

it becomes more important, if you 're looking at sort of two hundred three hundred
 out your priorities. Erm say if you 're looking at erm say if and you think well I k
 group in this case. So And it you 're looking at this chain here. So, that 's c
 in, and you say, oh from where you 're looking at it, it 's six rows with two in.
 'ful. Erm Or to do the projects you 're looking up things. Right. What projects, f
 ain was? Yeah. Yeah. You 're you 're looking at the carbonate trying to work out
 is people when you go back. So you 're looking for attention probably more than any
 look after you. They do. Well you 're looking very well. Aye well I do n't whethe
 olem was there. If that 's what you 're looking for, Yeah well that 's obviously c
 obviously important and er when you 're looking at retirement I mean it 's always be
 Basically. And therefore when you 're looking at the nutritional value you need to
 But I would not see it as when you 're looking along the because that 's the first
 k a golden rule must be, when you 're looking at financial er matters is er make
 copy into the memory. So when you 're looking at a file on the screen, what you 'r
 circumference. Good. So you you 're looking at It 's not a question of erm doing
 're you 're going and what 's, you 're looking at those figures yourself anyway. I
 you 're looking at it as two. You 're looking at it as one-to-one. If that 's a f

Figure 68. Extract from a BNC_spoken concordance of *looking*, illustrating the typical progressive-preposition collocation “looking at”

out that. This government is actually cutting back on everything here just so that
 e they do today, but we were actually cutting meat, you know, we were Cutting you
 yeah. There all. A bunch. There all cutting it now, so. . Cos the Producer you
 ay with ours they 're cutting back and cutting back all the time until I mean it 's
 nearly half of all surgical units are cutting back on operations a third are givin
 robably realise at the moment they 're cutting back on that particular budget. But
 s and the, and they sort of they 're cutting down you know, on the hospital pro
 ike and ours, I say with ours they 're cutting back and cutting back all the time u
 cated un cutting back. They 're cutting back on education. Oh, you 're tell

Figure 69. Extract from a BNC_spoken concordance of *cutting*, illustrating the typical progressive-preposition collocation “cutting back”

all sitting down! I know you 're all sitting down. What are you in? This was th
 cha chasing the duck and they were all sitting down. And a there was a domestic on
 supporter of Save The Children erm I am sitting down soon beside Elise who 's travel
 ber than what the, the others that are sitting the S A T S course. At some time, y
 pregnant married at. These two are sitting down. I think that Hazel, Charlene,
 it to convey that Mr and Mr Wogan are sitting down and having a conversation or Ms
 e one before that, there was a huge guy sitting about where Colin is sitting. And b
 shed. There. Right. Thank you. Is he sitting down mum? No he 's kneeling up. my

Figure 70. Extract from a BNC_spoken concordance of *sitting*, illustrating the typical progressive-preposition collocation “sitting down”

five per cent of the premium is actually paying for the level of cover. Mm. So you 're he, the direct debits that you 've been paying for years but you 're not sure where remember, I went up when I was first er paying for paying for an annual ticket with 'has the right to use er the land he is paying for and in the, in that matter of us And I often think to myself, when I 'm paying for something, think to myself, oh I 'you not approve of that? Not if I 'm paying for it! If they, if they, if they h quarter or whatever, minutes, and I 'm paying for a lunch, you 're thinking of dir 'd think they could because you are not paying for the service are you? That 's rig

Figure 71. Extract from a BNC_spoken concordance of *paying*, illustrating the typical progressive-preposition collocation “paying for”

oing something That, I mean they 're setting up the County Working, a County Work t 's, it 's because they 're, they 're setting up child care organizations, people a rather better arrangement, they 're setting up a joint team, they 're trying to you go and buy your ticket and you 're setting off on that day Mm. erm you 're warr cessarily meet all the needs as you 're setting out to meet. They always do in major you need to improve on just as you 're setting the clock just say something like by run on, okay? Small business you 're setting up as self employed Yeah. could we t u can be well annoyed, erm, if you 're setting up a, a joint life policy with waive . So it helps first of all when you 're setting up the spreadsheet I suppose you can f So you 've only got one bit. You 're setting up and prin the setting that 's the e and page sixty eight. They 're really setting out in in great detail what is conta of they may even kill you. And he 's setting the scene for the Resurrection he 's 's gon na be that busy today. if he 's setting it up if they 're not bringing while ng the scene for the Resurrection he 's setting the setting the scene for his triumph status. Oh yeah, yes. If, if you 're s setting up in business as a childminder, not then she ca n't do ours. But she 's setting it up and she will find us somebody t for the Resurrection he 's setting the setting the scene for his triumph out of desp s a question. could do, I mean if I was setting the exam I probably would, sort of, stions about the contingency fund, was setting up the contingency fund, had it got he basis that er we were, when we were setting the, the targets, the time was set or a repair because if you do, are you setting yourself up for a rent increase? Alr

Figure 72. Extract from a BNC_spoken concordance of *setting*, illustrating the typical progressive-preposition collocation “setting up”

: to the north of York, and here we 're dealing with here ma'am er dealing sir w one would be the opposite now if we 're dealing with with this angle? Er for dea : ganging up games are we we know we 're dealing with peoples ' lives and families their friends, but more and more we 're dealing with people Now you 're a caterer! 're dealing with voltage activated channels I 're dealing with are functions. So before eve 'ly dealing with exams. Right, we 're dealing with three questions in an exam ir nt in in our field, because er, we 're dealing with chemicals which we have some ost unanimously Alright, n what we 're dealing with is a small group of M P s re efficiency that we 've, which we 're dealing with the correspondence, I, I ar 've got to be always Right. Do you 're dealing with everything. All kinds. Yes. to show have n't you. But er you 're dealing with a person with with pot er pro of the piece I think, I mean if you 're dealing with fudd fuddy duddy old people y nical conversations. Erm but if you 're dealing with issues about human relations! cation, maybe less important if you 're dealing with er purely technical interacti should go on a front page if you 're dealing I du n no, but the sentences have 'er before the court today. So you 're dealing with a fifty one year old lady, sh given wrong advice. So make sure you 're dealing with a company that 's, that 's c who who those accounts are that you 're dealing with where that 's the situation v ol the internal medium, unless you 're dealing with er some very special er large ree please. He 's alright. You 're dealing. I want two. Ben 's in South Afr ou? Can I have four please. You 're dealing. Oh. I want three please. One, I ou get a weighting because everybody 's dealing with their own, but you 're actua ly has a special need, whereas Gail 's dealing with the peripheral ones, Mm have . I accept that erm and of course he 's dealing with it, approaching it from the e three members of management that he 's dealing with, are called Barbara, oh God at every level instead of local TEC 's dealing with local training needs. Labour in to see the planning officer that 's dealing with it, erm I 've written to ex d their files show they was, was still dealing with it. Mm. What was the outcor London University 's King 's College was dealing with the case. Cleared of rape, A tional Westminster Bank with whom he was dealing, Mr that that property was not a s that your Lordship came to and it was dealing with care to date. Yes. Your Lord eah ? er, no some other bod who was dealing with the accounts at in Enfield. I on the other Yeah, and the lad who was dealing with it all, we had some hassle l re right. Right. That was who was dealing with when I okay, went in there. together. What are the issues are we dealing with? And we 've had one meeting ly, had a number of people police were dealing with murder, then they knew stra were the policemen being told they were dealing with a hostage situation, or a ha is photographs of criminals that we were dealing with, we used to have to take th

Figure 73. Extract from a BNC_spoken concordance of *dealing*, illustrating the typical progressive-preposition collocation “dealing with”

Table 25. Shares of negated and non-negated progressive forms for each analysed verb (verbs in alphabetical order; average share of negation: 8.39%)








































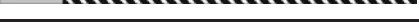


verb form	negated	non-negated	distribution graphically illustrated
			■ negated ■ non-negated
accepting	15.15%	84.85%	
adding	1.64%	98.36%	
agreeing	13.46%	86.54%	
asking	3.97%	96.03%	
being	37.50%	62.50%	
becoming	0.00%	100.00%	
believing	16.67%	83.33%	
betting	28.57%	71.43%	
bothering	51.79%	48.21%	
bringing	3.33%	96.67%	
buying	11.21%	88.79%	
calling	5.65%	94.35%	
carrying	5.61%	94.39%	
changing	10.11%	89.89%	
checking	2.35%	97.65%	
coming	10.00%	90.00%	
costing	6.30%	93.70%	
cutting	6.15%	93.85%	
dealing	9.38%	90.63%	
doing	7.27%	92.73%	
drawing	3.17%	96.83%	
eating	15.45%	84.55%	
expecting	20.61%	79.39%	
explaining	8.33%	91.67%	
feeling	14.67%	85.33%	
finding	2.20%	97.80%	
finishing	2.41%	97.59%	
following	7.32%	92.68%	
forgetting	0.00%	100.00%	
getting	9.72%	90.28%	
giving	15.08%	84.92%	
going	13.94%	86.06%	
happening	2.99%	97.01%	
having	12.86%	87.14%	
hearing	7.46%	92.54%	
helping	5.88%	94.12%	
holding	5.15%	94.85%	
hoping	0.87%	99.13%	
imagining	0.00%	100.00%	
keeping	4.94%	95.06%	
knowing	100.00%	0.00%	
learning	15.15%	84.85%	

Table 25. (continued)

leaving	7.27%	92.73%	
letting	18.75%	81.25%	
liking	25.00%	75.00%	
listening	16.03%	83.97%	
living	2.60%	97.40%	
looking	4.11%	95.89%	
making	3.85%	96.15%	
meaning	16.22%	83.78%	
meeting	0.00%	100.00%	
moving	8.79%	91.21%	
needing	9.52%	90.48%	
paying	16.67%	83.33%	
picking	10.81%	89.19%	
playing	10.00%	90.00%	
providing	10.17%	89.83%	
pulling	7.07%	92.93%	
putting	12.87%	87.13%	
reading	4.17%	95.83%	
remembering	6.67%	93.33%	
ringing	0.79%	99.21%	
running	3.13%	96.88%	
saving	3.51%	96.49%	
saying	14.38%	85.62%	
seeing	8.00%	92.00%	
seeming	0.00%	100.00%	
selling	9.62%	90.38%	
sending	3.92%	96.08%	
setting	5.56%	94.44%	
showing	8.08%	91.92%	
sitting	5.60%	94.40%	
sorting	0.00%	100.00%	
speaking	7.14%	92.86%	
spending	6.49%	93.51%	
standing	3.23%	96.77%	
starting	7.27%	92.73%	
staying	6.67%	93.33%	
stopping	23.15%	76.85%	
suggesting	20.63%	79.37%	
supporting	14.49%	85.51%	
supposing	0.00%	100.00%	
taking	5.00%	95.00%	
talking	6.91%	93.09%	
telling	6.25%	93.75%	
thinking	3.55%	96.45%	
trying	3.80%	96.20%	
turning	5.00%	95.00%	

Table 25. (continued)

understanding	11.11%	88.89%	
using	5.49%	94.51%	
walking	0.00%	100.00%	
wanting	8.11%	91.89%	
watching	8.77%	91.23%	
wearing	11.41%	88.59%	
winning	0.00%	100.00%	
wondering	0.42%	99.58%	
working	10.34%	89.66%	
worrying	13.33%	86.67%	
writing	6.45%	93.55%	

Being, bothering, expecting, letting, liking, meaning, paying, stopping, suggesting, and worrying count among the forms with comparatively high negative portions. On the other hand, only very few negated instances or none at all were found, for instance, in the concordances of *adding, becoming, bringing, checking, finding, finishing, hoping, imagining, living, meeting, ringing, sorting, supposing, walking, winning, and wondering*. Parts from BNC_spoken concordances of one frequently negated and one rarely negated verb form are displayed in Figures 74 and 75.

These two figures (Figure 74 and Figure 75) and the mini-diagrams in Table 25 clearly demonstrate that it would mean giving an oversimplified picture to simply state that 8.39% of the progressives in spoken English are negated. In fact, the picture is somewhat more complex. If we just talk about the average share of progressive negation, we do not do justice to the contextual preferences and affinities of the individual verbs.

busy looking at it they 're not even bothering to try and get on the roundabout.
I do n't know yet. I said things is bothering me a little bit. So she said well
a word? Just say, you know it 's just bothering me a bit about these long hours, d
y, so I do n't know why the heck I 'm bothering. Always try to be behind the othe
. Seven. No. Do n't know why I 'm bothering. Oh Jesus! Flaming bloody hell.
can if she wants. Yeah but I 'm not bothering though. Have it off her. Too muc
now mm I 'm just not, I 'm not bothering with coursework Mm. cos I mean th
rst, then you, then him then I 'm not bothering today cos I think mine is nearly
e you rang the bank? no no I 'm not bothering well if you can just put me mind
sleeping? Ah! Yes. Is it? I 'm not bothering to read all this, this is rubbish
ng though, look at it. Oh I 'm not bothering. No. I 'll do that on Tuesday. T
the times that I 've said oh I 'm not bothering this year. I actually got them ba
getting a carpet done, I 'm not, not bothering, we 'll just have to rough it for
are you doing the one o'clock or not bothering? No ca n't be bothered to do owt
ing of the tape? Because we 're not bothering watching any more now. That 's it
everything are n't ya? You 're not bothering with buses you need n't go any
you come out. Aye, aye. It 's not bothering me now. No. Just when when you y
he same as you, that er If it 's not bothering you, leave it alone. Aye. Said t
No I 'm not, it 's not, it 's not bothering me No it does n't bother me. Actua
it 's so new to them that they 're bothering to cost it, but how did we go on
I said, I do n't know what you 're bothering for. I said by heck you 've some
ed down with something that 's really bothering you. Because you do n't think th
p and had to stop and that 's what 's bothering me. Wan na be a man, man cub. Go
goodness! Oh! Did you get it It was bothering me, that 's why I wanted her to g
id she were n't wearing why are you bothering. Size ten. Mm? Size ten Karen. M

Figure 74. Extract from a BNC_spoken concordance of *bothering*, illustrating a high share of negation

that 's what I think it, it 's, I 'm wondering now if the nurses today get the sa
 Oh! Richard. What you got? And I 'm wondering if I should ring them actually, f
 ld do, but it never mentions that. I 'm wondering whether to go the A one way, or Lc
 rely thing. You 've got two Yeah. I 'm wondering whether the things from that will
 men. Can you hear her saying it? I 'm wondering if she 's saying down you know, li
 person would she have been, I 'm I 'm wondering whether she would be doing this f
 h with what you get Yes I mean I 'm wondering if perhaps Mm. I mean I may be at
 pact on our own service here. Now I 'm wondering how easy or whether there 's any n
 and things very well. Yes that 's I 'm wondering if, if it 's too acid for a good
 training session if necessary. So I 'm wondering if we need to have that in the tra
 : just. Yes. Erm. The other thing I 'm wondering whether I should do the Petroline
 e, I do n't I do n't know, what I 'm wondering is what, you know, what, you kno
 ell I can hear it now. Oh. Yeah I 'm wondering whether it 's playing to itself, y
 only wall Well, well I was just maybe wondering whether you might use a part of th
 me quick off them though if you 're not wondering if you 're not watching what 's b
 t be training railway safety we 're wondering You actually ? how this clashes
 xt to nothing! So er In fact, we 're wondering whether it 's going to be possible
 iths before they 'll reply. And you 're wondering, I mean I do n't think we can exp
 is you need go to the shop and you 're wondering if you 've got the right change. E
 ntlemen. Good evening in case you 're wondering why the conversation 's being recc
 imme. Right. Okay just in case you 're wondering who I 'm talking t Danny anything
 appy with this. You know I mean you 're wondering whether it 's going to blow the c
 irs you can stay in bed cos then you 're wondering and wondering and won so you 're n
 distributed because you, you, you 're wondering what to put in, what not to put i
 old brother who is on the job and he 's wondering why, are they gon na come and fet

Figure 75. Extract from a BNC_spoken concordance of *wondering*, illustrating a low share of negation

4.5.7 Verbs and other lexical-grammatical phenomena [BNC/BoE]

In Section 4.3.7 we looked at repeated occurrences of progressives in three types of lexical-grammatical constructions: questions, if-clauses, and relative clauses. We will now see whether the determined average shares of these three phenomena in the BNC/BoE dataset are representative of the majority of verb forms covered in this study or whether it is possible to detect typical co-selection patterns.

Table 26 displays the percentages of questions, if-clauses, and relative clauses based on the progressive datasets of 99 individual verbs (BNC_spoken and BoE_brspok tokens taken together). Figures which are clearly below or above average are highlighted with a grey shading.

One first thing worth noting with reference to questions is that 90 out of 99 verb form types show occurrences in interrogative contexts. However, the individual shares of verb-question co-selection differ considerably. A number of verbs apparently favour declarative contexts (e.g. *becoming*, *imagining*, *making*, *meeting*, *setting*, *showing*, *sorting*, *understanding*, *wondering*), while others frequently occur in questions (e.g. *bringing*, *eating*, *feeling*, *finishing*, *listening*, *living*, *playing*, *remembering*, *wanting*). Figure 76 shows concordance samples of two question-prone progressive forms, *feeling* and *listening*, and illustrates the repeated occurrence of the patterns “are you listening”, “What are you listening to?”, and “(how) are you feeling” in spoken English.

If we look at the shares of progressive-if-clause co-occurrences displayed in Table 26, we also find a lot of variation among different verb form types. The percentages for some verbs deviate considerably from the average share of 4.47%. *Buying*, *feeling*, *holding*, *ringing*, *selling*, *wanting*, and *wearing*, for instance, tend to occur often in if-constructions. The extract from a concordance of *wearing* presented in Figure 77 provides a couple of examples of the common cluster “if you’re wearing” in context.

oor Richard he just come back are you feeling niggly? , mm he just being naughty, t to enjoy the, the view. Er are you feeling in the mood to take it easy? Er some ke the car? Was he alright? How are you feeling now? What did she bring you for you 's well over the yard arm. How are you feeling? Is n't it time you stretched your want the marmalade on? That on? Are you feeling alright before you sort of? Tha lright. She said, oh! I said how you feeling? She said, I 'd like to see a. He en then, yeah. Yeah. are n't you feeling very well Jonathan? Got all my cake the experience of the people who were listening, so when a went down to, they al it in this tin in this draw. Are you listening? Yeah. In case I forget where I A la carte? That 's right! Are you listening? A la carte then? Just give me ews of G M B members. Well, are you listening Gordon, now you know. So collea another one at eight Yeah, are you listening to him I had another at ten to e o sit on your potty? Charlotte are you listening to me, ooh, ooh, wakey, wakey stay there Alan. All right? Are you listening to me? Yes I am I 'm all ears. S h. Hello little doggy! What are you listening to? Me? What was you listening come to think of it. Sam what are you listening to? Oh bloody fucking shite hard u got your headphones in? What are you listening to? I 'm listening to like a vir e you listening to? Me? What was you listening to? Nothing. Hello little dog it down look up and shut up. What you listening to? This. Come here a minute. C . I know. Now oh shit which what you listening to Cass? I 'm not listening to a ave some more teeth through. You listening? Yeah. Got to read all this? We

Figure 76. Extracts from BNC_spoken concordances of *feeling* and *listening*, illustrating above-average shares of questions

the token dishing out tour and you 're wearing them, you get your token, okay? ssibly twice, or three times if you 're wearing make up. Okay? We normally say c you have a skirt on , or if you 're wearing a dressing gown When sh, Sarah cor cca 's swimsuit, he said cor if you 're wearing that I think I 'll go swimming as you 've got that one, even if you 're wearing make up, you can use this cleanse something like that coat. He wo, and we put towels over you. If you 're wearing a bra we just undo your bra, okay? tions, whatever it is, yes? If you 're wearing the ear defenders when the supervi feel better with people knowing you 're wearing a wig They do n't know I 'm weari lightly different to Howard 's? You 're wearing the summer shorty for greater free rue coloured. You can tell what you 're wearing Yes. it 's the right colour. You why did n't use a coat, what you 're wearing put it on why do n't you get dress to have short length coats when you 're wearing longer skirts, you know. Yeah, v 's right. Erm but the other chap 's wearing a striped one. He 's he 's got a bother about a suit will he? If he 's wearing his uniform No, that 's what he s

Figure 77. Extract from a BNC_spoken concordance of *wearing*, illustrating a high share of if-clauses

. This by the way is er Donald who is making sound recordings of in the school tod er that time is the orange line. Just making the point that in the early years, yo . Yeah. And that 's the point I 'm making. Mm. These procedures should be man ors and crocodiles so the point I 'm making is it is just wrong to say that, that of phasing. Erm but the the point I 'm making is that existing figures erm for the t. The the That was the point I was making. Yeah but the favour we 're only tou ow they do But anyway the point I was making was that they call it distant intimac twenty three million the point I was making in the speech was that twenty three m of making all these errors that I was making, oh no what 's this? It 's seventeen ot of money Well this is that Ian was making in the first place is n't it, product I think, I think the point Ronnie was making before was that er there is an agreem

Figure 78. Extracts from a BNC_spoken concordance of *making*, illustrating a high share of relative clauses

Finally, the distribution of relative clauses across individual progressive verb types is rather heterogeneous too, ranging from zero to 25.42 per cent. A number of verbs show significantly above-average percentages (e.g. *buying*, *dealing*, *helping*, *providing*, and *using*), whereas others never occur in relative clauses in the examined datasets (e.g. *cutting*, *explaining*, *feeling*, *following*, *forgetting*, *remembering*, and *sorting*). Figure 78 shows two samples taken from a BNC_spoken concordance of *making*. Not only do the samples illustrate the repeated occurrence of relative clauses, they also highlight

Table 26. Shares of questions, if-clauses, and relative clauses for each analysed verb (verbs in alphabetical order)

verb form	ques- tions ø10.94%	if- clauses ø4.47%	relative clauses ø6.17%	verb form	ques- tions ø10.94%	if- clauses ø4.47%	relative clauses ø6.17%
accepting	6.06%	4.55%	4.55%	meeting	0.00%	0.00%	0.00%
adding	13.11%	8.20%	3.28%	moving	5.49%	6.59%	5.49%
agreeing	9.62%	3.85%	1.92%	needing	11.90%	4.76%	11.9%
asking	12.58%	1.99%	6.62%	paying	5.33%	5.33%	5.33%
being	0.00%	0.00%	12.50%	picking	15.32%	3.60%	5.41%
becoming	4.88%	1.22%	6.10%	playing	19.09%	3.64%	5.45%
believing	0.00%	0.00%	0.00%	providing	6.78%	3.39%	25.42%
betting	14.29%	28.57%	0.00%	pulling	12.12%	3.03%	3.03%
bothering	14.29%	7.14%	1.79%	putting	9.90%	5.94%	6.93%
bringing	21.11%	2.22%	12.22%	reading	6.94%	2.78%	4.17%
buying	9.48%	12.93%	12.93%	remembering	23.33%	3.33%	0.00%
calling	16.94%	1.61%	12.10%	ringing	7.94%	8.73%	1.59%
carrying	5.61%	0.93%	7.48%	running	8.33%	4.17%	8.33%
changing	5.62%	3.37%	1.12%	saving	7.02%	3.51%	7.02%
checking	9.41%	0.00%	1.18%	saying	13.73%	3.92%	3.92%
coming	9.00%	3.00%	7.00%	seeing	10.67%	1.33%	6.67%
costing	8.66%	1.57%	3.15%	seeming	0.00%	0.00%	0.00%
cutting	6.15%	3.08%	0.00%	selling	9.62%	11.54%	7.69%
dealing	4.69%	6.25%	21.09%	sending	8.82%	3.92%	5.88%
doing	16.36%	6.06%	10.30%	setting	3.70%	7.41%	7.41%
drawing	9.52%	0.00%	9.52%	showing	3.03%	4.04%	8.08%
eating	20.91%	5.45%	3.64%	sitting	8.00%	5.60%	4.80%
expecting	12.72%	3.07%	3.95%	sorting	0.00%	0.00%	0.00%
explaining	8.33%	4.17%	0.00%	speaking	15.71%	5.71%	8.57%
feeling	22.67%	12.0%	0.00%	spending	9.09%	6.49%	9.09%
finding	17.58%	4.40%	8.79%	standing	7.53%	6.45%	6.45%
finishing	18.07%	0.00%	6.02%	starting	9.09%	2.73%	2.73%
following	17.07%	4.88%	0.00%	staying	14.81%	5.19%	4.44%
forgetting	8.11%	0.00%	0.00%	stopping	15.74%	4.63%	0.93%
getting	8.33%	1.39%	4.86%	suggesting	8.52%	0.90%	4.93%
giving	11.11%	2.38%	7.14%	supporting	13.04%	4.35%	4.35%
going	14.90%	5.29%	5.77%	supposing	0.00%	0.00%	0.00%
happening	17.52%	0.85%	7.69%	taking	11.67%	6.67%	6.67%
having	7.14%	5.71%	7.14%	talking	10.11%	4.79%	9.57%
hearing	10.45%	0.00%	11.94%	telling	4.38%	1.88%	0.63%
helping	11.76%	1.18%	12.94%	thinking	4.96%	4.96%	2.84%
holding	5.15%	9.28%	9.28%	trying	7.07%	3.26%	5.43%
hoping	5.24%	0.00%	2.18%	turning	5.00%	3.75%	6.25%
imagining	0.00%	0.00%	9.09%	understanding	0.00%	0.00%	0.00%
keeping	17.28%	8.64%	6.17%	using	16.48%	6.59%	16.48%
knowing	0.00%	0.00%	0.00%	walking	10.31%	9.28%	5.15%

Table 26. (continued)

learning	13.64%	3.03%	3.03%	wanting	18.02%	9.91%	5.41%
leaving	16.36%	6.36%	5.45%	watching	11.40%	7.02%	3.51%
letting	13.54%	8.33%	5.21%	wearing	14.77%	10.74%	5.37%
liking	25.00%	0.00%	0.00%	winning	20.00%	4.00%	4.00%
listening	15.38%	6.41%	7.05%	wondering	4.18%	1.67%	0.84%
living	19.48%	3.90%	9.09%	working	12.93%	7.76%	5.17%
looking	13.70%	6.85%	8.90%	worrying	16.67%	6.67%	6.67%
making	1.92%	4.81%	23.08%	writing	16.13%	9.68%	1.61%
meaning	10.81%	2.70%	5.41%				

a typical collocation between “to BE making” and “the point”. It is apparently more common to say “... the point I’m/was making” than “I’m/was making the point...”.

Again, with respect to the three discussed lexical-grammatical phenomena (questions, if-clauses, and relative clauses), we observed clear tendencies of inter-verbform variation and some significant deviations from the “normal” distributions determined in Section 4.3. We shall see later whether EFL textbooks pay attention to at least some of the contextual patterns identified here.

4.5.8 Verbs and adverbial specification [BNC/BoE]

Before moving on to looking at some relations between verbs and progressive time references and functions, we will now turn to one final phenomenon in the lexical context of progressives: adverbial specification. It will be analysed how often each individual verb form in the BNC/BoE dataset co-occurs with an adverbial of time (e.g. *now, at the moment, still*), an adverbial of place (e.g. *here, there, in America*), or with an item from the “other adverbials” group (e.g. *actually, whether, really*; cf. Section 4.3.8). Another objective is to find out whether there are any particular verb-form-specific collocations, or whether the average co-occurrence values determined above (cf. 4.3.8) are roughly representative of progressives in general.

Table 27 lists the percentages of three different types of adverbial specification and graphically illustrates the adverbial-progressive co-selection profile for each of the 99 analysed verbs. As the mini-diagrams included in Table 27 show, the overall shares of adverbial specification differ considerably from verb to verb (cf. the different lengths of the bar charts). While some verbs show particularly high percentages (e.g. *adding, checking, finishing, sitting, standing, wondering*), others rarely co-occur with time, place, or other adverbials (e.g. *hoping, meaning, providing, suggesting*).

Not taking into account the very infrequent forms *liking, seeming, and supposing*, the shares of time adverbial specification (average: 25.04%) range from 5.83% (*suggesting*) to 60.24% (*finishing*). Time adverbials are particularly common in progressive constructions with the following verb forms: *adding, checking, finishing, learning, living, ringing, seeing, sorting, starting, wondering, and working*. Typical verb-adverbial collocations include “just checking”, “finishing” + “now”, “learning” + “now”, “start-

Table 27. Shares of different types of adverbial specification of each analysed verb (verbs in alphabetical order)





































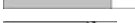



verb form	time adverbial ø25.04%	place adverbial ø3.04%	other adverbial ø7.68%	distribution graphically illustrated ■ time adv. ▨ place adv. □ other adv.
accepting	16.67%	0.00%	12.12%	
adding	37.70%	0.00%	14.75%	
agreeing	17.31%	3.85%	9.62%	
asking	23.18%	0.00%	7.28%	
being	37.50%	12.50%	0.00%	
becoming	31.71%	1.83%	15.85%	
believing	16.67%	0.00%	33.33%	
betting	14.29%	0.00%	0.00%	
bothering	28.57%	0.00%	23.21%	
bringing	22.22%	0.00%	5.56%	
buying	24.14%	0.86%	7.76%	
calling	15.32%	3.23%	7.26%	
carrying	31.78%	0.00%	7.48%	
changing	32.58%	0.00%	12.36%	
checking	55.29%	2.35%	10.59%	
coming	23.00%	5.00%	5.00%	
costing	18.90%	0.79%	18.11%	
cutting	30.77%	0.00%	9.23%	
dealing	20.31%	1.56%	10.94%	
doing	24.24%	4.24%	4.85%	
drawing	23.81%	0.00%	7.94%	
eating	30.91%	0.00%	7.27%	
expecting	19.30%	0.88%	3.07%	
explaining	25.00%	4.17%	18.75%	
feeling	28.00%	1.33%	5.33%	
finding	25.27%	6.59%	6.59%	
finishing	60.24%	0.00%	6.02%	
following	24.39%	2.44%	4.88%	
forgetting	10.81%	0.00%	10.81%	
getting	21.53%	0.00%	9.72%	
giving	15.08%	2.38%	11.11%	
going	19.23%	0.00%	5.77%	
happening	21.79%	11.97%	4.27%	
having	27.14%	2.86%	4.29%	
hearing	31.34%	1.49%	2.99%	
helping	9.41%	0.00%	8.24%	
holding	27.84%	7.22%	4.12%	
hoping	9.17%	0.44%	4.37%	
imagining	31.82%	0.00%	13.64%	
keeping	24.69%	2.47%	6.17%	
knowing	0.00%	0.00%	0.00%	—

Table 27. (continued)

learning	37.88%	0.00%	3.03%	
leaving	34.55%	3.64%	6.36%	
letting	17.71%	2.08%	5.21%	
liking	0.00%	0.00%	12.50%	
listening	21.15%	0.64%	5.77%	
living	37.66%	10.39%	3.90%	
looking	23.97%	3.42%	8.90%	
making	21.15%	3.85%	6.73%	
meaning	8.11%	0.00%	2.70%	
meeting	38.46%	7.69%	0.00%	
moving	18.68%	3.30%	8.79%	
needing	16.67%	2.38%	19.05%	
paying	18.00%	0.67%	8.00%	
picking	23.42%	1.80%	11.71%	
playing	31.82%	6.36%	1.82%	
providing	6.78%	1.69%	6.78%	
pulling	22.22%	0.00%	7.07%	
putting	23.76%	2.97%	5.94%	
reading	31.94%	1.39%	2.78%	
remembering	26.67%	0.00%	10.00%	
ringing	39.68%	0.00%	5.56%	
running	36.46%	1.04%	7.29%	
saving	19.30%	0.00%	3.51%	
saying	11.11%	0.00%	5.23%	
seeing	38.67%	2.67%	8.00%	
seeming	100.00%	0.00%	0.00%	
selling	22.12%	0.96%	4.81%	
sending	20.59%	0.00%	1.96%	
setting	14.81%	0.00%	0.00%	
showing	29.29%	2.02%	6.06%	
sitting	24.00%	36.80%	3.20%	
sorting	39.39%	0.00%	3.03%	
speaking	28.57%	4.29%	7.14%	
spending	23.38%	1.30%	7.79%	
standing	26.88%	32.26%	4.30%	
starting	40.91%	2.73%	4.55%	
staying	34.07%	22.96%	4.44%	
stopping	15.74%	10.19%	8.33%	
suggesting	5.83%	0.90%	4.04%	
supporting	11.59%	0.00%	10.14%	
supposing	0.00%	0.00%	0.00%	
taking	26.67%	1.67%	5.00%	
talking	27.66%	2.66%	4.26%	
telling	26.88%	0.00%	3.13%	
thinking	36.17%	0.00%	4.26%	
trying	22.83%	0.54%	7.61%	

Table 27. (continued)

turning	20.00%	2.50%	5.00%	
understanding	22.22%	0.00%	0.00%	
using	26.37%	1.10%	1.10%	
walking	36.08%	2.06%	5.15%	
wanting	14.41%	0.00%	5.41%	
watching	35.96%	0.88%	2.63%	
wearing	22.82%	0.00%	6.71%	
winning	20.00%	4.00%	4.00%	
wondering	41.00%	0.00%	46.86%	
working	37.93%	6.03%	8.62%	
worrying	21.67%	0.00%	6.67%	
writing	32.26%	0.00%	6.45%	

ing” + “now”, “working” + “when”, and “just wondering”. Extracts from BNC_spoken concordances which illustrate two of these patterns are displayed in Figure 79. Figure 80 highlights the frequent collocation of the adverbial “just” with *wondering*-progressives. Interesting to observe in this figure is a more complex phraseological pattern which can be described with the formula “I + ’m/was + just + wondering + interrogative particle” (*whether/if, when, what, why, how*).

Adverbials of place are on the whole rather infrequent in the context of progressives and the majority of verbs are very rarely (or never) specified by any of these adverbials. Some verb forms, however, show considerable shares of place adverbial specification (especially *happening, living, sitting, standing, and staying*), shares which lie significantly above the average value of 3.04 per cent. Worth highlighting in this context is the repeated co-selection of the items “sitting” + “here”, “standing” + “there”, and “staying” + “in PP”. A couple of concordance lines which exemplify these co-selection patterns are given in Figure 81.

Among the verb form types that commonly co-occur with items from the “other adverbials” group are *bothering, costing, explaining, imagining, needing, and wondering*. The only significant collocation that was found, however, is that between “wondering” and non-conditional “if” and its synonym “whether”, as shown in many of the

own. But he ta all the dirty tricks are starting now. Kinnock 's a funny he real e reactions, and to get over that we are starting now to use computers. We 're gi ion twenty three, they 're actually now starting to look at membership of a trade bigger cities that are criminals are now starting to discover that it 's easier to it 's not a bad hedge and it 's now starting to get you know you see, I 'm a, on because the nuclear dimension was now starting to become very important. The e it that 's happening. obviously they 're starting now to pick the bits off one by e. Mummy , when are we starting? We 're starting now so could you sit down on you ding it and once it was up. When we were working on it it was just a load of excav to your family did it make that you were working, financially? Well. When I went see. Grown up now? Mhm. When you were working here did you always kind of? When ch is what you said earlier when you were working them out you said, Oh so they get ming interested in politics when you were working at, we were your friends surprise for them and you qualified while you were working for them. Yeah. And as far as I

Figure 79. Extracts from BNC_spoken concordances of *starting* and *working*, illustrating typical verb-adverbial collocations.

heard what you said! I 'm just wondering why why Hazel, Hazel thinks y
sed to have . I know, but I 'm just wondering how it 's got on there. Probab
ah I no it do n't hurt. I 'm just wondering what waxing does though. The
; a vole? It could be, or I 'm just wondering if it 's actually squirrels,
women to have a child and, I 'm just wondering how many of us are thinking a
us going a full circle and I 'm just wondering when he 's gon if I 'll be al
'm of of this erm session, I 'm just wondering if we could have ten minutes
nineteen ninety one. And I 'm just wondering whether er the er point at ni
'm Madam Deputy Speaker. I 'm just wondering er the statutory instruments
rases, erm, not an asset, I 'm just wondering what the title below that, t
Divisional Manager West. I 'm just wondering whether one or two things tha
again does n't it ? Yeah I 'm just wondering if there that 's related to t
roduce as many conkers. So I 'm just wondering whether this is a different v
all night Doctor, Aye. I 'm just wondering if it 's my teeth or that it
all tend to go now. Erm. I 'm just wondering th I mean how how how do you
my own particular guild, I 'm just wondering if it 's the secretaries that
's got no small talk. And I 'm just wondering what the the knock on effect
that was a everyone but I 'm just wondering why it is that now what of Wa
se it must be must n't it? I 'm just wondering Well I go through one of the
like that and trap them but I 'm just wondering Stefan could it also be a fox
ital recently. Yeah. Erm I 'm just wondering if he ever gets the messages
/. Twelve thou, you see, I 'm just wondering whether I should quote on a C
we 've got to go on to. I 'm just wondering how much you know already. K
ies well you know Mhm and I 'm just wondering if maybe, maybe we will just
a research director or A.D. I 'm just wondering whether we can comply with th
be alright. What were we on er I 'm just wondering how yo do the what 's her
o it just has his name. I was just wondering should it not have a photogra
ree, nine , six. Yes, no I was just wondering, that 's all. Yep, I 'm gon
that you put to us. Well I was just wondering whether I could reserve that
l. Might the Mr Chairman I was just wondering if it 's worth just mentionin
another question? Yes, I was just wondering, er raise one point. The C
them midweek delivery. I was just wondering put that back to the same del
to contact, to find out I was just wondering whether it 's something, you
leave that alone. Right I was just wondering if they would i if they would
op entrance. Well and er I was just wondering we we ought to do something,
has gone up. Oh Sorry, I was just wondering if you actually asleep or not
entertainment you know. I was just wondering about well about that. I was
did you try? Foxes. Mm I was just wondering erm Is there anywhere else? '
It 's just erm I 'm I w I was just wondering if you wanted to write them a

Figure 80. Extract from a BNC_spoken concordance of *wondering*, illustrating typical contextual patterns

y till half past five at night I was sitting here yesterday mor day before, and
did n't he? No. It 's just I, I was sitting here, have n't phoned my parents
I had to go from here. My mother was sitting here on a chair yeah, when I come
he Fodors travel guides, and Stan was sitting here nodding his head in agreement
leave it on the hob, one day we were sitting here and I do n't normally, if I pu
football belongs to men. If we were sitting here just now debating tennis, or
', Jane I mean. No. She 's just standing there Really quite strong. Yes. takin
all its friends and then it was just standing there on the ground. But really it
Are yes, but you see, we were just standing there, and I was just showing off.
till one thousand five hundred people standing there, rain, hail, sleet or blow.
nce actually to yourself when you 're standing there and the audience participation
irty pupils so And how many are there standing there in the playground? There 's
pe and date rape. Mm. And er I was standing there I was n't really sure whether
le staying in the hotel and we're all staying in the hotel and we're a # family. We
<M09> massive suites people have been staying in # <M01> Huh. <M09> they've been p
gon. <F01> Mm. <F02> They had been staying in Vancouver. But of course Vancouver
side # around them or if they're just staying at home and watching the # telly all
2> At least I'm not in Birmingham I'm staying in Lichfield but my pupils # are. <M
moment so # <F02> Oh dear. <F01> I'm staying in the front room until they've finis
at # <F0X> Well on Saturday night I'm staying at FX's and on Sunday night I'm # sta
over in erm # St Peter's Grove # <F04
slade # <M07> I'm not in a hotel. I'm staying in an apartment in a place called # S
ink Well I've got job security. I'm # staying. <M01> Mm. <M02> That <ZF1> may <ZF
u phone up and say Oh # you <ZGY> I'm staying in <ZZ1> <!--place--> <!--name--> <ZZ

Figure 81. Extracts from BNC_spoken and BoE_brspek concordances of *sitting*, *standing*, and *staying*, illustrating typical verb-place adverbial collocations

corpus examples in Figure 80 above. There are a couple of occurrences of “actually costing”, and “actually needing” (cf. Figure 82), though not enough to justify speaking of “typical clusters”.

er as a people and which is # actually costing our industry millions upon millions
 approximate idea of # what it's actually costing. <N01> Yeah. Mm the only trouble is
 on a free number I E it wasn't actually costing you that meant that # it was costing
 FOX> Mm. <FOX> when I wasn't actually needing the stuff # <FOX> Mm. <FOX> for all
 t # <F02> No because I wasn't actually needing things I just made up what I # needed
 w <ZF1> you're <ZF0> you're # actually needing to kind of reassert that as the I mea
 for your prices that you're actually # needing # <F01> Yes. <F02> All right that's

Figure 82. Extracts from BoE_brspok concordances of *costing*, and *needing*, showing the repeated occurrence of “actually costing” and “actually needing”

As our observations on verb-adverbial co-selection in progressive constructions show, there is no such thing as a common pattern which could be applied to the majority of progressive verb forms in BNC/BoE. Instead, shares of adverbial specification differ a lot from verb to verb and a number of typical patterns can be determined for individual verbs and frequent adverbials. Again, our findings indicate that average percentages may be misleading as they do not necessarily tell us a lot about the actual use of all, or most progressives but may obscure important collocational and colligational tendencies of lexical items.

4.5.9 Verbs and time reference [BNC/BoE]

Let us now turn from phenomena in the context of progressives to the function side. In analogy to Section 4.4 where we looked at progressive functions of the group of 99 types, we will start with the analysis of time reference, this time, however, with a shift in perspective from the verb group to individual verbs. I will examine the possibility of significant relations between different time reference types (past, present, future, present/future “indeterminate”; cf. 4.4.1) and the progressive forms of particular verbs.

In order to identify such relations, the progressive form tokens of each verb type were filtered according to time reference and subdivided into the four categories “past time reference”, “present time reference”, “future time reference”, and “present/future time reference”. The resulting percentages for each verb are given in Table 28, together with miniature diagrams which serve to visualise the time reference distributions.

If we first concentrate on the first column of percentages and the left-hand parts of the diagrams in Table 28, we notice a great amount of inter-verb variation concerning shares of past time reference progressives. As was to be expected, most of the verbs which often occur in the PastProg (cf. Table 21 in Section 4.5.2) are also significantly common in past time contexts. Particularly high percentages were found for *carrying*, *expecting*, *explaining*, *hearing*, *imagining*, *living*, *meaning*, *reading*, *standing*, *telling*, *thinking*, *walking*, *watching*, and *working*. On the other hand, a number of verbs, such as *changing*, *going*, *moving*, *remembering*, and *wondering*, show shares of past time reference that lie far below average.

Quite some variation can also be observed with respect to individual verbs and present time reference. While verb forms such as *coming*, *going*, *leaving*, *meeting*, *providing*, *sending*, *staying*, and *winning* are rarely used to refer to present actions or events, others predominantly relate to the moment of speaking. Clearly above-average

Table 28. Time reference distribution for each analysed verb (verbs in alphabetical order)

verb form	past time ref.	present time ref.	future time ref.	present/ future time ref.	distribution graphically illustrated
	ø28.87%	ø38.09%	ø17.60%	ø15.45%	<div> <div></div> past t. ref. <div></div> present t. ref. <div></div> future t. ref. <div></div> present/future t. ref. </div>
accepting	9.09%	74.24%	13.64%	3.03%	
adding	18.03%	68.85%	11.48%	1.64%	
agreeing	26.92%	59.62%	3.85%	9.62%	
asking	39.74%	53.64%	1.99%	4.64%	
being	25.00%	62.50%	12.50%	0.00%	
becoming	19.51%	73.17%	5.49%	1.83%	
believing	50.00%	50.00%	0.00%	0.00%	
betting	28.57%	28.57%	28.57%	14.29%	
bothering	19.64%	67.86%	7.14%	5.36%	
bringing	20.00%	38.89%	33.33%	7.78%	
buying	25.00%	30.17%	27.59%	17.24%	
calling	25.81%	64.52%	5.65%	4.03%	
carrying	40.19%	47.66%	5.61%	6.54%	
changing	14.61%	46.07%	31.46%	7.87%	
checking	31.76%	42.35%	8.24%	17.65%	
coming	25.00%	13.00%	56.00%	6.00%	
costing	22.05%	28.35%	24.41%	25.20%	
cutting	30.77%	29.23%	20.00%	20.00%	
dealing	21.09%	48.44%	4.69%	25.78%	
doing	22.42%	43.64%	15.15%	18.79%	
drawing	19.05%	53.97%	6.35%	20.63%	
eating	32.73%	32.73%	14.55%	20.00%	
expecting	45.61%	37.72%	11.84%	4.82%	
explaining	58.33%	29.17%	2.08%	10.42%	
feeling	32.00%	41.33%	5.33%	21.33%	
finding	23.08%	48.35%	8.79%	19.78%	
finishing	22.89%	30.12%	39.76%	7.23%	
following	24.39%	31.71%	14.63%	29.27%	
forgetting	18.92%	56.76%	2.70%	21.62%	
getting	26.39%	38.89%	16.67%	18.06%	
giving	19.05%	26.19%	39.68%	15.08%	
going	13.46%	5.29%	77.40%	3.85%	
happening	23.08%	32.05%	26.50%	18.38%	
having	28.57%	21.43%	30.00%	20.00%	
hearing	44.78%	25.37%	7.46%	22.39%	
helping	38.82%	14.12%	11.76%	35.29%	
holding	24.74%	37.11%	14.43%	23.71%	
hoping	26.20%	72.05%	1.75%	0.00%	
imagining	45.45%	50.00%	0.00%	4.55%	
keeping	16.05%	41.98%	16.05%	25.93%	
knowing	100.00%	0.00%	0.00%	0.00%	
learning	24.24%	59.09%	4.55%	12.12%	

Table 28. (continued)

leaving	15.45%	9.09%	61.82%	13.64%	
letting	21.88%	25.00%	38.54%	14.58%	
liking	0.00%	87.50%	0.00%	12.50%	
listening	32.05%	54.49%	1.28%	12.18%	
living	49.35%	35.06%	0.00%	15.58%	
looking	27.40%	50.68%	4.11%	17.81%	
making	26.92%	43.27%	12.50%	17.31%	
meaning	70.27%	21.62%	2.70%	5.41%	
meeting	15.38%	7.69%	61.54%	15.38%	
moving	13.19%	23.08%	41.76%	21.98%	
needing	23.81%	23.81%	33.33%	19.05%	
paying	25.33%	12.67%	18.67%	43.33%	
picking	17.12%	34.23%	23.42%	25.23%	
playing	40.00%	28.18%	17.27%	14.55%	
providing	18.64%	5.08%	23.73%	52.54%	
pulling	36.36%	33.33%	14.14%	16.16%	
putting	19.80%	27.72%	28.71%	23.76%	
reading	48.61%	20.83%	6.94%	23.61%	
remembering	6.67%	63.33%	10.00%	20.00%	
ringing	34.92%	37.30%	11.11%	16.67%	
running	40.63%	34.38%	12.50%	12.50%	
saving	38.60%	15.79%	22.81%	22.81%	
saying	34.64%	50.33%	1.31%	13.73%	
seeing	25.33%	38.67%	14.67%	21.33%	
seeming	100.00%	0.00%	0.00%	0.00%	
selling	29.81%	12.50%	25.96%	31.73%	
sending	17.65%	10.78%	49.02%	22.55%	
setting	22.22%	24.07%	44.44%	9.26%	
showing	27.27%	47.47%	12.12%	13.13%	
sitting	47.20%	25.60%	14.40%	12.80%	
sorting	42.42%	21.21%	24.24%	12.12%	
speaking	38.57%	32.86%	8.57%	20.00%	
spending	23.38%	11.69%	20.78%	44.16%	
standing	53.76%	25.81%	5.38%	15.05%	
starting	22.73%	35.45%	33.64%	8.18%	
staying	31.85%	10.37%	47.41%	10.37%	
stopping	18.52%	12.04%	56.48%	12.96%	
suggesting	17.94%	77.13%	0.00%	4.93%	
supporting	28.99%	43.48%	7.25%	20.29%	
supposing	100.00%	0.00%	0.00%	0.00%	
taking	24.17%	17.50%	37.50%	20.83%	
talking	35.11%	48.94%	4.26%	11.70%	
telling	52.50%	23.13%	11.25%	13.13%	
thinking	41.84%	45.39%	2.13%	10.64%	
trying	21.74%	54.89%	7.07%	16.30%	
turning	22.50%	38.75%	27.50%	11.25%	

Table 28. (continued)

understanding	22.22%	55.56%	22.22%	0.00%	
using	34.07%	30.77%	8.79%	26.37%	
walking	54.64%	15.46%	12.37%	17.53%	
wanting	29.73%	37.84%	7.21%	25.23%	
watching	50.88%	23.68%	9.65%	15.79%	
wearing	34.23%	26.17%	13.42%	26.17%	
winning	28.00%	6.00%	58.00%	8.00%	
wondering	6.28%	89.54%	0.00%	4.18%	
working	47.41%	27.59%	6.90%	18.10%	
worrying	26.67%	58.33%	1.67%	13.33%	
writing	38.71%	33.87%	4.84%	22.58%	

percentages of present time reference are, for instance, found with *adding*, *bothering*, *calling*, *hoping*, *liking*, *remembering*, *suggesting*, and *wondering*. Another group of progressive verb forms typically refers to the future. This group includes the following items, most of which have just been described as non-present time reference verbs: *coming*, *going*, *leaving*, *meeting*, *sending*, *setting*, *staying*, *stopping*, and *winning*. Except for *staying* and *winning*, these forms share the semantic feature of “movement”. Apparently, future reference expressed by progressives is closely connected with dynamic action.

In this context, I should say a few words about progressives with *going*. As the concordance sample in Figure 84 shows, in the large majority of examples, *going* is immediately followed by *to* plus the infinitive of a verb. The construction “TO BE *going to* + infinitive” is usually referred to as “*going to* future” and, as the name indicates, expresses futurity, however of a different kind than progressives that are not followed by a *to*-infinitive. A closer look at the *going to* examples in Figure 84 reveals that *going to* future is typical in the context of speaker intentions (note the large number of first person subjects). The “normal” progressive future, i.e. the future reference expressed by means of progressives, esp. those listed in the preceding paragraph, does not appear to be restricted to that kind of intention/purpose contexts. Instead, what is often referred to, are certain plans that a person or group of people has for the future. These plans are in general much more firm and binding than the intentions expressed by the *going to* future. As the *leaving* examples presented in the concordance extract in Figure 83 show, reference is often made to a fixed time or date, e.g. “eight o’clock”, “ten to eight”, or “Christmas time”, so speakers in fact use progressives to talk about some arranged events that are taking place in the future. These corpus-driven findings are in accordance with the findings of other scholars who state that TO BE *going to* mainly expresses intentionality while progressives with future time reference are used for plans or arrangements (cf. e.g. Huddleston & Pullum 2002:212; Quirk et al. 1985:214f.; Williams 2002:198).

There is reason to argue that the lexical sense of *going* in *going to* future constructions has become weakened and that the meaning is grammatical rather than lexical. And indeed, the *going to* future is often quoted as a typical example of grammati-

I 'm getting old Yeah Charming I 'm leaving at eight o'clock Are you? Yeah, gon
 'm leaving at ten to eight Eh? I 'm leaving at ten to eight You 've got precisely
 ot going home then ? No and erm I 'm leaving a turkey in the freezer, an Paul is q
 ve I got now Clare? Cos I do, I 'm leaving at ten to eight Eh? I 'm leaving at t
 an me Mum I 'm leaving that, I 'm leaving that, I 'm Mm and that I know we went
 they I wonder why. Your deal. I 'm leaving my two in. They get paid enough goodn
 each, do n't worry I the exam, I 'm leaving at Christmas time. Alright, oh I see
 Mr. Where is The. Opposite I 'm leaving all my mess here because I 'm coming b
 ave n't done the supper dishes. I 'm leaving them now for you, I 'm not doing th
 u do n't change me from this job I 'm leaving, so they says oh we do n't want you t
 hree years younger than me Mum I 'm leaving that, I 'm leaving that, I 'm Mm and
 clock. Ten o'clock! Oh hang on I 'm leaving at ten cos I want to be there a little
 as long as I 've gone by, or I 'm leaving here by eleven give me a half hour to

Figure 83. Extract from a BNC_spoken concordance of *leaving*, illustrating future time reference

by the Causeway? Yes there is. Who is going to check that? Well I can check it. Yea
 wearing protective gear. But I 'm just going to demonstrate to you just what happens.
 als. Yeah. Erm Erm erm. We 're just going to write to local councils arbitrarily a
 go to sleep find time Say they just going for a walk, yeah, and you come across
 be exactly five centimetres. And I 'm going to put a guide line which is one erm er
 ich I shall then put in here, and I 'm going to put the the plate, which I describec
 ibility in the series of diagrams I 'm going to show you in this lecture that we car
 l the boy. Ah, said the old man. I 'm going to live for as long as this forest remai
 id, all females are ladies. I 'm, I 'm going to try a little experiment which er, whi
 , yeah that is, you 're right! I 'm going! What the hell are you doing like that?
 the conference is n't it. Yes. I 'm going to be busy tomorrow. Erm how about I Ar
 ter the most careful clean up. I 'm going to take a, a roller these are very expe
 the things we have to go through. I 'm going back to the erm controversial emotive su
 ie table, at the end of the lesson I 'm going to have those in centre of the table no
 our triangular bandage here. Now I 'm going to show you now how you going to put it
 your breath. Erm other than that I 'm going to write to. And do the other things tr
 ist one minute, erm the poems that I 'm going to read to you, very few of them are ac
 sorry it is the question now that I 'm going to come to was going back to organisat
 not exactly hay-fever but I think I 'm going to be sneezing for a few minutes. Going
 m interested, and in about a week I 'm going to follow it up by writing to them. Right
 note here that you 're you two are n't going away till May. I Uh-huh. presume that
 person who was doing this it was n't going to be appropriate Yeah that 's why it gc
 eld in Glasgow, so I think we were n't going to that one anyway, but it just giving
 ou? Oh No? No. You said you were n't going to. Well why on earth should I say that
 never going to control, you 're never going to be on top of it but I mean No oh no i
) be more common because people are not going for help when they need it. Right, I thi
 ild is level five, I 'm certainly not going to accept that without evidence that I
 / as an English Mm. teacher I 'm not going to argue that but what I 'm saying is tr
 ow you going to use it? Well I 'm not going to actually tell you because I 'll ruin
 ! for yes, button two for no. I 'm not going to pursue anybody on this one so I mean
 oll your trousers up a bit you 're not going to have much bandage to do much bandagir
 well you have, that little bit 's not going to make them any worse, they 've got to

Figure 84. Extract from a BNC_spoken concordance of *going*, exemplifying repeated occurrences of “*going to future*”

calization, i.e. the diachronic development whereby lexical items gradually take on grammatical functions and become semantically bleached and often phonologically reduced, as in the case of *going to* → *gonna* (cf. e.g. Hopper & Traugott 1993 or Heine et al. 1991). The person who says “I’m going to try a little experiment” (cf. line 9 in Figure 84) is not necessarily *going* anywhere to do that, i.e. moving to a different place, but simply expresses what she or he intends to do. Hence, from a semantic point of view, “TO BE *going to* + infinitive” is certainly not a regular progressive but a special kind. In the present analysis, however, this construction has been included. The approach I follow, i.e. CDL, forbids me to dismiss “TO BE *going to*” as not a progressive right from the start, without looking closely at the data.⁷⁴

In this context it is interesting to see that *going* is not the only form that is frequently followed by a *to*-infinitive. Five other verb forms in our set of 99, *expecting*, *hoping*, *meaning*, *trying*, and *wanting*, also show this typical “TO BE V-ing *to* + infinitive” pattern (cf. Figure 85 for illustrative concordance samples). It appears significant

got in here screw hole. Oh I 've been meaning to ask and I keep forgetting. I ca n
me I pass I fume inwardly. I 've been meaning to write to The Post letter page about
recks. Yeah. Well done. I 've been meaning to offload those on Give us a fiver an
him! We let him out. No, I 've been meaning to. Come on Mark! See you later! Marl
I can get. Yeah. Cos I 've been meaning to get myself some. does it round abou
er that 's something else we 've been meaning to do. If you had made a will who wou
Have you? All the jobs you 've been meaning to do Yeah, yeah Yes I know Hello th
ng around since Christmas. What? Been meaning to give them to me. Ever so slimming
ring this current round of er. I 'm trying to establish put down a figure of forty
bragging or anything, Michael, I 'm trying to tell you the situation as it was and
tives for your course. Yeah. I 'm trying to remem you 've put me on the spot now
iring on us right now. Alright, I 'm trying to make contact with the forces now, st
and style with favours. You know I 'm trying to push more you know. I hope er that
a when you get when you get I 'm I 'm trying to do too much at once now it 's okay o
erm to work out I think I think I 'm trying to do too much at once I 'm going to be
ion that now then? I 'm trying to I 'm trying to tie this this erm Actually that 's
ing it very difficult I mean what I 'm trying to say is for goodness sake eyes should
Does anybody not understand what I 'm trying to Oh yes, yes. yes fully I think. It
y good marks, it 's not, they are not trying to test your erm your deep knowledge ab
s day, this was an instance, I 'm not trying to blow me trumpet do n't think that at
o n't you do it this way? they 're not trying to make it harder for you. You might th
red Mm Mm er I mean I 'm not sort of trying to point the finger at her, but No, bec
n't bloody deny it. Yeah I 've been wanting to ask you, what do you think of my l
They 're animals! Yeah! I 've been wanting to get out. Like me, I always Ah! Oh
t the majority of them are definitely wanting to put money into promoting girl 's an
chat. U unless there 's somebody else wanting serving is n't it. If you happen to b
s and carers with trick or treat, is wanting to do is is to make sure that the fun
but she 's recovered. But now she is wanting to be home. Yeah. But er Doctor had
ces. okay right, anything else? just wanting to get away from traditional. Yeah, t
know, people who are, who are, just wanting a crutch because their coming to the e
have a cuddle from me, you 're just wanting to look at what pet? I just want to l
o, just to go out, cos I 'm wanting just to go out on a Monday with her, c
hat I think, but er if people are n't wanting your opinion then you ca n't really gi
but Mm . er at the moment I was n't wanting to s s change any anything that, which
it comes , just And I mean I 'm not wanting to make up any. Ah but Doctor, it 's
t of an idea at this stage, I 'm not wanting to take up ogres of time on this, but
ete bag of nerves and I 'm, I 'm not wanting to be walking along the, the edge of
have got commitments and they 're not wanting to take them on. And then there 's o
and good. Yeah. We 're, we 're not wanting a wage out of it or anything, you know
ad a funny feeling possibly he 's not wanting to do it anyway. does n't know it but

Figure 85. Extracts from BNC_spoken concordances of *meaning*, *trying*, and *wanting*, illustrating “TO BE V-ing to + infinitive” patterns

that the five verbs are semantically related in that they all refer to mental processes which refer to desires or intentions of differing degrees of intensity. The main function expressed in the examples in Figure 85 is probably “emphasis”. The progressive serves to emphasise or put more stress on the actual mental process and on the desire and intention of the speaker.

On the semantic change side, compared to TO BE *going to* + infinitive, the same construction with *expecting*, *hoping*, *meaning*, *trying*, or *wanting* is certainly still much more transparent and not very abstract or even grammaticalized. It might be interesting to see, however, how the patterns “TO BE *expecting to* + infinitive”, “TO BE *hoping to* + infinitive”, “TO BE *meaning to* + infinitive”, “TO BE *trying to* + infinitive”, and “TO BE *wanting to* + infinitive” will develop, what meaning features they may acquire over time, or, in Bloomfield’s (1933:429) terms, what “refined and abstract meanings” may “grow out of more concrete meanings”. What can be said for the time being and with reference to our spoken British English data from the 1990s is that the functional focus in the above-listed constructions is less on the traditional progressive meaning of “something being in progress at a particular time” than on the emphatic meaning of stressing a speaker’s intention or desire.

Finally, a look at the right-hand parts of the diagrams in Table 28 displaying present/future time reference shows that the distribution of these “indeterminate” cases is not at all equal across verb forms. With some verbs we find considerable de-

hey are proud of the forces they are providing. They differ from the authorities not only adequate, have quality and we are providing safe service to the client. He is also understanding the needs of people whom we are providing a service on Thursday, Friday, Saturday. I understand that you know we are providing public service here and we have got a dissatisfaction was and the service we are providing. That's proved to be quite difficult to have a number of dwellings that we are providing, through the structure plan, through out service at all. I think we are providing a good service here and we're meeting for years to come. That way, we are providing those front line services, that Council that people need to know what we are providing so we need to be monitoring, continuing to set up their own new schools which are providing excellent education and really meeting to even challenge the people who are providing the services, or could restrict the

Figure 86. Extract from a BNC_spoken concordance of *providing*, illustrating the connection between present/future time reference and “general validity”

violations from the average value of 15.45 per cent and either comparatively high shares of present/future time reference (e.g. with *costing*, *dealing*, *following*, *keeping*, *paying*, *providing*, *selling*, *spending*, *using*, and *wearing*) or very low shares of tokens that refer to both the future and the present (e.g. with *accepting*, *adding*, *going*, *suggesting*, and *wondering*). It was noted earlier on (cf. 4.4.1 and 4.4.5) that progressives with an “indeterminate” time reference often refer to something that is valid not only at the time of speaking but also for situations to come. Thus, there is obviously a strong overlap between the additional function “general validity” and present/future time reference. Figure 86 displays a few examples from a concordance of the form *providing* which illustrate this connection. Significant here is the frequent collocation with “service” or “services”. We will see later (cf. Section 4.5.11) whether the same verbs that often express an indeterminate time reference also show above-average shares of the additional function “general validity”.

To briefly sum up our findings concerning individual verbs and time reference, we can say that certain verbs favour certain time references and that the average values determined in Section 4.4.2 do not help us to predict the behaviour of a large number of progressives.

4.5.10 Verbs and central functions of the progressive [BNC/BoE]

In Section 4.4 I identified two central function features of progressives, continuousness and repeatedness, and specified the distributions of the four possible feature value combinations, “continuous + repeated”, “continuous + non-repeated”, “non-continuous + repeated”, and “non-continuous + non-repeated”, in the BNC_spoken and BoE_brspok datasets (cf. 4.4.3).

Due to their frequencies of occurrence, the two continuous functions were described as central functions (CFs) of progressives in spoken English, with CF1 referring to continuous and non-repeated actions and events, and CF2 expressing continuousness and repeatedness. It will now be examined if there are any typical relations between individual verb forms and the two central functions. In addition, I will analyse whether the comparatively low shares determined for the occurrence of non-continuous functions (“non-continuous + repeated” and “non-continuous + non-repeated”, cf. Figure 38 in 4.4.3) are representative of all progressive types in the

and he 'd love a bit of yeah, he was asking about you and Ah, I only said to Chr
 us n't I? Yeah Yeah but the way he was asking, again it was, it was a natural floo
 I? Well it 's up, you know, all I was asking I was just querying you said you, you
 buy anything from Chile, and er I was asking the young man who was is this from Ch
 es in my scrapbook where erm it, I was asking, telling Jean she would have to look
 morrow than to London Zoo. Oh. I was asking, Would you? Probably. Yeah. to Lond
 ar either. Turn that thing off. I was asking what sub committee it was. Er th
 is asking, for a number of reasons I was asking for floor plans for our, you see. Di
 l yes, I we I explained why I wa I was asking the question, I said it was for my s
 finitely hey. Well sometimes when I was asking you things like erm say if I asked y
 ll, why? No, it 's just that Jill was asking the other day and I said I ca n't rem
 but for a Sunday lunch. Oh, Ken was asking us about that. Three ninety five. Oh
 then you go back with what Lord was asking you about and I just want you to expl
 it. And I think what the questioner was asking for is, is probably adequately cover
 nd at all. Yeah I know cos Robert was asking for his So I I, I, I went in today I
 as in bed a couple of weeks ago she was asking erm how grateful she was to your parei
 asking wh who were they? Think she was asking your mum who were they. But er anyhoi
 jood, eh I just wanted to find She was asking me what Brian it was, you were talkin
 und the world ticket? Well it 's only costing Eddie 's brother a hundred erm one tho
 what good value it was, it was only costing them hundred and Well I really ca n't
 've paid for it. But it 's probably costing us about er, three quarters of a mill
 ll, I do n't suppose this contract 's costing us twenty-one million, but it 's cost
 ny thing else or else this holiday 's costing me a fortune. I think you can manag
 ng from all over the county and it 's costing us a fortune. What But you say somet
 ing us twenty-one million, but it 's costing us several million, and I would have
 ne hundred pounds. Only! But it 's costing Eddie about nine hundred pound, and t
 at What what is that costing? It 's costing nothing, the day. It 's just the day
 : says we ca n't, he says well I 'm not paying you hundred and forty pound for a part
 ead oh, he might have one on I 'm not paying that price, he charges about twice as
 uncly stuff there. Cos she said I 'm not paying a fortune for stuff that I wo n't wear
 a day, a fortnight ha No. I 'm not paying another month 's rent, I know that we
 u? Yeah but I 'm saying they 're not paying they have n't got ta pay for that mach
 year? And again like we no we 're not paying er personally for our bridal magazine
 said I 've You paying? Ay You 're not paying for mine are you? No did you want r

Figure 87. Extracts from BNC_spoken concordances of *asking*, *costing*, and *paying*, illustrating the function “non-continuousness + non-repeatedness”

BNC/BoE dataset. The percentages of CF1, CF2, and both non-continuous functions for each of the 99 verb forms are listed in Table 29 below. The bar charts included in the table visualise the individual distributions.

A quick glance at the charts in Table 29 reveals that there are big differences between verb form types with respect to progressive functions. If we just look at the shares of CF2 (continuous + repeated actions and events), we notice a high degree of inter-verb variation. Some verb forms never or extremely rarely express continuousness and repeatedness (e.g. *asking*, *costing*, *paying*), while others predominantly refer to continuous and repeated events (e.g. *cutting*, *helping*, *working*). This observation also applies to the other three function feature combinations. Individual percentages vary a lot and deviate significantly from the determined average values.

Besides, a closer look at the mini-diagrams shows that a distinction can be made between “continuous” verbs and “non-continuous” verbs, i.e. between verbs which either favour CF1 and CF2 or the two non-continuous functions. The following forms clearly fall into the latter category: *asking*, *buying*, *calling*, *costing*, *paying*, *ringing*, *saving*, *sending*, *spending*, *suggesting*, and *winning*. Whereas *asking*, *calling*, *ringing*, *suggesting*, and *winning* show highest shares of non-continuousness + non-repeatedness, the other six types typically occur in non-continuous + repeated contexts. Figure 87 combines extracts from BNC_spoken concordances of *asking*, *costing*, and *paying* which exemplify this typical co-selection of verb form and function.

The large majority of verbs, however, belong to the “continuous” group, and we find high shares of the central function CF1 for many of the forms listed in Table 29 (in particular with *expecting*, *finishing*, *hoping*, *meaning*, *staying*, *wanting*, and *wondering*).

Table 29. The distribution of verb form tokens across the four combinations of the central function features “continuousness” and “repeatedness” for each analysed verb (verbs in alphabetical order)

verb form	cont. + repeated (CF2)	cont. + non-rep. (CF1)	non- cont. + repeated	non- cont. + non-rep.	distribution graphically illustrated
	ø27.47%	ø54.21%	ø8.78%	ø9.24%	■ CF2 ■ CF1 □ non-cont + rep. ■ non-cont + non-rep.
accepting	54.55%	45.45%	0.00%	0.00%	
adding	67.21%	32.79%	0.00%	0.00%	
agreeing	23.08%	76.92%	0.00%	0.00%	
asking	0.00%	0.00%	32.45%	67.55%	
being	37.50%	62.50%	0.00%	0.00%	
becoming	37.20%	62.80%	0.00%	0.00%	
believing	0.00%	100.00%	0.00%	0.00%	
betting	85.71%	14.29%	0.00%	0.00%	
bothering	21.43%	78.57%	0.00%	0.00%	
bringing	37.78%	62.22%	0.00%	0.00%	
buying	1.72%	1.72%	60.34%	36.21%	
calling	0.00%	0.81%	45.16%	54.03%	
carrying	40.19%	59.81%	0.00%	0.00%	
changing	47.19%	52.81%	0.00%	0.00%	
checking	20.00%	23.53%	27.06%	29.41%	
coming	36.00%	64.00%	0.00%	0.00%	
costing	0.00%	0.00%	74.80%	25.20%	
cutting	78.46%	21.54%	0.00%	0.00%	
dealing	58.59%	41.41%	0.00%	0.00%	
doing	59.39%	40.61%	0.00%	0.00%	
drawing	53.97%	46.03%	0.00%	0.00%	
eating	53.64%	46.36%	0.00%	0.00%	
expecting	12.28%	87.72%	0.00%	0.00%	
explaining	18.75%	81.25%	0.00%	0.00%	
feeling	24.00%	76.00%	0.00%	0.00%	
finding	52.75%	47.25%	0.00%	0.00%	
finishing	6.02%	93.98%	0.00%	0.00%	
following	41.46%	58.54%	0.00%	0.00%	
forgetting	54.05%	45.95%	0.00%	0.00%	
getting	40.28%	59.72%	0.00%	0.00%	
giving	23.02%	23.02%	26.19%	27.78%	
going	23.56%	76.44%	0.00%	0.00%	
happening	43.16%	56.84%	0.00%	0.00%	
having	44.29%	55.71%	0.00%	0.00%	
hearing	59.70%	40.30%	0.00%	0.00%	
helping	76.47%	23.53%	0.00%	0.00%	
holding	27.84%	72.16%	0.00%	0.00%	
hoping	2.62%	97.38%	0.00%	0.00%	
imagining	13.64%	86.36%	0.00%	0.00%	
keeping	30.86%	69.14%	0.00%	0.00%	

Table 29. (continued)


























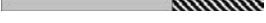

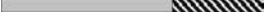
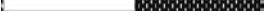

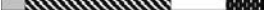
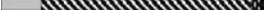
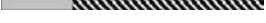
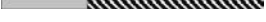

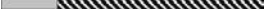
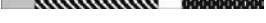




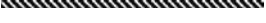
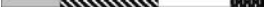
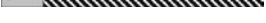
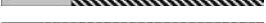

knowing	0.00%	100.00%	0.00%	0.00%	
learning	62.12%	37.88%	0.00%	0.00%	
leaving	18.18%	59.09%	4.55%	18.18%	
letting	51.04%	48.96%	0.00%	0.00%	
liking	37.50%	62.50%	0.00%	0.00%	
listening	12.18%	87.82%	0.00%	0.00%	
living	0.00%	100.00%	0.00%	0.00%	
looking	27.40%	56.85%	5.48%	10.27%	
making	36.54%	39.42%	6.73%	17.31%	
meaning	5.41%	94.59%	0.00%	0.00%	
meeting	23.08%	23.08%	0.00%	53.85%	
moving	29.67%	70.33%	0.00%	0.00%	
needing	35.71%	64.29%	0.00%	0.00%	
paying	0.67%	3.33%	82.00%	14.00%	
picking	19.82%	36.04%	29.73%	14.41%	
playing	30.00%	70.00%	0.00%	0.00%	
providing	59.32%	13.56%	13.56%	13.56%	
pulling	43.43%	56.57%	0.00%	0.00%	
putting	28.71%	38.61%	14.85%	17.82%	
reading	34.72%	65.28%	0.00%	0.00%	
remembering	26.67%	73.33%	0.00%	0.00%	
ringing	3.97%	12.70%	26.98%	56.35%	
running	35.42%	64.58%	0.00%	0.00%	
saving	14.04%	7.02%	61.40%	17.54%	
saying	24.84%	75.16%	0.00%	0.00%	
seeing	65.33%	34.67%	0.00%	0.00%	
seeming	50.00%	50.00%	0.00%	0.00%	
selling	65.38%	34.62%	0.00%	0.00%	
sending	0.00%	0.98%	50.00%	49.02%	
setting	25.93%	74.07%	0.00%	0.00%	
showing	9.09%	55.56%	21.21%	14.14%	
sitting	15.20%	80.00%	0.00%	4.80%	
sorting	27.27%	72.73%	0.00%	0.00%	
speaking	32.86%	67.14%	0.00%	0.00%	
spending	20.78%	2.60%	67.53%	9.09%	
standing	21.51%	78.49%	0.00%	0.00%	
starting	12.73%	47.27%	9.09%	30.91%	
staying	11.11%	88.89%	0.00%	0.00%	
stopping	2.78%	46.30%	26.85%	24.07%	
suggesting	0.00%	0.00%	13.45%	86.55%	
supporting	43.48%	56.52%	0.00%	0.00%	
supposing	0.00%	100.00%	0.00%	0.00%	
taking	22.50%	37.50%	27.50%	12.50%	
talking	15.96%	84.04%	0.00%	0.00%	
telling	26.88%	73.13%	0.00%	0.00%	
thinking	15.60%	84.40%	0.00%	0.00%	

Table 29. (continued)

trying	30.43%	69.57%	0.00%	0.00%	
turning	18.75%	56.25%	7.50%	17.50%	
understanding	33.33%	66.67%	0.00%	0.00%	
using	63.74%	36.26%	0.00%	0.00%	
walking	21.65%	78.35%	0.00%	0.00%	
wanting	17.12%	82.88%	0.00%	0.00%	
watching	27.19%	72.81%	0.00%	0.00%	
wearing	39.60%	60.40%	0.00%	0.00%	
winning	4.00%	12.00%	10.00%	74.00%	
wondering	4.60%	95.40%	0.00%	0.00%	
working	77.59%	22.41%	0.00%	0.00%	
worrying	33.33%	66.67%	0.00%	0.00%	
writing	40.32%	59.68%	0.00%	0.00%	

Still, it is important to note that these shares in most cases deviate considerably from the average value specified earlier on. I therefore consider it important, when talking about progressive functions, to distinguish between different types or groups of verbs, such as “continuous” vs. “non-continuous” verbs, instead of treating all verbs as one and making misleading statements about functions of “the progressive”.

4.5.11 Verbs and additional functions of the progressive [BNC/BoE]

I noted above that the specified central functions, CF1 and CF2, do not suffice to fully capture what progressives in spoken English denote, and identified a couple of other functions that may co-occur with the central ones. These additional functions were dealt with in Section 4.4.5. I shall now look at relations between individual verbs and the functions “general validity”, “politeness or softening”, “emphasis or attitude”, “shock or disbelief”, “gradual change and development”, “old and new habits”, and “framing”.

Table 30 gives the distributional values of these functions for each analysed verb form and highlights shares which lie 3% or more above average. The following paragraphs will comment on the distributions of each additional function across verbs and discuss in how far the functions are lexically determined.

Verbs and general validity. The discussion of “general validity” in Section 4.4.5 based on the group of 99 verbs lead to the conclusion that, like CF1 and CF2, “general validity” can be called a function of the progressive as such. We found that this frequent function occurs with a large number of verbs and not just with a few particular ones.

However, if we look at the “general validity” values for each verb form in Table 30, it becomes apparent that some verbs show a clear preference for this function while others very seldomly refer to generally valid actions or events. Evidently, there is a connection between “general validity” and the types *buying, costing, dealing, eating, forgetting, helping, keeping, paying, providing, saving, selling, spending, supporting,*

Table 30. Frequencies of occurrence of additional progressive functions across verb forms (shares of $\geq 3\%$ above average are shaded grey)

verb form	general validity ø22.35%	polite- ness/sof- tening ø12.82%	empha- sis/ shock ø9.46%	gradual change ø5.17%	old habit ø1.10%	new habit ø0.44%	framing ø0.89%
accepting	40.91%	0.00%	0.00%	1.52%	0.00%	9.09%	0.00%
adding	19.67%	0.00%	3.28%	6.56%	8.20%	1.64%	0.00%
agreeing	17.31%	17.31%	19.23%	1.92%	0.00%	0.00%	0.00%
asking	7.28%	27.15%	49.67%	0.00%	0.00%	0.00%	0.00%
being	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
becoming	0.00%	0.00%	0.61%	98.78%	0.00%	0.00%	0.00%
believing	50.00%	0.00%	16.67%	0.00%	0.00%	0.00%	0.00%
betting	57.14%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
bothering	16.07%	3.57%	28.57%	1.79%	0.00%	5.36%	0.00%
bringing	14.44%	4.44%	0.00%	6.67%	3.33%	1.11%	1.11%
buying	36.21%	0.86%	1.72%	0.00%	0.00%	4.31%	0.00%
calling	12.90%	2.42%	5.65%	0.00%	6.45%	1.61%	0.00%
carrying	16.82%	1.87%	0.93%	0.00%	3.74%	0.00%	4.67%
changing	20.22%	0.00%	6.74%	52.81%	3.37%	0.00%	1.12%
checking	14.12%	32.94%	1.18%	0.00%	0.00%	0.00%	1.18%
coming	12.00%	1.00%	0.00%	8.00%	0.00%	1.00%	1.00%
costing	59.84%	12.60%	62.20%	0.00%	0.00%	0.00%	0.00%
cutting	27.69%	0.00%	1.54%	3.08%	0.00%	0.00%	1.54%
dealing	35.16%	0.00%	0.00%	0.78%	0.00%	0.00%	0.78%
doing	28.48%	0.00%	0.61%	0.00%	0.00%	0.00%	0.61%
drawing	25.40%	0.00%	0.00%	0.00%	0.00%	0.00%	3.17%
eating	33.64%	0.91%	0.91%	0.91%	0.00%	0.91%	5.45%
expecting	11.40%	4.39%	0.00%	0.00%	0.00%	0.00%	0.00%
explaining	10.42%	0.00%	0.00%	0.00%	0.00%	0.00%	4.17%
feeling	30.67%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
finding	31.87%	9.89%	2.20%	0.00%	0.00%	0.00%	1.10%
finishing	4.82%	10.84%	0.00%	0.00%	0.00%	0.00%	3.61%
following	29.27%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
forgetting	45.95%	62.16%	0.00%	18.92%	0.00%	0.00%	0.00%
getting	22.22%	2.78%	0.69%	43.75%	0.00%	0.00%	0.00%
giving	26.98%	6.35%	11.90%	0.00%	0.00%	0.00%	2.38%
going	16.83%	2.88%	1.92%	1.92%	0.00%	0.00%	1.44%
happening	18.38%	0.00%	0.00%	0.00%	0.00%	0.00%	0.85%
having	20.00%	1.43%	0.00%	0.00%	0.00%	0.00%	1.43%
hearing	22.39%	10.45%	4.48%	0.00%	0.00%	0.00%	0.00%
helping	40.00%	2.35%	1.18%	1.18%	0.00%	0.00%	0.00%
holding	23.71%	0.00%	0.00%	0.00%	0.00%	0.00%	1.03%
hoping	3.06%	20.09%	79.48%	0.00%	0.00%	0.00%	0.00%
imagining	4.55%	36.36%	0.00%	0.00%	0.00%	0.00%	4.55%
keeping	35.80%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Table 30. (continued)

knowing	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
learning	21.21%	0.00%	0.00%	10.61%	0.00%	0.00%	0.00%
leaving	14.55%	0.00%	7.27%	0.91%	0.00%	0.00%	3.64%
letting	30.21%	8.33%	16.67%	0.00%	0.00%	0.00%	0.00%
liking	37.50%	62.50%	0.00%	12.50%	0.00%	0.00%	0.00%
listening	8.97%	5.77%	10.90%	0.00%	0.00%	0.00%	1.28%
living	28.57%	3.90%	1.30%	0.00%	50.65%	2.60%	0.00%
looking	16.44%	0.00%	1.37%	0.68%	1.37%	0.00%	0.00%
making	20.19%	15.38%	7.69%	7.69%	0.00%	0.00%	0.96%
meaning	5.41%	56.76%	32.43%	0.00%	0.00%	0.00%	0.00%
meeting	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
moving	24.18%	0.00%	5.49%	20.88%	0.00%	0.00%	0.00%
needing	35.71%	16.67%	69.05%	0.00%	0.00%	0.00%	0.00%
paying	70.67%	84.00%	4.00%	0.00%	0.00%	0.00%	0.67%
picking	31.53%	7.21%	8.11%	7.21%	0.00%	0.00%	0.00%
playing	17.27%	0.00%	0.91%	0.91%	0.00%	0.00%	1.82%
providing	66.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
pulling	24.24%	2.02%	4.04%	1.01%	0.00%	0.00%	4.04%
putting	22.77%	3.96%	4.95%	4.95%	0.00%	0.00%	1.98%
reading	26.39%	0.00%	4.17%	0.00%	0.00%	0.00%	4.17%
remembering	26.67%	20.00%	0.00%	6.67%	0.00%	0.00%	0.00%
ringing	15.08%	36.51%	10.32%	0.00%	0.00%	0.00%	1.59%
running	22.92%	0.00%	2.08%	6.25%	0.00%	0.00%	0.00%
saving	35.09%	59.65%	1.75%	0.00%	0.00%	0.00%	0.00%
saying	18.30%	58.17%	9.15%	0.00%	0.00%	0.00%	0.00%
seeing	29.33%	0.00%	22.67%	1.33%	17.33%	4.00%	0.00%
seeming	0.00%	0.00%	100.0%	0.00%	0.00%	0.00%	0.00%
selling	40.38%	0.00%	0.00%	0.00%	0.00%	1.92%	0.00%
sending	30.39%	1.96%	3.92%	0.00%	0.00%	0.00%	0.00%
setting	12.96%	0.00%	0.00%	46.30%	0.00%	0.00%	1.85%
showing	17.17%	4.04%	12.12%	0.00%	0.00%	0.00%	1.01%
sitting	17.60%	0.00%	0.00%	0.00%	0.00%	0.00%	2.40%
sorting	15.15%	0.00%	9.09%	0.00%	0.00%	0.00%	6.06%
speaking	22.86%	14.29%	0.00%	0.00%	0.00%	0.00%	4.29%
spending	45.45%	12.99%	0.00%	0.00%	0.00%	0.00%	0.00%
standing	18.28%	0.00%	0.00%	1.08%	0.00%	0.00%	0.00%
starting	15.45%	7.27%	0.91%	50.91%	0.00%	0.00%	0.91%
staying	11.85%	0.00%	2.22%	0.00%	14.81%	0.00%	0.74%
stopping	20.37%	1.85%	8.33%	2.78%	0.00%	0.00%	0.00%
suggesting	6.28%	65.92%	29.60%	0.00%	0.00%	0.00%	0.00%
supporting	37.68%	1.45%	4.35%	0.00%	0.00%	0.00%	0.00%
supposing	0.00%	100.0%	0.00%	0.00%	0.00%	0.00%	0.00%
taking	23.33%	0.83%	7.50%	9.17%	0.00%	0.00%	0.00%
talking	14.36%	0.00%	3.72%	0.00%	0.00%	0.00%	0.53%
telling	16.88%	30.00%	31.25%	0.00%	0.00%	0.00%	0.63%
thinking	15.60%	54.61%	3.55%	0.00%	0.00%	0.00%	0.00%

Table 30. (continued)

trying	25.54%	15.22%	5.98%	1.63%	0.00%	0.00%	0.54%
turning	12.50%	2.50%	1.25%	18.75%	0.00%	0.00%	0.00%
understanding	33.33%	11.11%	55.56%	11.11%	0.00%	22.22%	0.00%
using	35.16%	1.10%	2.20%	0.00%	0.00%	0.00%	0.00%
walking	19.59%	4.12%	0.00%	1.03%	0.00%	0.00%	4.12%
wanting	30.63%	17.12%	74.77%	0.00%	0.00%	0.00%	0.90%
watching	21.93%	0.00%	3.51%	0.00%	0.00%	0.00%	1.75%
wearing	32.21%	0.00%	7.38%	0.00%	0.00%	0.00%	0.00%
winning	8.00%	2.00%	4.00%	0.00%	0.00%	0.00%	0.00%
wondering	4.18%	88.70%	1.26%	0.00%	0.00%	0.00%	0.00%
working	55.17%	1.72%	1.72%	0.00%	0.00%	0.00%	0.00%
worrying	33.33%	6.67%	15.00%	0.00%	0.00%	0.00%	0.00%
writing	20.97%	1.61%	1.61%	0.00%	0.00%	0.00%	3.23%

bloke there. That bloke, the council are paying all that on his mortgage. He 's mov
1 term corruption in that the landlords are paying backhanders to them so that the lar
2 them off then? The old age pensioners are paying for a caravan including these, you
3 Post every day you know where people are paying a few hundred pounds, or it may be
4 the ones then that like Linda said are paying two pound fifty then they are So the
5 their rates it 's because they are paying that 's all! Yeah. Phoned up to see
6 Yeah. the bowling Club. We are, we are paying nine hundred pound at the moment. S
7 I started to ba er buy my house, we are paying, we were paying twenty seven pound
8 with people now. Yeah. People who are paying seventy eighty pound a month mortga
9 are already paying are the one 's who are paying and nobody seems to chasing after th
10 at about the other eighty thousand who are paying it and you 're making them pay more
11 or the benefit, and she thought he 'd been paying the rent. And she came in with and
12 er out the factory! Cos he had n't been paying his maintenance you see. Mother use
13 it 's been like that for oh she 's been paying round about twelve pounds a month n
14 er, the rates yesterday, cos I 've been paying my mortgage over ten ye, over the
15 out then dun n it? does it? I 've been paying for me petrol with me money me fami
16 bearing in mind the premiums I 've been paying before my endowments would have bee
17 I even noticed then like when they 're wearing all your clothes right? And I do n
18 material to so many other uses. We 're wearing them, one housewife remarked. Thi
19 the token dishing out tour and you 're wearing them, you get your token, okay? C
20 sibly twice, or three times if you 're wearing make up. Okay? We normally say cl
21 you have a skirt on, or if you 're wearing a dressing gown when sh, Sarah come
22 cca 's swimsuit, he said cor if you 're wearing that I think I 'll go swimming as w
23 you 've got that one. I'll do it, he usually does it if you 're wearing make up, you can use this cleanser
24 and we put towels over you. If you 're wearing something like that coat. He wo, h
25 tions, whatever it is, yes? If you 're wearing a bra we just undo your bra, okay?
26 feel better with people knowing you 're wearing the ear defenders when the supervi
27 're wearing a wig They do n't know I 'm wearing

Figure 88. Extracts from BNC_spoken concordances of *paying* and *wearing*, illustrating high shares of the additional function “general validity”

using, *wearing*, and *working* (cf. also Table 13 in 4.4.5). A very similar list of verb forms was given above in connection with “indeterminate” time reference (cf. 4.3.9), which means that the same verbs which refer to present/future (or timeless) actions or events also commonly express general validities. The concordance extracts in Figure 88 present examples of progressives of *paying* and *wearing*, referring to generally valid situations. Significant in this context is the frequent collocation of *wearing*-progressives with *if*.

Verbs and politeness or softening. Table 30 also highlights a number of verbs which commonly convey politeness or soften the threat of an utterance. The forms *asking*, *checking*, *forgetting*, *imagining*, *meaning*, *paying*, *ringing*, *saying*, *suggesting*, *thinking*,

d, well I said I 've got one oh you 're paying me too much now. Said I 'm, you er
 in a , a least and you you 're you 're paying ten and he said, no he said, he sai
 nst er, doing two plans, which you 're paying the true cost of the benefits you wou
 then you can do the tape while you 're paying ca n't you? Yeah. Tape while you pay
 you think about it, it 's wrong you 're paying tax on your wages at source and ther
 f then. But again, you know, you 're paying here premiums for twenty thousand pol
 paid, no other bills paid, all he 's paying is the gas, electric and the water r
 fe who has no income and her husband 's paying twenty five percent or forty percent
 le? We 'll have a female, and she 's paying, let's say escalating premiums. The
 s a member of the family because she 's paying so much over there. The boy sitting
 er they had to repossess it and he was paying. And I do n't know how, what 's happ
 that rate and they got it. But he was paying them sixpence above the rate so there
 eek or something for him. Mm. He was paying it and I must make sure that he is pa
 of twenty five quid on the seat. I 'm suggesting we send that man a voucher for fi
 h just because of the weekend. I 'm suggesting it may be bigger. Yeah. It may a
 ustry day with them I 'm trying, I 'm suggesting er, in January. Right. But I ne
 at you 're proposing. Yeah, I, I 'm suggesting that the facilities management co
 where, er as opposed to, I mean, I 'm suggesting at the moment we take stock a lo
 every woman to take me H R T, I 'm suggesting that it 's preventative medicine.
 done wickedly, and deliberately, I 'm suggesting with hindsight it 's sloppy and,
 e congress system. Was that in? I 'm suggesting that by perhaps the eighteen twen
 nd that the erm er although, now I 'm suggesting you do this, but if you went thr
 r pupils er relative success. So I 'm suggesting that that, we could take that bo
 B and erm in my response to that I 'm suggesting, and I hope it 's not just seman
 still in obedience is n't it? what I 'm suggesting really, is let's get it on the a
 Council on how to do that. what I 'm suggesting is is that that is perhaps someth
 wing us that one miracle but what I 'm suggesting to you here is that you can look
 at relatively little cost. what I 'm suggesting is that it ought to be possible t
 hat. Well Do n't do that. what I 'm suggesting do n't fall into that category.
 'd ask you to go along with what I 'm suggesting, that 's an absolute discharge a
 when we come to them. So what I 'm suggesting is that A is in fact your address
 riots. Erm I think I think what I 'm suggesting is that the the the the panel sh
 orward a little bit. Well, what I 'm suggesting doing is thi it, what are my out
 So, and disobedience. So what I 'm suggesting you do is, decide roughly what y
 e done in the No. No, what I 'm suggesting is lesson. that the kids erm as
 You see I am, you can see what I 'm suggesting to you, there 's not much point

Figure 89. Extracts from BNC_spoken concordances of *paying* and *suggesting*, illustrating high shares of the additional function “politeness/softening”

and *wondering* are comparatively frequent in politeness contexts. It is interesting to see that many of these verbs frequently collocate with “just”, an adverbial which helps to enhance the downtoning effect (cf. 4.5.8).

Some of the verbs, *asking*, *checking*, *ringing*, *saying*, and *suggesting*, express an activity of the speaker that might mean a disturbance for the addressee. In such situations the progressive can take over the required softening function (cf. some of the *suggesting* examples in Figure 89). In many of the examples with *paying* in the same figure we also notice a softening effect. Apparently, people wish not to sound too direct when they talk about money and choose progressive instead of simple forms.

Verbs and emphasis or attitude, shock or disbelief. We noted earlier on that although 62 out of 100 verbs show occurrences in the “emphasis/attitude” or “shock/disbelief” function, only a few of them have a typical preference for these functions. Table 30 shows which verb forms are particularly often used to emphasise an event/action or to express strong surprise or severe doubt about something.

Clearly above-average percentages for the emphasis or shock function can be found with *asking*, *bothering*, *costing*, *hoping*, *meaning*, *needing*, *seeing*, *suggesting*, *telling*, *understanding*, and *wanting*. It may be surprising that some of the “politeness” verbs identified in the previous section also appear in the “emphasis” list (e.g. *asking*, *hoping*, *suggesting*, and *telling*). We might wonder how it can be that the same form is used for two opposing purposes, to soften an utterance and to put emphasis on it. I argue that the concordance extract of *suggesting* in Figure 90 provides a solu-

've got a heavy mortgage, and I 'm not suggesting that many of you will have a heavy mortgage. Well although I mean it I 'm not suggesting that is n't cheap but erm to proceed to finish at four o'clock but I 'm not suggesting we can get through all the business before it be out of date and therefore I 'm not suggesting deliberately, but innocently, a suggestion at all costs no I 'm not I 'm not suggesting for a moment it does n't cross my mind to work on that presumption, and I 'm not suggesting you creep around frightened of what it has to go from there. I 'm not suggesting that you, you 're right, I do n't want ribbons and things I do n't no, I 'm not suggesting will you find out now, but there is what we have to start doing and I 'm not suggesting you 're not doing, but I just want to erm criteria two. Perhaps I 'm not suggesting this as a sort of erm a definite suggestion. So if we do under-fund, and I 'm not suggesting in these proposals that there is a possibility, yes it could yeah, I 'm not suggesting it was done wickedly, and deliberately, and sacrifices and all sorts. I 'm not suggesting we should drag all these things into it, it 's all marvellous, and I 'm not suggesting The big Christmas, big Christmas, let me make it absolutely clear I 'm not suggesting er on behalf of Julie that there is er, one of two of the guests, I 'm not suggesting your mother 's guests er over-imbudged or reasons. Oh, no, oh, no, I 'm not suggesting you do, but it just seems unkind to me to have to Havstock Park but I 'm not suggesting you know the switch just because it just has to be shared. You 're not suggesting pregnancy 's a disease there are a lot of women I 'm scared now I hope you 're not suggesting that er women take to their cars

Figure 90. Extract from a BNC_spoken concordance of *suggesting*, illustrating high shares of the additional function “emphasis/attitude”

I see that, I see that, are, are you suggesting that those provisions in the membership relax into your presentation. Are you suggesting that Absolutely, and you will find out of bone less in a year! Are you suggesting that every woman going through the pregnancy well what sort of time are you suggesting? well I have n't discussed this with them. They will hear. They wo n't Are you suggesting that Jean has nothing interesting to say, okay so what are you s what are you suggesting then? Well if I can just draw your

Figure 91. Extract from a BNC_spoken concordance of *suggesting*, illustrating the additional function “shock/disbelief”

tion to this problem: All examples in this concordance are emphatic, and all of them are negated, a pattern that was also found with *asking* and *telling*. This means that there is a close relation between form and function, between a particular pattern and a particular meaning. In affirmative contexts, *suggesting* progressives frequently express politeness or softening, whereas their negative forms sound emphatic.

Examples of *suggesting* in the “shock or disbelief” function are given in Figure 91 (see also the last two lines in Figure 90). An apparent pattern here is “are you suggesting”, so again we find a pattern-meaning connection between the form of a question and the function of strong surprise or disbelief.

Verbs and gradual change and development. In a sub-section of Section 4.4.5 we identified a set of verbs that repeatedly indicate the gradual change of a situation or refer to a development of some kind (cf. Table 17). The values in Table 30 now show us which of the verbs listed in Table 17 are the most typical “gradual change” verbs.

The highest percentage by far is found with *becoming*, unsurprisingly the most prototypical verb form in this function. Part of the BNC_spoken concordance of *becoming* is displayed in Figure 92. *Changing*, *getting*, *moving*, *setting*, and *starting* also show very high shares of “gradual change” instances. This function is of course part of the semantics of these verbs, but the progressive certainly helps to emphasise the aspect of an ongoing development and stresses the temporal extension of the respective actions (cf. Figure 93 for illustrative evidence). Interesting also are the above-average shares of the “gradual change and development” function found with the semantically

rumbling away. And so it was actually becoming dangerous at the same time. The, the then easing out the sheet. It's all becoming a more natural process. Well how was it so often, that they're also almost becoming well can you think of any words that with which Your Lordships are already becoming very familiar. My Lords, I I'm lik peasants who were taking the lead and becoming members of the Party. Because i in, ruitment base, large factories, are becoming a thing of the past. Fourteen years ons of peasants are joining Mn. are becoming involved in a peasant movement, are mean I think erm the the cutbacks are becoming more and more and more and people I think nowadays that the doctors are becoming very much more educated and perhaps t the dangers is that many drivers are becoming impatient with waiting their turn in ource. Meanwhile, our employers are becoming more involved, and for a variety of ped to deal with them and firemen are becoming more expert in, in handling these er e are now dependant on aid. Many are becoming increasingly desperate and turning th children! When children nowadays are becoming more responsible to our society, and tions means traditional coal ones are becoming surplus to requirements. Productio s being built and a lot of people are becoming christians in Kenya. Nepal is diffe . So that now at least the people are becoming more aware, and people are more awar quite important. I think people are becoming more and more and it's not true finz me more, not less common. They are becoming a lot more common, yes when you look of buying travel services. They are becoming increasingly more important as they v trade union movement that I think are becoming somewhat of a rarity these days I kno representing the whole workforce are becoming an increasing interest to the U K trs had becoming er had, had, had been becoming frighteningly er accurate say in eight island in which we live is in effect becoming smaller day-by-day, as it is becomir

Figure 92. Extract from a BNC_spoken concordance of *becoming*, illustrating the additional function “gradual change and development”

the standards that are being set are changing and at the end of the day the hunt ant sees the world. Er that they are changing a whole range o of non-economic va her and further behind but things are changing economic and political form is tak , that something is, that things are changing and that the parish council is of changing, the scriptures tell us that the at this particular time, things are changing, present time er which has so much, have changed round here, things are changing the basis of the calculation of th Er and and therefore er unless we are changing the application or deleting Yes th th Q P5, I am not sure whether we are changing erm over recent decades. There is x because the law of the sea has been changing their ways, they now go half way the last twenty years they have been changing the pipes or digging them up or so and the chlorine. If they've been changing some every bloody minute I think, 's always playing! He 's changing, changing their own dwellings, but it 's si do suggest there are more households, changing to the caring professions. I mean nd so on. And the nature of caring is changing as time goes on. This is over the ry, that 's its, how its distance is changing

Figure 93. Extract from a BNC_spoken concordance of *changing*, illustrating the additional function “gradual change and development”

related verb forms *learning*, *forgetting*, and *understanding*. Obviously, the progressive aids to stress the gradual procedural sense that is part of the meaning of these items.

Verbs and old and new habits. It was noted in Section 4.4.5 that only a limited number of verbs in our spoken BrNSE dataset express the rather rare progressive function of habituality. The values displayed in Table 30 support this finding. Just nine out of 99 verbs refer to what we labelled “old habits”. Highest shares were found with *living*, *seeing*, and *staying*, while the percentages for *adding* and *calling* are much lower but still clearly above the average value of 1.10%. Figure 94 presents two samples from concordances of *living* and *staying* that serve to illustrate this function.

Close relations between function and lexical items also occur in the case of “new habits” where a more or less significant co-selection was only detected for the new habit function and *accepting*, *bothering*, *buying*, *seeing*, and *understanding*, and even with these items the values are comparatively low (cf. Table 30). A few examples which refer to new “buying” habits are displayed in Figure 95. Typical here is the pattern “I’m/am not buying” and the collocation with “no more” or “any more”.

reason why I went there, my brother was living two or three here. And he said s
 I think his description of where he was living in the tree house was very clear :
 neck down. This is the one who 's was living with another woman and his Yeah. v
 got a husband? As far as I know she was living on her own with this little old be
 here they on holiday? No, she was living out there. she had a Oh I see! He
 ' Ah, no and you went to and somebody was living in there no, no love, because was
 citizen. You know cos, it you was living at Yeah. Like her mum, she 's ser
 it any local authority wherever Paul were living would find it possible to that he
 re on a course at Donnington, they were living in, they invited me over, I said
 them, alright, perhaps whilst they were living with Naomi perhaps she could pull
 er their weekly ration in But you were living on credit that was like credit to
 this very common amongst people you were living with, were they, was it very cor
 : have a soft bed if er you were you were living rough. And of course er er i i I
 t affect your life much? No. Were you living in the s in the h house with the c
 of food did you get, and where were you living at the time? Er, well there was
 did n't want to leave? Yes. Oh! He was staying there. And we I was I was trying
 en. God! What, what what exactly I was staying on I said oh er, yeah cos yo, had
 well received, and the person that I was staying with, lived in a flat, a very simp
 'es I well remember er I use when I was staying down at Trimley there was erm a lac
 I went, my mother, where I, where I was staying at er Dunyask, there was five metr
 Well that 's right. That 's where I was staying. as well, tha we went there did r
 rmany f at Easter. Aha. A woman I was staying with made a wreath like that Aha. s
 l she? Mm. Ooh . Yeah. Yeah cos she was staying here. Really? Oh she left this mo
 got them insured? So I 've got She was staying with the tech was n't she? Was she
 lady laughed at me and she said, she was staying in the same hotel and she said er :

Figure 94. Extracts from BNC_spoken concordances of *living* and *staying*, illustrating the additional function “old habit”

re 's granddad 's kiss? Not. I 'm not buying you no more then! Oh! I 'll get you
 ina get me then? something. I 'm not buying you anything. Oh thanks a lot. No I
 kick you out. I was saying I 'm not buying no more! Can you. No, it we over th
 about ten pencil cases okay. I am not buying any more pencil cases. They are Amy
 ave n't got I have n't got I am not buying any more pencil cases. No you go and
 bought so many yesterday, I 'm never buying another pair in my life. Why? So doe

Figure 95. Extract from a BNC_spoken concordance of *buying*, illustrating the additional function “new habit”

He died on the way down while we were carrying him to pit bottom. Where that was
 : okay. Now when you went in you were carrying out a plan with you very close behi
 music while you 're eating while you 're eating. and then winds it up when you Ye
 o four. So there they are while you 're eating that another quarter each. And wh
 Mm mm. Dave. Do n't sing while you 're eating. They 're very now are n't they? :

Figure 96. Extracts from BNC_spoken concordances of *carrying* and *eating*, illustrating the additional function “framing”

Verbs and framing. Although “framing” was identified to be a very infrequent function in our collection of 9,468 progressive form tokens (see the discussion in 4.4.5), a comparatively large number of semantically rather diverse verbs can express this function. The 82 framing tokens in our BNC/BoE concordances are distributed across 42 verb types. There are hence no particularly significant lexical relations worth highlighting, though some verbs show percentages which are somewhat above average (e.g. *carrying*, *eating*, and *sorting*).

In contrast to most of the other additional functions, we did not find any important typicalities concerning “framing” and verb relations. A pattern, however, was found in the immediate lexical context of “framing” progressives. The large majority of forms collocates with the adverbials “while” or “when” and shows an above-average share of “you” in subject position. This co-selection pattern is illustrated in Figure 96.

To sum up our findings on verbs and additional functions, we can say that, except for “framing”, all functions show some typical relations with certain sets of verbs. Our

systematic functional analysis for each verb form (see Table 30 for results) made clear that, as was the case with context features and central functions, we cannot really trust average values that are determined for a larger group of progressives, as verbs show an often very different individual behaviour. The good news is that this behaviour is not entirely chaotic but that certain regularities can be determined. A number of important co-selection patterns and typical collocations have been mentioned throughout Chapter 4 (especially in 4.5). Some central findings on lexical-grammatical relations will be summarised in the following section.

4.5.12 Summary of the findings [spoken English – verbs and progressives]

The aim of the present subsection (4.5) has been to analyse the relationship between progressives and individual verb forms and to tackle the question “How lexical is grammar?”, i.e. “How strong are the connections between a grammatical construction and its lexical items?”

Our analyses included several types of co-selection in progressives, for instance that of verbs and progressive verb forms, verbs and prepositions, verbs and negation, or verbs and functions of the progressive. The analyses highlighted a number of typical co-occurrence patterns of particular verbs and particular context and function features, and revealed some significant distributional differences between the 99 included verb forms, not only with respect to lexical collocations but also concerning progressive functions and time reference. For example, we found that verbs differ considerably when it comes to frequencies of use in different progressive tense forms such as the past progressive (cf. Table 21 in 4.5.2). The shares of PastProgs range from below 10 per cent to over 40 per cent. Also, there are many significant lexically-determined deviations from the average share of progressive negation. Progressives with verbs like *bothering*, *paying*, or *suggesting* are frequently negated, whereas other forms (e.g. *meeting*, *walking*, or *winning*) exclusively occur in affirmative contexts. The distribution of progressive functions across verb forms (cf. 4.5.10 and 4.5.11) brought similar results which indicated that it may be unwise to treat all verbs as one and to make general statements about functions of “the progressive”. Functions such as “gradual change and development” were found to be highly verb-specific and restricted to a limited group of progressive types.

The examined phenomena certainly show that, with respect to progressives, a lot is lexically determined. What we found out about the relationship between the progressive and individual verbs implies an apparent need to question the existence of a purely grammatical progressive in favour of a lexical-grammatical one. Our findings demonstrate that it is difficult, maybe even impossible, to treat the progressive as a grammatical construction independent of lexis. I have provided ample evidence for the interrelatedness of lexis and grammar and observed that individual verbs show clear preferences for particular contexts and functions and deviate in their behaviour from the specified “normal” distributions.

The results of our corpus-driven analysis strongly support the argumentation by Sinclair, Hunston, Francis, and others in favour of a lexical grammar. As Francis (1993: 142) rightly states,

[...] syntax is driven by lexis: lexis is communicatively prior. As communicators we do not proceed by selecting structures and independently choosing lexis to slot into them. Instead, we have concepts to convey and communicative choices to make which require central lexical items, and these choices find themselves syntactic structures in which they can be said comfortably and grammatically.

We could see that some verbs indeed choose particular progressive structures while others prefer different ones and that, on the other hand, not every verb can be found in every construction. This second observation nicely confirms what Tognini-Bonelli (2001: 33) said about corpus-driven grammar, namely that it “shows very clearly that any grammatical structure restricts the lexis that occurs in it”. I would hence argue that, if we aim at developing an accurate description of language in use, or “used language” in Brazil’s (1995: 24) terms, we need to treat grammar and lexis together, or, as Aarts (2000: 27) puts it, that we “should allow an integrated description of syntactic, lexical and discourse features”.

To answer my initial question, I would say that grammar is *very* lexical, too lexical in fact to be treated as a separate component of language and make general statements about it without paying attention to lexical diversity. In Section 5.7 I will investigate how lexical the grammar is in German EFL coursebooks. I will analyse whether the lexical-grammatical patterns and co-selection phenomena of spoken BrNSE are also found in so-called “school” English.

CHAPTER 5

Progressive teaching (?)

Progressives in the German EFL classroom

The present chapter constitutes the second part of my empirical analysis of progressive forms and their functions and contexts. Having explored the use of progressives in spoken BrNSE in the previous chapter, it will now be interesting to find out how the topic is presented in foreign language teaching and to see what teaching materials say.

In this pedagogical part of the investigation I will focus on English language teaching in Germany and examine the treatment of the progressive in EFL coursebooks and learner's grammars that are widely used in German secondary schools. Of course one could argue that language teaching does not exclusively consist of coursebook work but ideally comprises a range of different pupil-centred activities, based on various types of materials. However, in reality, and teachers will confirm that, the first few years of English language instruction in German schools are very much determined by a particular coursebook series. Most teachers closely stick to the contents of the textbooks and to the grammatical progression suggested there. The language samples in the books make up a considerable part of what I have called "school English" elsewhere (cf. Römer 2004b and Forthcoming). This centrality of the coursebook in ELT justifies the selection of this type of material as the basis of my investigations. Mindt (1996:232), commenting on the similarities between different textbook series, notes that "[t]here is obviously a kind of school English which does not seem to exist outside the foreign language classroom". This is in accordance with my observation that in coursebooks "we tend to find [...] a simplified, non-authentic kind of English" and that "[p]upils are mainly presented with invented sentences, sentences which probably have not occurred in any natural speech situation before (and which probably never will)" (Römer 2004b:153).

The steps taken in the analysis of progressives in textbook English run in parallel to those applied to spoken BrNSE corpus data. I will look at contexts and functions of progressive forms, and examine a range of lexical-grammatical phenomena. Hence I will ensure direct comparability with the results obtained in Chapter 4 in the analysis of progressives in spoken English. In addition to contextual, functional, and lexical-grammatical features, the grammatical progression in the coursebooks will be considered. It will be interesting to see at what stage in the course what information about progressives is presented to the learner. Throughout the different analytic stages, comparisons will be made between the two selected textbook series, *Green Line New*

and *English G 2000 A*. As both series should follow the same pedagogical principles, the differences between them should not be significant. My analyses will show whether this hypothesis can be confirmed or not.

Before turning to the examination of progressives in EFL textbooks though, I will first briefly discuss what lies behind the “learning problem ‘progressive’”, and then deal with issues of teaching materials selection, textbook corpus compilation, and data collection and processing.

5.1 Learning problem “progressive”

Earlier on I hinted at the existing problems related to learning and teaching the progressive. I referred to Williams (2002: 18) who notes that this language feature “constitutes one of the most basic and ubiquitous problems facing language teachers.” More significant than for language teachers, however, are the progressive-related problems for language learners.

Bald, Carstensen and Hellinger (1972: 14) describe the use of the progressive vs. the simple form as an insurmountable obstacle for learners and call for more linguistically founded materials in the EFL classroom. The extreme difficulty learners, even on a comparatively high level of proficiency, often have in appropriately using the progressive is also highlighted by Berry (2001: 1), van Ek (1969: 580), Schneider (1980: 201, 211), and Zydatiś (1976a, 1976b, 1977). Zydatiś (1976b: 352) states that this form “is certainly one of the elements within the English language whose syntax and semantics have remained rather elusive concepts for most learners of English as a foreign language”. In his early learner corpus analysis of spoken and written German learner performance data the author found that in every other occurrence the progressive form was used inappropriately (cf. Zydatiś 1976a: 2). Examples (125) to (128) provide illustrative evidence for existing problems concerning the appropriate use of progressives. They are taken from supervised in-class essays written by, and from classroom discussions among intermediate and advanced German learners of English.⁷⁵

- (125) We saw the Houses of Parliament and we saw Big Ben. Most people *are thinking* [think] that the tower’s name is Big Ben but Big Ben is only the name of the bell.
- (126) (question addressed to a homeless person) What *are you doing* [do you do] every day?
- (127) I’m *thinking* [think] that George W. Bush is a representative of the concept of Manifest Destiny.
- (128) Two years later Grace *expected* [was expecting] our first son, Philip.

Johansson and Stavestrand (1987) make very similar observations based on Norwegian learner data and report on problems Norwegian learners have with the use of the pro-

gressive. As a reason for the inadequate use and the common overuse of this form, they refer to the missing direct counterpart in the learners's native language and state that "[t]he Progressive aspect is a new category for both German and Scandinavian learners of English" (Johansson & Stavestrand 1987: 144). Several other linguists comment on the same reason, i.e. the inexistence of an equivalent grammatical construction in the L1 of the learners (cf. Klein 1995: 140; Königs 2000: 335; Markus 1977: 115; Zydaitis 1976b: 352; cf. also Lenko-Szymanska 2004 on the problems Polish learners have with English progressive forms).

A detailed study of German progressive equivalents is provided by Krause (2002). The author discusses the use of a couple of constructions in German that have a comparable function to the progressive in English, such as *am/beim/im V sein*, (*gerade dabei sein*, *X zu V*, or the so-called absentive, e.g. *einkaufen sein*. To date however, none of these constructions has been entirely grammaticalized, and it is arguable whether the *am*-progressive as in *ich bin einen Brief am schreiben* (*I am writing a letter*), which is typical of Rhineland dialects, should be considered standard German or rather vernacular (for a recent discussion on that topic, see Rödel 2003). Besides, these constructions fulfil partly different functions and are not necessarily translation equivalents of English progressives. The absentive in a phrase like *Sie ist einkaufen*, for instance, uttered in response to *Wo ist Mama?* (*Where is Mum?*), would translate as *she went shopping*, not as *she is shopping*.

The inexistence of a grammaticalized equivalent in German certainly serves as an explanation of the learning problem "progressive"; however, it might not be the only one. We could claim that another reason for learner problems lies in inadequate descriptions of language phenomena in teaching materials (cf. Römer Forthcoming). Maybe what learners get in their coursebooks and grammars is not the most appropriate account of progressives in English. Maybe learners would find it less difficult to handle the progressive in communicative situations if it was presented in the same way as it is used by native speakers, i.e. in its most typical contexts and functions. In a recent forward-looking article on corpus evidence and language teaching, Sinclair (2004c: 273) discusses some central problematic features of language description (ambiguity, variation, terminology, and incompleteness) and asks the question "Are they inherent in the language or do they arise in the description?" This question will also be addressed in the present study, though only with reference to one feature, the progressive. My corpus-driven analysis of EFL textbook language (compared to spoken BrNSE) will help to find out how appropriate the descriptions are which teaching materials offer the learner and whether a reason for learning problems may lie in faulty language descriptions.

5.2 Selection of teaching materials

I noted above that the first step in any corpus linguistic research project must be the selection of appropriate data collections that can serve as the basis of the analysis. The considerations behind the selection of the BNC_spoken and BoE_brspok corpora have been discussed earlier (cf. 4.1). In the selection of EFL teaching materials and the compilation of a pedagogical database I then had to consider which materials are most representative of the type of English used in the majority of German EFL classrooms.

Responding to this consideration, the main criterion which determined the selection process was the distribution of available EFL teaching materials across Germany. I decided to use those books in my analysis that are most widely used in German secondary schools. Also, I wanted to make sure that I only examine recently published materials. These considerations lead to the selection of two coursebook series (and their accompanying grammars) by the two leading publishing companies on the German EFL market: *Learning English Green Line New* (henceforth GLN; Ashford et al. 1996, 1997, 1998, 1999, 2000; Aston et al. 1995) published by Klett Verlag (Stuttgart) and *English G 2000 A* (henceforth EG 2000; Schwarz 1997, 1998a, 1999a, 1999b, 2001, 2002) published by Cornelsen Verlag (Berlin).

GLN and EG 2000 are two best-selling series with a very wide distribution across the fifteen German federal states. At the time of materials selection for this study, GLN and EG 2000 had just been published as the most up-to-date EFL coursebook series available. Each series consists of six coursebook volumes which are used from grades 5 to 10. The pupils who use these books are about ten years old at the beginning of the course (in grade 5, immediately after having finished primary school) and about sixteen when they finish the course (in grade 10). At the end of grade 10, pupils either leave school or stay on for another three years before they can take their A-levels. The altogether twelve selected volumes of GLN and EG 2000 can be said to provide a representative sample of the coursebooks used in German schools at the specified level of English language instruction.

Also included in the analysis are two learner's grammars which accompany the textbooks. The first one is the *Learning English Grundgrammatik* (Ungerer et al. 2001) by Klett Verlag, the second one the *Cornelsen English Grammar* (English edition, Fleischhack & Schwarz 2001). In my analysis of progressive contexts, functions, lexical-grammatical relations, and the progression in the coursebooks, I will make repeated reference to both grammars and to the so-called "Grammatische Beihefte" of GLN and EG 2000. These Grammatische Beihefte are thin grammar booklets, each of which summarizes the grammar sections of two coursebook volumes. The systematic corpus-driven investigation, however, does not cover the contents of the learner's grammars or the Grammatische Beihefte but is solely based on the twelve GLN and EG 2000 volumes. The following section will describe the design and compilation of a small electronic corpus, GEFL TC, which consists of a large number of texts from all volumes of *Green Line New* and *English G 2000 A*.

5.3 The German English as a Foreign Language Textbook Corpus (GEFL TC) – a collection of EFL textbook language

5.3.1 Corpus design and composition

The German English as a Foreign Language Textbook Corpus (henceforth GEFL TC) was compiled to enable the same analytic steps for the analysis of “school” English as were carried out for spoken BrNSE.⁷⁶ There was a need for such a corpus of German EFL textbook language, as no other ready-made computerised collections of this text type were available⁷⁷ and as a manual examination of the GLN and EG 2000 volumes, in analogy to the BNC_spoken and BoE_brspok analyses, would not have been feasible. When it comes to counting, sorting, and calculating, the computer is much faster and much more reliable and infallible than any human brain.⁷⁸ In order to optimise and systematise my empirical analysis, I thus decided to transform the printed teaching materials into electronic format and to treat this computerised text collection as a corpus. This corpus can be classified as a “pedagogic corpus”, defined by Hunston (2002a: 16) as a collection of data which “can consist of all the course books, readers etc. a learner has used”.⁷⁹

The most important design criterion of GEFL TC relates to the choice of text type. To allow for comparisons of the retrieved results from this “school” English corpus with the findings based on spoken English, only such and all such texts from the EFL coursebooks were selected which represent spoken language. Included in GEFL TC are, for instance, dialogues, interviews, speech bubbles, and those narrative texts that consist mainly of dialogue (cf. Römer 2004b). All sorts of exclusively written material, such as narratives, letters, or excerpts from novels, as well as exercises or grammar boxes are not made part of the corpus.

I decided not to take samples from the texts but use full texts and to include all texts from the twelve coursebook volumes which fit my text-type criterion. Altogether, this corpus of spoken-type texts from GLN and EG 2000 has a size of a bit more than 100,000 words, i.e. tokens. Judged by today’s standards, this is a relatively unimpressive size. We have to keep in mind, though, that GEFL TC had to be compiled by just one person in a relatively short period of time and in a quite time-consuming process (cf. 5.3.2). The inclusion of texts from a wider range of textbooks was hence not possible. Besides, we are here dealing with a rather specialised corpus which is supposed to represent the type of language used in German secondary school level EFL course materials. Such specialised corpora are usually much smaller than general or reference corpora. In his preface to a collection of papers on small corpus studies (Ghadessy, Henry & Roseberry 2001), Sinclair (2001: xi) describes the difference between corpora as one of method, rather than just of size. He refers to the editors’s introduction to the book in which they express that

[a] small corpus is seen as a body of relevant and reliable evidence, and is either small enough to be analysed manually, or is processed by the computer in

a preliminary fashion, [...], thereafter the evidence is interpreted by the scholar directly. There is no need to collect the quantities of data needed in order to delay the direct participation of the human being.

This methodological difference between small and large corpora is hence one between “early human intervention (EHI)” and “delayed human intervention (DHI)” (Sinclair 2001:xi). In approaching GEFL TC, I take an EHI approach, which combines initial computerised procedures and subsequent manual analytic steps.

The composition of my small and specialised corpus of textbook English, GEFL TC, is illustrated in Figure 97. As we can see, the corpus consists of two subcorpora of almost identical size and internal structure (see statistics boxes): “English G 2000” and “Green Line New”. The subcorpora contain spoken-type texts from volumes 1 to 6 of the respective coursebook series. Their similarity in composition enables direct comparisons between the two textbook subcorpora. I will now briefly describe the process in which GEFL TC was compiled.

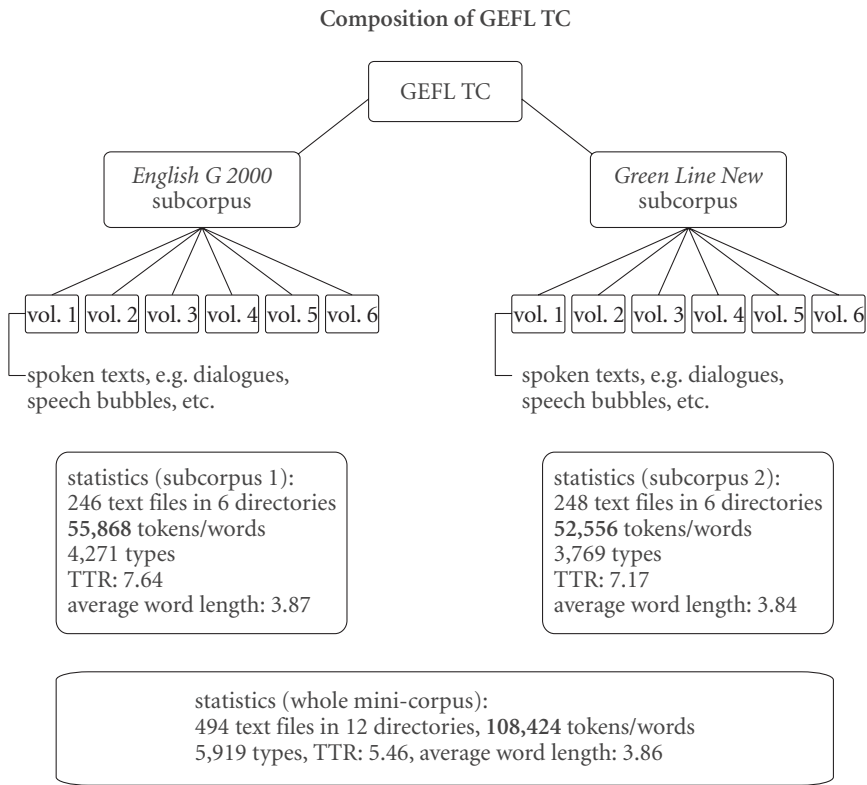


Figure 97. The composition of GEFL TC

5.3.2 Corpus compilation

GEFL TC was compiled in four steps. First of all, those pages from the twelve course-book volumes of the GLN and EG 2000 series were selected which included spoken-type textual material. Each selected textbook page was then digitised by means of a scanner. The result of the scanning were image files in tif-format.

In a third step these files were processed with *TextBridge Pro* (version 8.0), an optical character recognition (OCR) software. The highlighting tool in *TextBridge Pro* served to select all parts of the scanned pages which were to be included in the corpus. Figure 98 gives a screenshot of the OCR software and shows that from this page only the two short dialogues have been selected (see grey shadings). Pictures and narrative passages on the same page could thus be excluded from the corpus.

The final step in the compilation of GEFL TC involved a conversion of the highlighted textual material in the image files into txt-file format. It was then possible to analyse the collected text files (one for each scanned textbook page) with the help of a concordance program (cf. 5.4.2). Separate queries in GLN and EG 2000 texts could easily be carried out because the text files had been saved in two corpus subfolders.

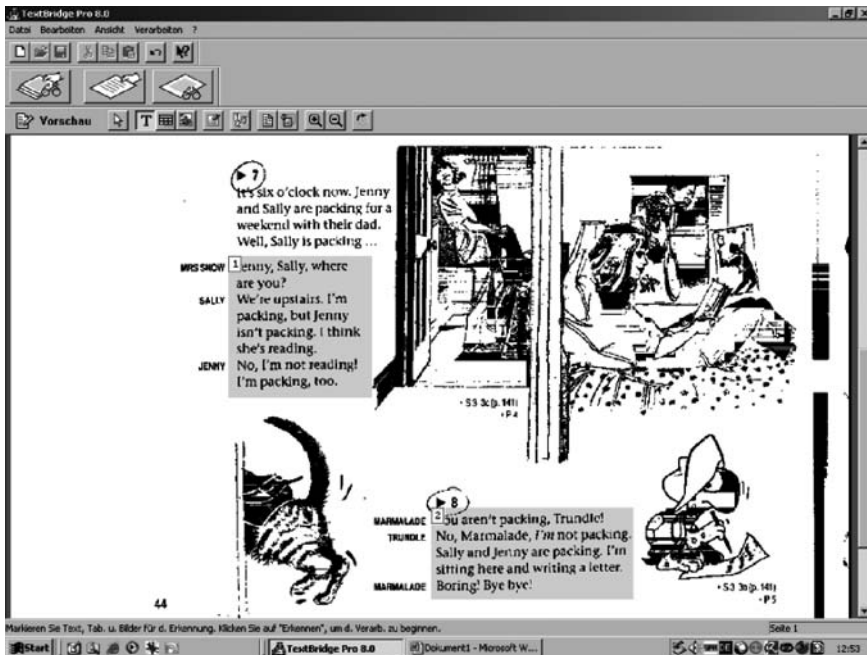


Figure 98. Processing of a textbook page image with the OCR software *TextBridge Pro*

5.4 The empirical method: GEFL TC data collection, processing, and evaluation

In analogy to Section 4.2, which dealt with the actual corpus-driven work based on the BNC_spoken and BoE_brspok corpora, the next short sections discuss the steps that were taken in the GEFL TC analysis.

5.4.1 Verbs under analysis

The aim of the textbook corpus analysis was to retrieve findings on the use of progressives in “school” English which could be compared with the results obtained in the analysis of BNC and BoE data as described in Chapter 4.

Starting point for the collection of corpus data from GEFL TC thus had to be the same list of 100 -ing forms of high-frequency English verbs which was used in the examination of progressives in spoken BrNSE. The 100 selected -ing forms have been listed in Table 4 above but are displayed once again in Table 31 for the convenience of the reader.

Table 31. 100 -ing forms under analysis (in alphabetical order)

1. accepting	26. finding	51. meaning	76. spending
2. adding	27. finishing	52. meeting	77. standing
3. agreeing	28. following	53. moving	78. starting
4. asking	29. forgetting	54. needing	79. staying
5. becoming	30. getting	55. paying	80. stopping
6. being	31. giving	56. picking	81. suggesting
7. believing	32. going	57. playing	82. supporting
8. betting	33. happening	58. providing	83. supposing
9. bothering	34. having	59. pulling	84. taking
10. bringing	35. hearing	60. putting	85. talking
11. buying	36. helping	61. reading	86. telling
12. calling	37. holding	62. remembering	87. thinking
13. carrying	38. hoping	63. ringing	88. trying
14. changing	39. imagining	64. running	89. turning
15. checking	40. keeping	65. saving	90. understanding
16. coming	41. knowing	66. saying	91. using
17. costing	42. learning	67. seeing	92. walking
18. cutting	43. leaving	68. seeming	93. wanting
19. dealing	44. letting	69. selling	94. watching
20. doing	45. liking	70. sending	95. wearing
21. drawing	46. listening	71. setting	96. winning
22. eating	47. living	72. showing	97. wondering
23. expecting	48. looking	73. sitting	98. working
24. explaining	49. making	74. sorting	99. worrying
25. feeling	50. mattering	75. speaking	100. writing

5.4.2 Data collection: Querying GEFL TC with *WordSmith Tools*

What was noted above with reference to large corpora such as the British National Corpus or The Bank of English also applies to a relatively small text collection like GEFL TC: Computerised corpora would be rather useless tools if the researcher did not have the appropriate corpus-analytic program at hand. *WordSmith Tools* (WST), the same software package that was used in the collection of BNC_spoken data, was also the program of choice to carry out searches in the EFL textbook corpus.

In order to retrieve concordances of progressive forms from the GEFL TC subcorpora, “Green Line New” (GLN) and “English G 2000 A” (EG 2000), I used the WST Concord tool (cf. Section 4.2.2). Concordance queries were carried out for each of the 100 -ing forms listed in Table 31, first in GLN, then in EG 2000. For each type I downloaded all occurring form tokens, as the maximum number of occurrences was only 104 in GLN and 108 in EG 2000 (in both cases for *going*), and as all other -ing forms showed much smaller token numbers (between 1 and 41). No instances at all were found for 33 of the 100 types in EG 2000 and for 37 types in GLN. These forms which retrieved no results are listed in Table 32. The downloaded concordance sets of the remaining 67 and 63 verb forms respectively were sorted to the left and saved as text files.

Table 32. -ing forms of the list of 100 which do not occur in GEFL TC

no occurrences in the GLN subcorpus				
1. accepting	9. checking	17. knowing	25. pulling	33. suggesting
2. adding	10. costing	18. liking	26. remembering	34. supporting
3. agreeing	11. dealing	19. mattering	27. saving	35. supposing
4. becoming	12. expecting	20. meaning	28. seeming	36. understanding
5. believing	13. finishing	21. moving	29. sending	37. wanting
6. betting	14. giving	22. paying	30. setting	
7. bothering	15. imagining	23. picking	31. sorting	
8. changing	16. keeping	24. providing	32. spending	
no occurrences in the EG 2000 subcorpus				
1. accepting	8. buying	15. forgetting	22. picking	29. suggesting
2. adding	9. costing	16. hearing	23. providing	30. supporting
3. agreeing	10. cutting	17. knowing	24. seeming	31. supposing
4. becoming	11. dealing	18. liking	25. sending	32. understanding
5. believing	12. drawing	19. mattering	26. setting	33. wanting
6. betting	13. explaining	20. needing	27. sorting	
7. bothering	14. finishing	21. paying	28. stopping	

5.4.3 Data filtering

Following the procedure described in 4.2.3, the saved concordances of -ing forms were filtered manually. I deleted all those instances in which the searchword (e.g. *living*) was not part of a progressive, but, for instance, a noun as in examples (129) and (130).

(129) “Ooh, look,” David says. “Prizes for the healthy *living* competition!” (GLN)

(130) Hey, Jenny, I’m in the *living*-room. (GLN)

In this filtering process, the number of GEFL TC concordance lines was reduced from 1,044 (475 GLN and 569 EG 2000) to 702 (306 GLN and 395 EG 2000). From the original 130 data files only 104 remained. Unlike in the case of the BNC/BoE examples, the saved 200 characters per concordance line always provided enough context to make decisions about the “progressive” status of an example. After the filtering the reduced concordances were again saved as txt-files. These files could then be imported to the *Access* corpus database, the construction of which has already been described in Section 4.2.4 above.

5.4.4 Data processing and encoding: The addition of GEFL TC concordance lines to the Access database

In the next step, each of the 104 GLN and EG 2000 concordance files was imported separately into the already existing *Access* database which stored the 9,468 progressives from BNC_spoken and BoE_brspok. The database then had 10,171 entries altogether, 702 from GEFL TC texts and the rest from the spoken BrNSE corpora.

For the annotation of the 702 GEFL TC progressives, the same function and context variables were used that had been defined to encode BNC and BoE data (cf. Tables 6 and 7 in 4.2.4). Once all progressive concordance lines had been annotated according to variables such as time reference, repeatedness, time adverbial, negation, etc., it was possible to retrieve the respective values for the GLN and EG 2000 datasets.

5.4.5 Data evaluation

The evaluation of GEFL TC corpus data was carried out in the same way as that of BNC_spoken and BoE_brspok data (cf. 4.2.5).

From the annotated database, figures were retrieved mainly with the help of the *Access* filter-by-form function. This function renders only those concordance examples from the set of 10,171 that match certain specified criteria, such as searchword: “going”, corpus: “Green Line New”, form of TO BE: “are”, and subject: “we”, as e.g. *So we are going to go out in the boat tomorrow*. Figure 99 displays the seven concordance lines from GEFL TC which meet the filter just defined.

Important also in the data evaluation was the *Access* sorting option (in table view only), which enabled immediate access to figures and patterns in larger datasets that

concordance line	searchword	corpus	subject/p	BE form	time advert
n and asked. "No, sorry," the man in the pet shop answered. "Oh, no!" said Robert. "Now what are we going to do?" Then Becky had an	going	Green Line	we	are	now
, too - all over the front seats. "Oh no, look! He left the sunroof open," said Sarah. "What are we going to do now?" "Well, we'll have to tell	going	Green Line	we	are	now
climb Ben Nevis tomorrow. And we'll go to Loch Ness the day after. Everybody happy now? What are we going to do tomorrow? What	going	Green Line	we	are	tomorrow
museum now?" "Yes," Katrin said. "I'm not interested in playing video games any more." "What are we going to say to Mrs Grant?"	going	Green Line	we	are	
Mr and Mrs MacTa vish have got a boat for the guests. Dad can row (well, he says he can!). So we are going to go out in the boat tomorrow!	going	Green Line	we	are	tomorrow
nging?" he asked. "You aren't singing," said Paula. "Can you take the tin?" said Mrs Davis. "We are going to the streets behind the	going	Green Line	we	are	
ow, Becky: Hey, David. Are you going to stay in Nottingham this summer? David: Yes, I am. We aren't going to have a holiday this year.	going	Green Line	we	are	this year

Figure 99. The application of an Access “filter by form” in table view (GLN data)

resulted from some of the less complex filtering processes. The numbers retrieved by means of these query strategies serve as a basis of the discussion of progressive use in “school” English in Sections 5.5 to 5.7.

5.5 The use of progressives in “school” English (I) – contexts

In the discussion of the contexts of progressives in spoken English in Section 4.3, I said that not only is *context* at the centre of all sorts of corpus-driven studies, but also that it is advisable to pay attention to context phenomena, such as collocation and colligation, in language pedagogy. Reference was made to one of Sinclair’s (1997: 34) pedagogical precepts, “[i]nspect contexts”, which is, I would argue, largely neglected in language teaching.

Having stressed the pedagogical importance of a contextual approach to language features and having examined the contexts of a large number of progressives from spoken BrNSE corpora in some detail, I will now turn to analysing the contexts of progressive verb forms in “school” English, i.e. in the datasets retrieved from the two GEFL TC subsections. My analysis closely follows the model of the BNC_spoken and BoE_brspok data examination described in Section 4.3 above. In the following subsections, I will concentrate on the same context features as have been dealt with in the

analysis of real spoken English, and later on compare the results of both parts of the investigation (cf. Section 6.1).

5.5.1 Distribution of different tense forms

In the contextual analysis we will first of all look at the shares of different types of progressive forms in GEFL TC. It will be seen how tokens of the past progressive (PastProg), the present progressive (PresProg), the present perfect progressive (Pres-PerfProg), and the past perfect progressive (PastPerfProg) are distributed in the GLN and EG 2000 datasets.

To determine this tense form distribution, the progressives in the database were filtered for corpus (GLN, EG 2000) and form of TO BE. This means that a new filter had to be specified for each occurring TO BE form. The thus retrieved absolute numbers of occurrence and their relative frequencies are displayed in Table 33. If we add up the values of those TO BE forms which belong to the same tense form, we arrive at the distribution that is visualised in Figure 100.

As Figure 100 shows, the results based on the two GEFL TC subcorpora are roughly comparable. By far most frequent in GLN (with 70.03%) and in EG 2000 (with 75.95%) are present progressives. If we look at the frequencies of individual forms of TO BE within the PresProg (cf. Figures 101 and 102), we see that the most common patterns in both GEFL TC subcorpora are *are V-ing* and *'m V-ing*, although the actual values for GLN and EG 2000 differ to a certain extent (34.88% vs. 25.67% and 19.53% vs. 24.67%). The shares of the PresProg patterns *'s V-ing* and *is V-ing* are also rather

Table 33. The co-selection of progressives and forms of TO BE (GEFL TC)

Form of TO BE	GLN (307 examples) absolute and relative frequencies	EG 2000 (395 examples) absolute and relative frequencies
<i>am</i>	4 (1.30%)	5 (1.27%)
<i>'m</i>	42 (13.68%)	74 (18.73%)
<i>are</i>	75 (24.43%)	77 (19.49%)
<i>'re</i>	36 (11.73%)	57 (14.43%)
<i>is</i>	41 (13.36%)	31 (7.85%)
<i>'s</i>	17 (5.54%)	56 (14.18%)
<i>was</i>	60 (19.54%)	48 (12.15%)
<i>were</i>	20 (6.51%)	25 (6.33%)
<i>has been</i>	0	1 (0.25%)
<i>'s been</i>	0	3 (0.76%)
<i>have been</i>	4 (1.30%)	2 (0.51%)
<i>'ve been</i>	6 (1.95%)	8 (2.03%)
<i>had been</i>	1 (0.33%)	3 (0.76%)
<i>'d been</i>	0	0
none (empty box)	1 (0.33%)	5 (1.27%)
<i>ai</i> (in <i>ain't</i>)	0	0

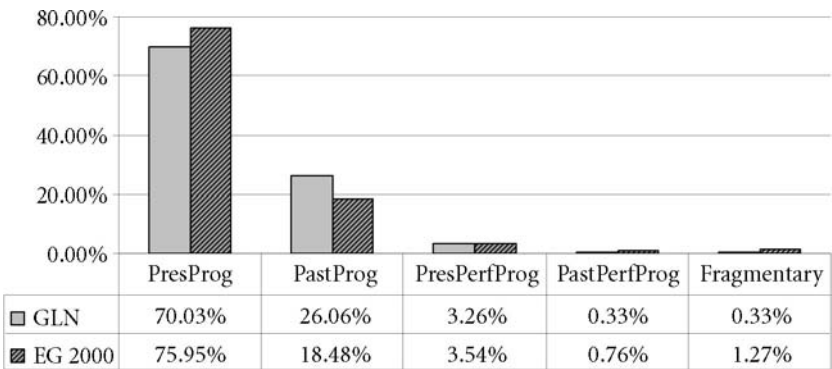


Figure 100. The distribution of different tense forms in GLN and EG 2000

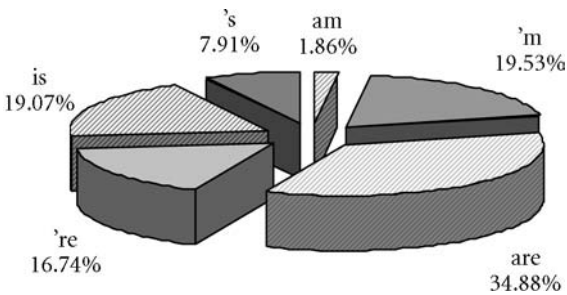


Figure 101. The distribution of forms of TO BE within the PresProg (GLN)

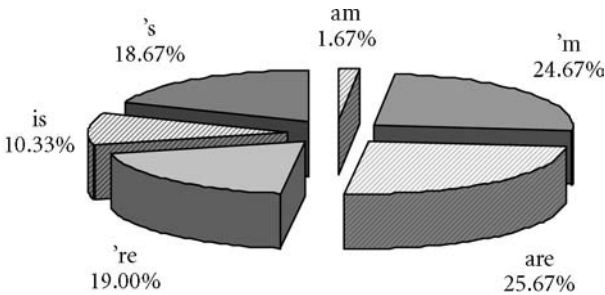


Figure 102. The distribution of forms of TO BE within the PresProg (EG 2000)

different in the two textbook corpora (7.19% vs. 18.67% and 19.07% vs. 10.33%). We can thus say that while the difference between the overall shares of PresProgs in GLN and EG 2000 is not significant (chi-square tested), there is quite some variation between the two corpora when it comes to the respective lexical realisations of this structure.

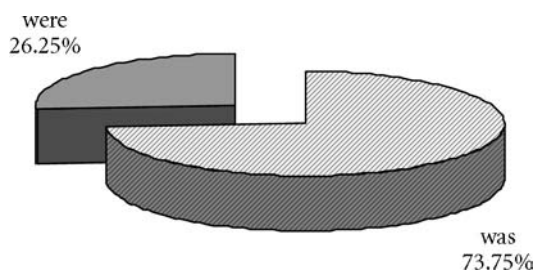


Figure 103. The distribution of forms of TO BE within the PastProg (GLN)

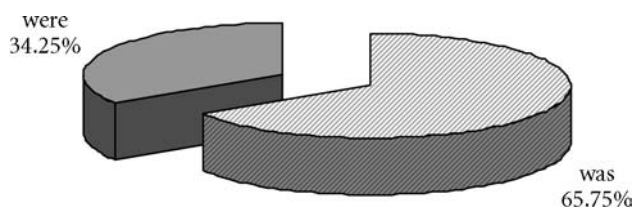


Figure 104. The distribution of forms of TO BE within the PastProg (EG 2000)

Past progressive forms are much less frequent in the coursebook datasets than PresProgs. They make up 26.06% in GLN and 18.48% in EG 2000, which again is not a statistically significant difference (at .05 level). As Figures 103 and 104 illustrate, the collocation *was V-ing* is clearly more common in both corpora than *were V-ing*. The actual frequencies, however, differ to some degree (26.25% vs. 34.25% and 73.75% vs. 65.75%). For the PresPerfProg, the PastPerfProg, and for forms that have been labelled “fragmentary” (e.g. “You keeping that as well?” EG 2000), the shares are very low and very similar in GLN and EG 2000 (cf. Figure 100). There is only a small number of instances of these types in the coursebooks, and they are mainly realised by the forms *’ve been V-ing*, *have been V-ing*, and *had been V-ing*. The patterns *has been V-ing*, *’s been V-ing*, and *’d been V-ing* do not occur at all in GLN and are very rare in EG 2000. As the absolute numbers are very low, we cannot really talk about significant differences of individual PresPerfProg and PastPerfProg distributions between the two textbook corpora.

5.5.2 Tense form contractions

We saw in the previous section that the shares of contracted PresProg forms are higher in EG 2000 than in GLN. We will now look at short vs. long form realisations of all GEFL TC progressives and briefly discuss whether there is a general systematic difference between the distributions in the two subcorpora. GEFL TC examples of contracted and non-contracted progressives are given in (131) to (134).

- (131) 'We don't have time to be real teenagers. We're either *doing* math and science all the time or getting up early to practice with the swim team. (GLN)
- (132) David: What *are you doing*, Mark? Mark: I'm drawing a birthday picture for Jenny. (GLN)
- (133) Sorry, I can't come to the phone right now. I'm *taking* a bath. So leave a message after the beep. (EG 2000)
- (134) Well, why *am I telling* you all this? Let's go!' (EG 2000)

Figure 105 illustrates the shares of short forms (e.g. *'m V-ing*) and long forms (e.g. *am V-ing*) in GLN and EG 2000. The group-internal distributions of contracted and non-contracted forms are visualised in Figures 106 to 109. As we can clearly see in Figure 105, there are significant differences between the shares of long and short forms of TO BE in the two GEFL TC subcorpora.⁸⁰ With 66.78%, GLN favours non-contracted progressives and shows a comparatively low percentage of contracted forms, whereas in EG 2000 contracted form tokens are more common than long forms. Among the contracted forms, *'m V-ing* has the largest share in both corpora (cf. Figures 106 and 107). Second most frequent is *'re V-ing*, followed by *'s V-ing*.

As was the case with PresProg and PastProg tense form realisations (cf. 5.5.1), we find rather different individual percentages for GLN and EG 2000 concerning some of the contracted forms. The same is true with respect to the different non-contracted progressive types, though to a smaller extent. Figures 108 and 109 show that although the order of frequency of long constructions in both datasets is *are V-ing* → *was V-ing* → *is V-ing* → *were V-ing*, the actual shares diverge. With respect to verb form contractions, it would hence be unwise not to distinguish between GLN and EG 2000 and treat both corpora as one in the later comparison with BNC/BoE-values.

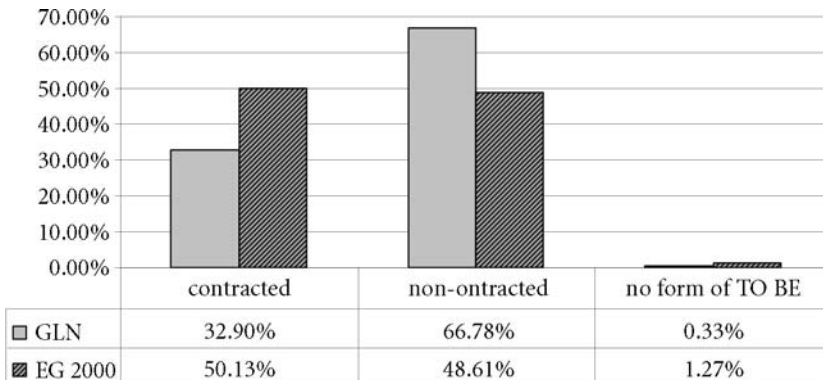


Figure 105. Contracted and non-contracted progressive forms in GLN and EG 2000

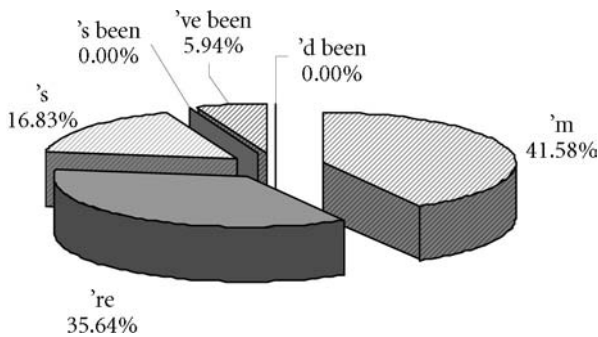


Figure 106. The distribution of different contracted forms (GLN)

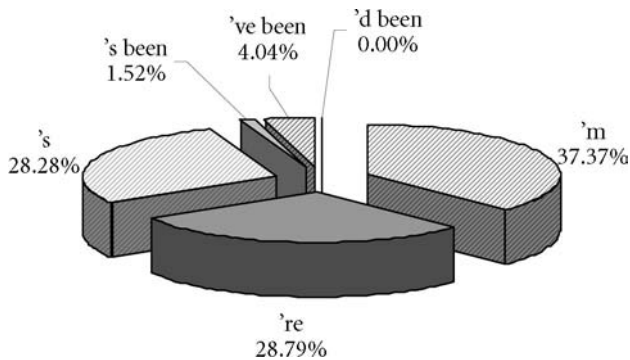


Figure 107. The distribution of different contracted forms (EG 2000)

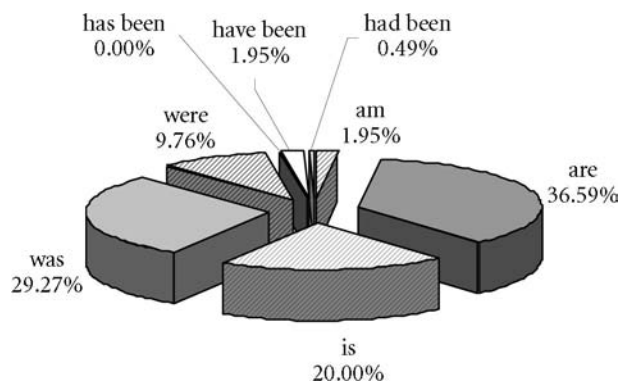


Figure 108. The distribution of different non-contracted forms (GLN)

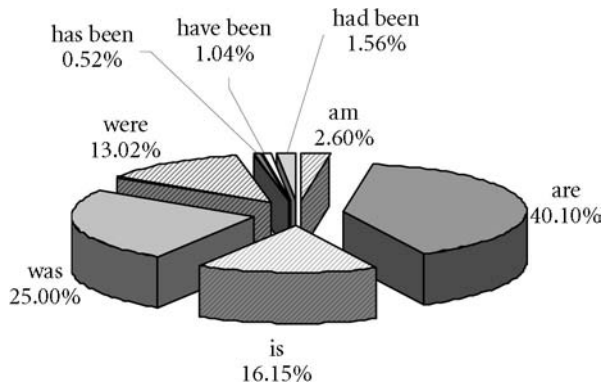


Figure 109. The distribution of different non-contracted forms (EG 2000)

5.5.3 Progressives and subjects

Another important context feature that was analysed in the BNC/BoE data evaluation was the co-selection of progressive forms and their subjects. I will now describe the connections between subjects and progressives in GEFL TC.

By far most frequent in subject position in the GLN and EG 2000 datasets are personal pronouns (cf. Figure 111). With 70.63% and 68.08%, the two coursebook corpora show rather similar shares. Some differences, however, can be found when we consider the relative frequencies of occurrence of the individual pronouns in GLN and EG 2000, as displayed in Figure 110. Most common in both corpora are *I* and *you* as progressive subjects, though for both pronouns the shares are higher in EG 2000 than in GLN (35.13% vs. 25.84%, and 22.94% vs. 20.57%). On the other hand, the GLN dataset shows comparatively higher percentages of *we*, *they*, and *he*. Examples (135) to (138) serve to illustrate some of these common co-selection types. Again, we are not surprised to observe higher figures for *he* than for *she* in both datasets. Apparently, the general male bias in English is mirrored in the teaching materials.

- (135) “We are going to the streets behind the supermarket.” (GLN)
- (136) *They’re* putting all their rubbish from their lunch in the water. (GLN)
- (137) I took part in competitions and *I* was getting on fine. (EG 2000)
- (138) You ... *you’re* really asking for it. (EG 2000)

After personal pronouns, names of people are the second most frequent type of subjects in EFL textbook progressives. Common in GLN are, for instance, *Emma*, *Mark*, and *Mr. Zorzi*, while *Nick*, *Debbie*, and *Jenny* count among the featured characters in EG 2000. Other subjects that occur repeatedly in the two analysed datasets are *the* + noun or noun group, e.g. *the ghost* or *the twins*, and *what* and *who*. All other subject types, such as *people*, or *a* + noun or noun group are very infrequent in both subcorpora of GEFL TC.

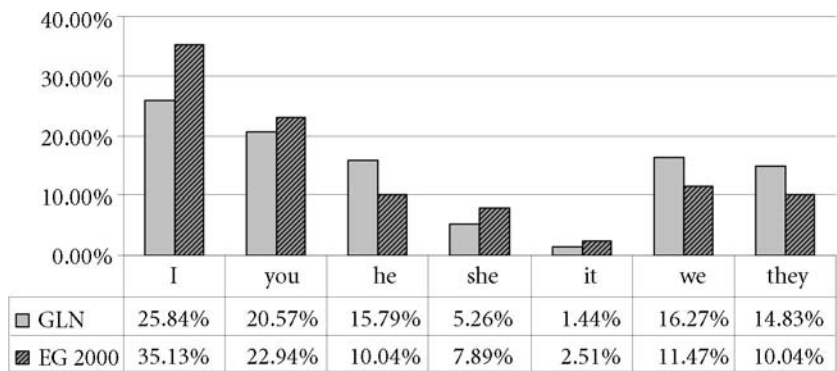


Figure 110. The distribution of personal pronouns as subjects of progressives (GLN and EG 2000)

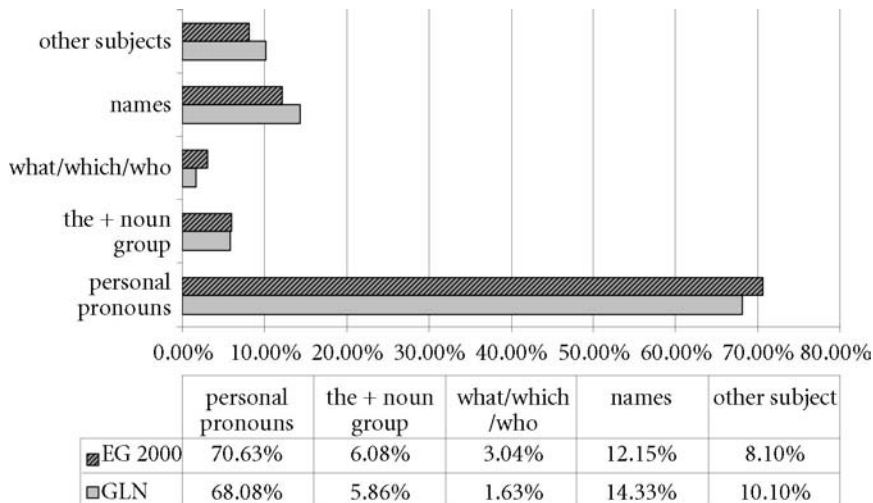


Figure 111. The distribution of subjects in progressive constructions (GLN and EG 2000)

5.5.4 Progressives and objects

It will now be examined whether in the textbook English data there are any typical co-occurrences of progressives and particular items in object position. Figure 25 in Section 4.3.4 illustrated the shares of some commonly used objects in spoken BrNSE. A comparable illustration, based on GLN and EG 2000 data, is provided in Figure 112 below.

We see that there is quite some variation among words and phrases in object position (cf. the large number of “other objects”) and that nouns or noun groups introduced by *the* and *a* (e.g. *the birds*, *a party*) are particularly frequent. Two typical examples are given in (139) and (140).

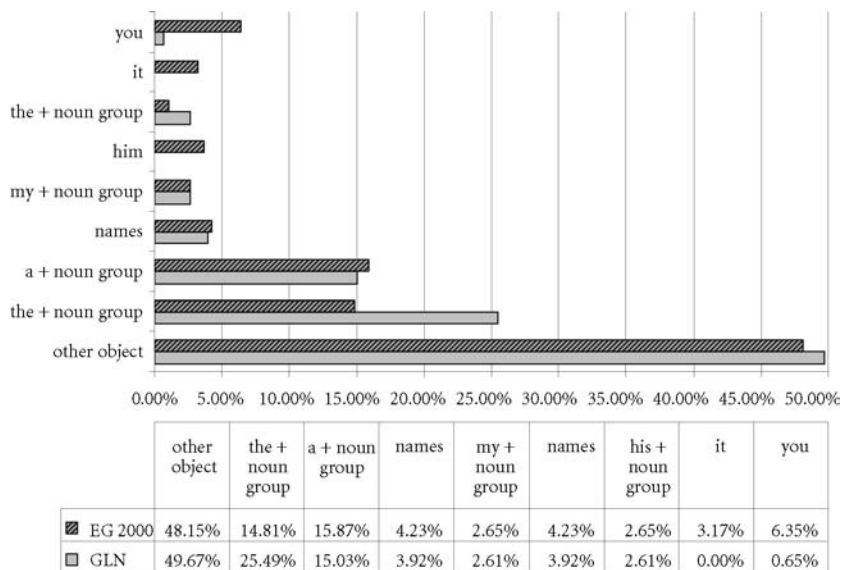


Figure 112. The distribution of objects in progressive constructions (GLN and EG 2000)

- (139) Sarah is buying *the tickets* because she has got money from her parents. (GLN)
 (140) I’m getting *a cramp* in the leg from running. (EG 2000)

Repeatedly used also are personal names, such as *Debbie* or *Aunt Claire*, and some pronouns, sometimes followed by a noun or noun group, as in examples (141) and (142).

- (141) “What’s Simon doing with *his brother*?” (GLN)
 (142) I’m looking for *my head* and I must find it. (EG 2000)

Noteworthy are the different shares of some objects, especially *you*, *it*, *him*, and *the + noun* or noun group, in the two subcorpora (cf. Figure 112). It seems that the lexical differences between the coursebook series are bigger than expected.

5.5.5 Progressives and prepositions

If we now turn our attention to prepositions which postmodify the progressive forms, we note that the shares of progressive concordance lines with prepositions in 1R position, i.e. immediately on the right of the -ing form, are comparable in GLN (30.29%) and EG 2000 (28.61%). The percentages of the most frequently occurring prepositions are given, and graphically illustrated, in Figure 113.

As Figure 113 shows, many of the values for the GLN and EG 2000 datasets differ noticeably from each other. The most common prepositions in GLN progressives are *at*, *to*, *about*, and *in*, whereas in EG 2000 the prepositions *to*, *about*, *on*, and *for* are most frequently used following progressive verb forms (cf. examples (143) to (146)).

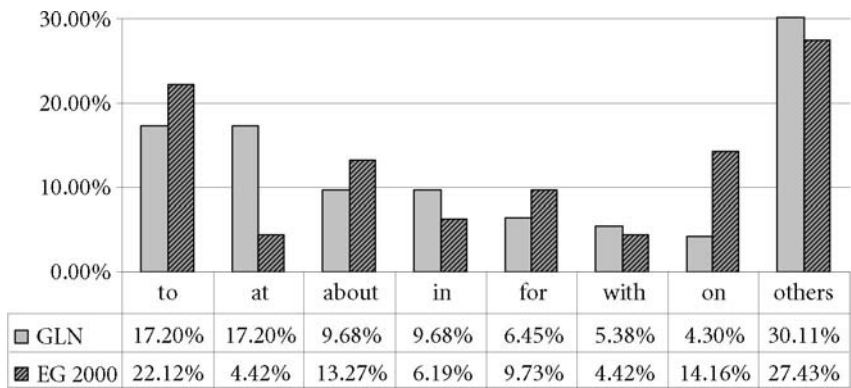


Figure 113. The distribution of prepositions in progressive constructions (GLN and EG 2000)

- (143) “Are you staying *at* Tulloch House?” (GLN)
- (144) The other kids in my class were coming *in* by now. (GLN)
- (145) JOE: The wrong time? What are you talking *about*? (EG 2000)
- (146) She was working *on* the project with the twins. (EG 2000)

Of course, what is true for the occurrence of prepositions in progressives in BNC_spoken and BoE_brspok also applies to GEFL TC data: Prepositions are highly lexically determined items and their frequency of use depends on the use of individual verbs. Therefore, a more detailed analysis of the results displayed in Figure 113 is probably not very sensible, as these results are based on the whole set of different progressive types taken together. In Section 5.7.5 we will look into the connections between prepositions and individual verb forms and see what the most typical preposition-progressive patterns in the coursebooks are.

5.5.6 Progressives and negation

We noted above that progressives in negative contexts are apparently more frequent in speech than in writing and found shares of 8.59% and 7.98% of negated progressives in BNC_spoken and BoE_brspok.

As Figure 114 shows, the values of negation in the two textbook English datasets are significantly lower than in our spoken BrNSE corpora. Of the progressives in GLN only 4.56% are negated, and with 3.80% the share for EG 2000 is even smaller (for GEFL TC examples of negated progressives see (147) and (148)). This finding may hint at a type of language in the spoken-type coursebook texts which is closer to written than spoken English. However, such a register-related statement can of course not be made on the basis of just one distinctive feature. More indicators are needed to make a sound judgement concerning the register attribution of GEFL TC texts. It will also be

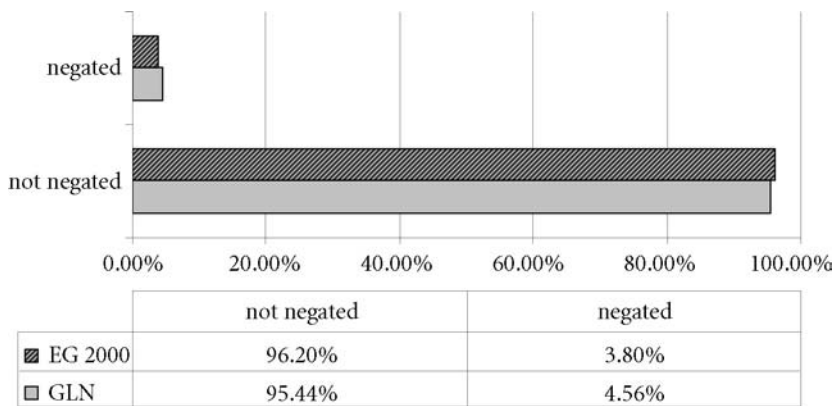


Figure 114. Shares of negated and non-negated examples in the GLN and EG 2000 datasets

interesting to see whether there are any relations between particular verbs and negated progressives. Section 5.7.6 will examine whether the shares of negation are equally low for all verbs or whether some of them show particularly high percentages.

- (147) I'm *not* listening to Radio 1. It's Radio Nottingham. (GLN)
- (148) Jenny, why *aren't* you helping your dad? (EG 2000)

5.5.7 Progressives and other lexical-grammatical phenomena

In Section 4.3.7 above I discussed the co-occurrence of progressive forms and three types of lexical-grammatical constructions in spoken BrNSE: questions, if-clauses, and relative clauses. If we check the same co-selection features in our GEFL TC data, we notice that, of the three constructions, only questions form a significant colligational pattern with progressives (cf. Figure 115). In other words, a comparatively large number of progressive forms from the GLN and EG 2000 subcorpora occur in interrogative contexts, as in examples (149) and (150).

- (149) “What’s Simon *doing* with his brother?” (GLN)
- (150) *Are* you *taking* her to City Hospital? (EG 2000)

If-clauses and relative clauses are, on the other hand, as the values displayed in Figure 115 demonstrate, very infrequent in the progressive datasets. In each subcorpus there are just four instances of progressives in if-clauses, two of which are given below in (151) and (152). In half of the eight examples the progressive form occurs in an embedded (that)-clause. Two of the seven relative clause concordance lines in which progressives feature (three from GLN and four from EG 2000) are displayed in (153) and (154). Due to the very small absolute numbers of occurrence of if-clauses and rel-

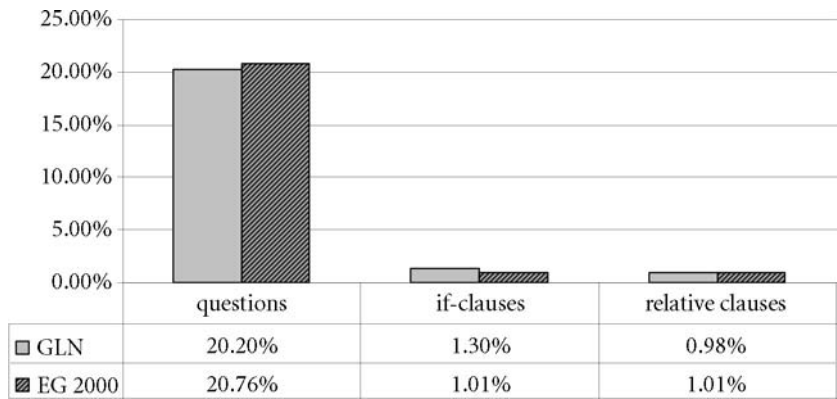


Figure 115. Shares of questions, if-clauses, and relative clauses in the GLN and EG 2000 datasets

ative clauses, it is not possible to detect any lexical-grammatical peculiarities here. We will see later on whether there are any typical patterns in the co-selection of questions and individual progressive verb forms (cf. Section 5.7.7).

- (151) If they knew he was going with a Korean girl, they’d throw him out of the gang – or worse! (GLN)
- (152) They wouldn’t talk if nothing was going on. And something is going on, isn’t it? (EG 2000)
- (153) A big ’Hi!’ to all you kids who are listening out there. (GLN)
- (154) Your father and I are members of an organization that’s bringing him into the country. (EG 2000)

5.5.8 Adverbial specification

The last context feature I will be looking at is the specification of progressives with different kinds of adverbials. Earlier on in Chapter 4 (cf. Section 4.3.8) I referred to the high percentage of “time adverbial and progressive” co-occurrences but also found progressives in spoken BrNSE to collocate repeatedly with other adverbials, such as *actually*. Figure 116 below illustrates the shares of adverbial specification of progressives in coursebook English.

As we see in Figure 116, time adverbials (e.g. *just*, *now*, or *tomorrow*) are most frequent in both GEFL TC corpora but much more common in GLN than in EG 2000. The determined shares of 32.57% vs. 23.54% differ significantly from each other. Significant also is the difference between those progressive tokens in the two datasets that are specified by a place adverbial, e.g. *here* or *in the kitchen*. Only 2.93% of the GLN

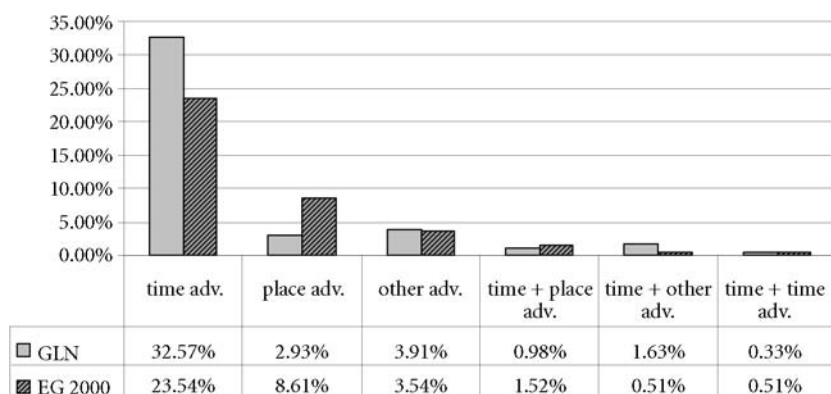


Figure 116. Shares of adverbial specification of progressives in the GLN and EG 2000 datasets

concordance lines contain such an adverbial, while this type of specification is considerably more common in EG 2000 where I found a value of 8.61%. All other types of adverbials and combinations of adverbials are rather infrequent (cf. Figure 116).

Considering the different temporal adverbial types, there are a couple of particularly common ones in the two coursebook English datasets. As Figure 117 shows, *now* and *just* are most frequent among the time adverbials in GLN progressives (for typical examples see (155) and (156)). *When*, *while*, and *today* also occur repeatedly but, as the high percentage of “other” time adverbials indicates, there is quite some variation here, and we find a number of different types which only occur once or twice in all progressives from the GLN coursebooks.

(155) So this is what I’m getting my enjoyment from *now*. (GLN)

(156) The milkman was *just* leaving with the empty bottles. (GLN)

Most frequent among the time adverbials in the EG 2000 dataset are *now* and *when* (cf. examples (157) and (158)). The token numbers of all other types are comparatively small and lie between one and seven, which means that we cannot talk about any other typical collocation patterns. There are a few differences between the time adverbial shares found in GLN and EG 2000 progressives, but due to the low absolute figures none of them is statistically important by the chi-square method.

(157) Take your chance. I’m walking away *now*. (EG 2000)

(158) Yes, Grandma, and I met a cute guy *when* I was getting our drinks. How can I get to know him? (EG 2000)

Place adverbials are not only much less frequent than time adverbials but also show much less type variation. Figure 118 displays the distribution of repeatedly occurring place adverbials in GEFL TC progressives. We see that *here* and *there* occur a few times in both datasets. EG 2000 also has a couple of examples which contain a

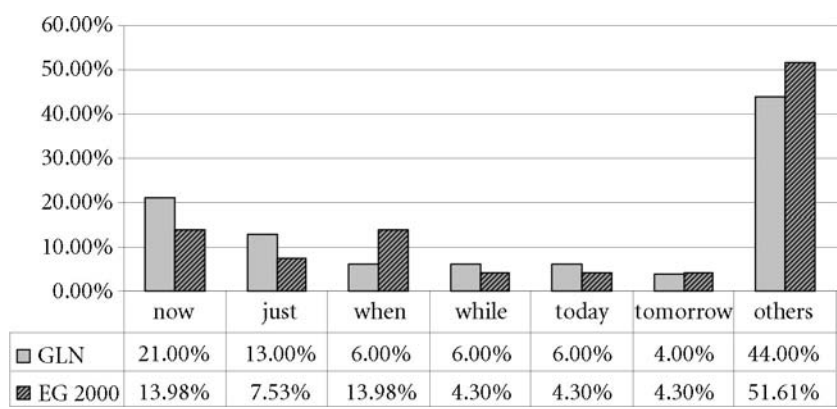


Figure 117. The distribution of time adverbials in the GLN and EG 2000 datasets

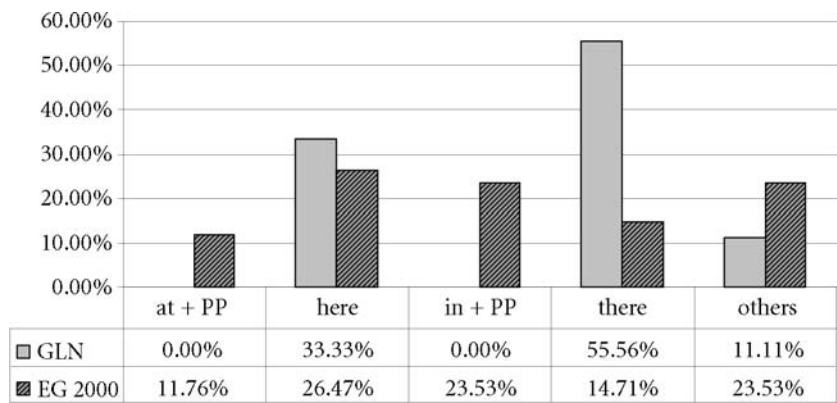


Figure 118. The distribution of place adverbials in the GLN and EG 2000 datasets

prepositional phrase introduced by *in* or, less frequently, by *at*, e.g. *in Blackpool* or *at school*. Concerning the remaining group of “other” adverbials, there is not much to say patternwise. Numbers of occurrence in GLN and EG 2000 are very small (cf. the low shares in Figure 116), and each type only occurs once or twice in the two datasets. Among the adverbials used in GLN are *actually*, *faster*, *only*, *probably*, and *well*. EG 2000 uses largely different types, such as *badly*, *excitedly*, *if* (in the “whether” sense), *quietly*, and *really*.

In Chapter 6 the findings on progressive-adverbial co-occurrences in “school” English will be compared with the respective results obtained from the analysis of spoken English data (cf. Chapter 4). I will now briefly summarise what has been found out about context phenomena in the GEFL TC progressives.

5.5.9 Summary of the findings [GEFL TC – contexts]

In the last few subsections we have made a couple of interesting observations concerning recurring features in the contexts of progressives in textbook English, always following the analytic model that was used in the examination of progressive collocations in spoken BrNSE in Chapter 4.

First of all we looked at the distribution of different tense forms (cf. 5.5.1) and found that, in the coursebook datasets, the most frequent form by far is the present progressive (cf. Figure 100). Particularly common among the PresProg forms are the patterns *are V-ing* and *'m V-ing*, though the percentages of occurrence of these patterns in the GEFL TC subcorpora, GLN and EG 2000, differ to a certain extent (cf. Figures 101 and 102). Past progressives are much less common than PresProgs, and only very few form tokens were found of PresPerfProgs and PastPerfProgs. We then distinguished between contracted and non-contracted progressive forms and found a rather divergent distribution of short and long forms in the two coursebook subcorpora. In EG 2000 we observed a slightly higher share of contracted than non-contracted progressives, while GLN clearly favours non-contracted forms (cf. Figure 105). Also, there were some significant differences between GLN and EG 2000 concerning the shares of individual short and long forms (cf. Figures 106 to 109).

Among the items in subject position of GEFL TC progressives, personal pronouns were clearly most frequent, in particular *I* and *you* (cf. Figures 110 and 111). Repeatedly used also were personal names and determiner phrases. Again, the actual shares of common subjects in GLN and EG 2000 were somewhat different. With respect to objects in progressives we found quite some type variation and observed that *the* + noun group and *a* + noun group are the most frequent recurring types in both textbook corpora. Less than one third of the progressive tokens from GLN and EG 2000 are followed by a preposition. The percentages of the most common prepositions in the two datasets differ considerably, and whereas *at*, *to*, *about*, and *in* are the top four items in GLN, highest shares in EG 2000 are found with *to*, *about*, *on*, and *for*.

The shares of negated progressives in the GEFL TC subcorpora are comparatively low (cf. Section 5.5.6), and of the three other examined lexical-grammatical phenomena, questions, *if*-clauses, and relative-clauses, only the first one shows significant shares of occurrence in GLN and EG 2000 progressives (cf. Figure 115). Finally, among the different kinds of adverbial specification, we found highest values for time adverbial collocates and only rather low shares for place and other adverbials (cf. Figure 116). The share of progressives that contain a time adverbial, such as *now*, *just*, or *when*, is much higher in GLN than in EG 2000.

It is worth noting that the contextual analysis of progressives in GEFL TC has led to some similar but also to some rather different results for the two sets of concordance lines from GLN and EG 2000. If we take into account that both coursebook series are supposed to serve the same purposes and follow the same pedagogical principles, these divergent findings may be unexpected. We will see in the comparison of contexts of progressives in “school” English and spoken BrNSE in Section 6.1 which

of the two textbook series is closer to real English. The next sections of the present chapter will discuss the distribution of progressive functions in the GLN and EG 2000 datasets.

5.6 The use of progressives in “school” English (II) – functions

In Section 4.4 I analysed what was expressed by the 9,468 progressive concordance lines from the BNC_spoken and BoE_brspok corpora. I developed a systematic way of describing the most commonly occurring functions of progressives and determined the distributions of progressive functions and time references in spoken BrNSE.

In the following the focus will be on the same function-related aspects of the use of progressives, i.e. time reference, central functions, additional functions, but this time I will look at “school” English instead of real spoken English and examine my collected datasets from the GLN and EG 2000 textbook subcorpora.

5.6.1 Time reference

As in the function analysis of BNC and BoE data, the first step in analysing progressives from GEFL TC was the determination of the time reference expressed by each form token. I wanted to find out how many of the collected concordance lines refer to events or actions in the past, the present, or the future.

This time reference attribution was much easier for the GEFL TC examples than for progressives from BNC_spoken and BoE_brspok. In the textbook data there were hardly any fragmentary concordance lines or unclear examples and indeterminate cases were also quite rare. I will now look at the distribution of GEFL TC progressives across time references and then discuss the relation between time reference and different types of progressives.

Distribution. In order to determine the percentages of GLN and EG 2000 progressives which refer to the past, present, or future, the respective datasets in the *Access* database were filtered according to the four possible values in the “time reference” box, “past”, “present”, “future”, and “present/future (indeterminate)”. The relative frequencies of occurrence of these four time reference types are visualised in Figure 119.

Highest shares of progressives in both datasets refer to present actions or events, though the actual values for GLN (35.50%) and EG 2000 (43.54%) differ significantly. For GEFL TC examples with present time reference, see (159) and (160). Second most frequent, and equally common in EG 2000 and GLN, are examples which express futurity, such as those displayed in (161) and (162). Rather different again and altogether lower are the shares of progressives with past time reference in the two subcorpora (see for instance (163) and (164)).

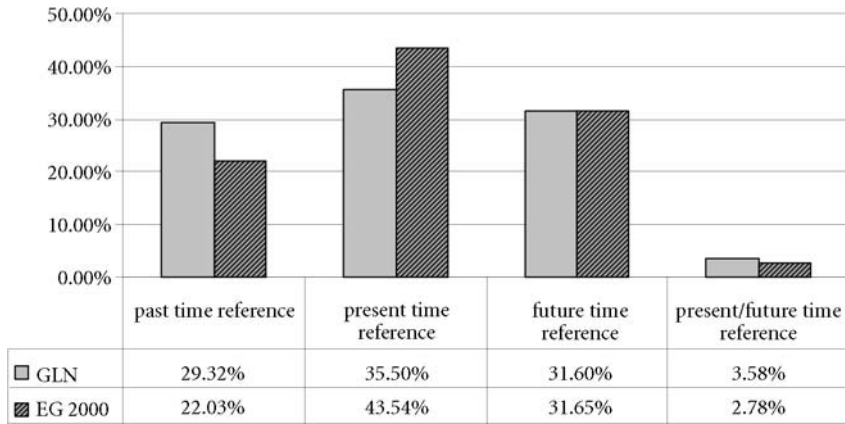


Figure 119. The distribution of time references of progressives in GLN and EG 2000

- (159) “We *are talking* in my tongue now! How long have we been doing that?” (GLN)
- (160) NAZMA: It’s *coming* from behind that bench (EG 2000)
- (161) Jenny: Well, I’m *meeting* Robert at 6 o’clock. (GLN)
- (162) His cousins from Berlin, Katrin and Julia, *are coming* to England. (EG 2000)
- (163) I’ve *been learning* Welsh for three years but I still find it very difficult! (GLN)
- (164) Oh, I *was pulling* a man out of a fire. (EG 2000)

The percentages of tokens with an indeterminate present/future reference are very low (3.58% for GLN and 2.78% for EG 2000), which means that only a small number of GEFL TC progressives refer to something that is valid at the time of speaking and/or at a later point in time, or even valid in general as expressed in example (32) from BNC_spoken, repeated in (165). Two of these “indeterminate” concordance lines from GEFL TC are given in (166) and (167).

- (165) Because that ’s what it ’s all about. Yes. And when you ’re *listening* you ’re watching for body language as well. (BNC_spoken)
- (166) What do you do on an oil rig when you *aren’t working*? (GLN)
- (167) “When you feel it coming, stop what you’re *doing*,” (EG 2000)

It is significant that textbooks favour clear cases of time reference and rarely include instances that may be interpreted in different ways.

Form and function relationship. Having determined the distribution of time references in the GLN and EG 2000 datasets, it will now be interesting to see whether there is a one-to-one relation between the function of temporal orientation and progressive forms, i.e. whether e.g. past time reference is always expressed by means of the

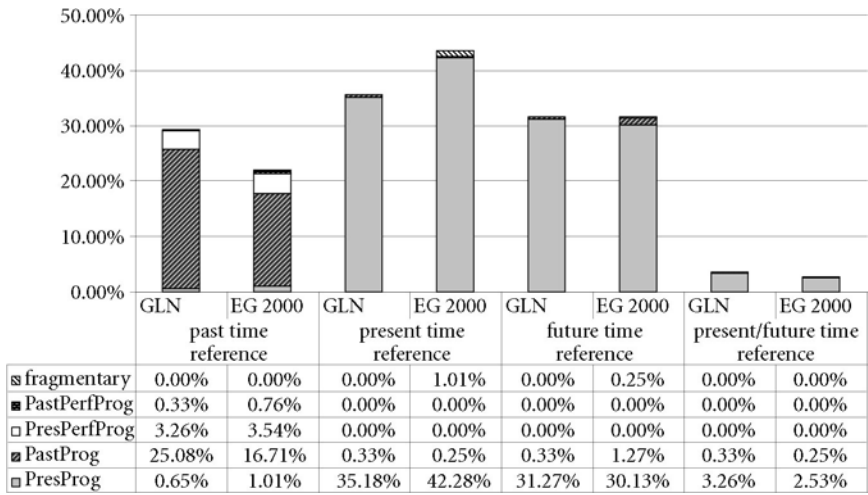


Figure 120. The relationship between time reference and progressive tense form in GLN and EG 2000

PastProg. I hence filtered the GEFL TC progressives in the database according to time reference and sorted them for forms of *to be* which could then be grouped into the respective tense forms. The results of this filtering, sorting, and grouping process are displayed in Figure 120. All results displayed in this figure are very similar for GLN and EG 2000 data.

The two left-hand columns in Figure 120 show that, unsurprisingly, past time reference is mostly expressed by PastProgs, sometimes by PresPerfProgs, and only very rarely by PastPerfProgs. However, there are a couple of PresProg form tokens which relate to the past and are taken from storytelling contexts (see e.g. (168) and (169)).

- (168) So I'm *taking* my pets home. I've got the tickets for everyone. I'm showing them the way to the station. (GLN)
- (169) Anyway. They're *sitting* there laughing, and suddenly this man's in the doorway. I didn't even see him come in. (EG 2000)

With respect to the other types of time reference, the connection between forms and functions is even more straightforward. The tense form which is predominantly used to refer to present, future, and indeterminate actions or events is the present progressive, as shown in examples (170) to (172). Very rarely PastProgs or fragmentary progressive forms with a missing form of *to be* are used to refer to the present or future. Examples of PastProgs with present and future time reference are displayed in (173) and (174). In (173) the past progressive conveys a sense of politeness, and (174) expresses a hypothetical future plan. PresPerfProg and PastPerfProg tokens never express anything but past time reference.

Table 34. Relations between time references and progressive forms in “school” English

Time reference	Forms of progressives
past time reference	predominantly PastProg, some Pres/PastPerfProgs, very few PresProgs
present time reference	predominantly PresProg, very few PastProgs and fragmentary forms
future time reference	predominantly PresProg, very few PastProgs
present/future time reference	predominantly PresProg, very few PastProgs

- (170) Jenny, why *aren't* you *helping* your dad? (EG 2000)
- (171) Great that you're *coming* to visit us for the summer, Daniel! (GLN)
- (172) you're *going* and going and it never seems to get closer. (EG 2000)
- (173) 'Yasmin! Perfect! I *was wondering* who to ask and now you're here!' (EG 2000)
- (174) It was a present from Gaius, the man she *was getting* married to next year. (EG 2000)

We can say that there is a clear relationship between progressive forms and time references which is summarised in Table 34. Deviations from these clear correlation patterns are extremely rare and restricted to a very small number of lexical items.

5.6.2 Two central function features: Continuousness and repeatedness

In my search for a central function or central functions of progressives in spoken BrNSE I found two function features that served particularly well to capture what was expressed by the sets of concordance lines under investigation: continuousness and repeatedness. The realisation of these features in the progressive tokens from GLN and EG 2000 (i.e. continuous vs. non-continuous and repeated vs. non-repeated) will be looked at in the following two short subsections.

Continuousness. To decide whether a progressive form token expresses continuousness, I used the above-specified criteria (cf. 4.4.2) of non-interruption and extension over a certain period of time. That means that examples which referred to very short-termed, punctiform events or to actions that were carried out over a longer time span but with interruptions or breaks were labelled “non-continuous”.

As Figure 121 shows, such non-continuous instances very infrequently occur in the examined coursebook data. Only 4.56% of the GLN progressives and 2.53% of the progressives from EG 2000 refer to non-continuous actions or events, as exemplified in (175) and (176). The large majority of progressives in GEFL TC expresses continuousness. Typical examples are given in (177) and (178). There is only a very small number of verbs that are used in non-continuous contexts. The forms *asking*, *calling*, *looking*, and *putting* belong to this group.

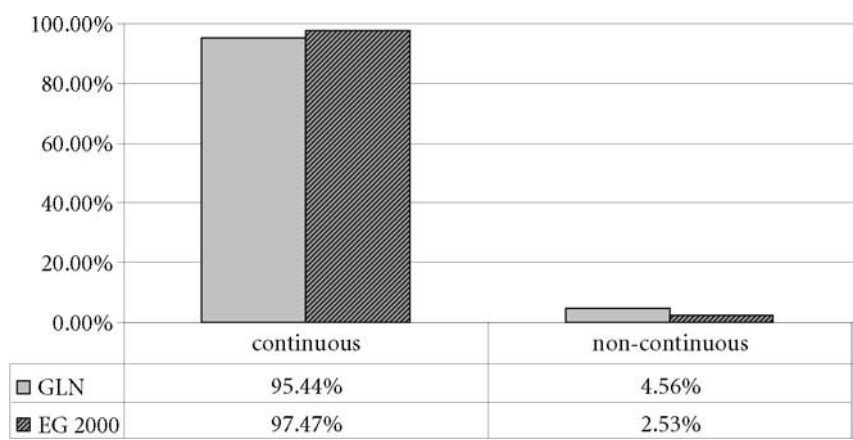


Figure 121. The distribution of progressives referring to continuous and non-continuous actions or events in GLN and EG 2000

- (175) They’re *putting* all their rubbish from their lunch in the water. (GLN)
- (176) So if you need him, you’ve got to ask.’ I’m *asking* now, Doug. (EG 2000)
- (177) The pupil who did it *is* probably *listening* to me now. Usually this is a case for the police. (GLN)
- (178) ‘He’s speaking Latin.’ ‘Yes, but what’s a Roman *doing* here?’ (EG 2000)

On the whole, what the examined textbooks do is provide a clear picture of progressives in which the forms, with very few exceptions, are used with reference to continuous actions and events.

Repeatedness. As the second central feature in the functional interpretation of progressives in spoken BrNSE I determined “repeatedness” and observed that, contrary to some traditional accounts of the progressive which emphasise the expression of single, continued events as the central function of the form, a large number of progressive concordance lines in the BNC_spoken and BoE_brspok datasets refer to repeated actions or events (cf. Figure 37). We see in Figure 122 that this is not quite the case in the EFL textbook data.

More than 90% of the progressives in GLN and EG 2000 express non-repeatedness, such as the two given in (179) and (180). Repeated actions are referred to in less than one tenth of the GEFL TC examples, two of which are displayed in (181) and (182). We will see in Section 5.7.10 which verbs are used in “repeated” contexts in GLN and EG 2000. For the time being, let us just note that, according to what is conveyed by the textbooks, the reference to repeatedly occurring actions is a rather uncommon function of the progressive.

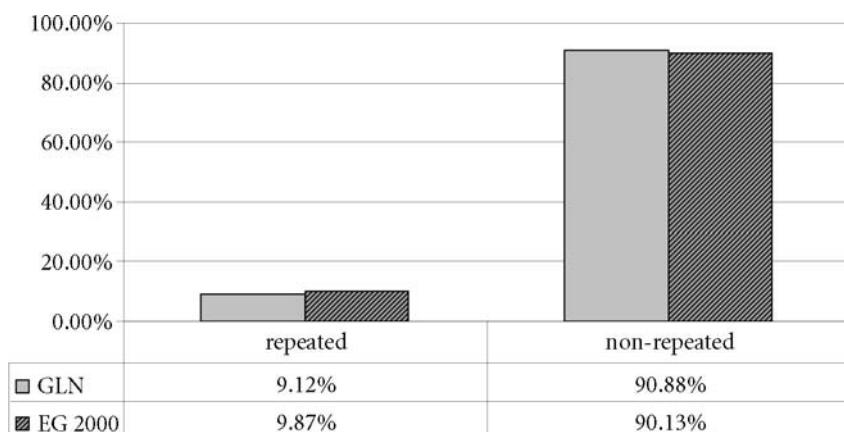


Figure 122. The distribution of progressives referring to repeated and non-repeated actions or events in GLN and EG 2000

- (179) Everyone's in trouble, but Dad's *getting* the worst of it. (EG 2000)
- (180) But then they heard a noise. Someone *was coming* into the classroom! (GLN)
- (181) I listen to my walkman all the time – even when I'm *doing* my homework. (GLN)
- (182) "Where *is* he *learning* these things, Julia?" "He's your son, too, Kenneth," (EG 2000)

5.6.3 Central functions

In Section 4.4.3 we looked at the four possible combinations of the features "continuousness" and "repeatedness" in spoken English and found that two of these combinations, continuousness + non-repeatedness and continuousness + repeatedness, occur particularly frequently in the examined datasets and that they are expressed by almost all included types of verbs. We hence described the two feature combinations as central functions of the progressive: CF1, which refers to continuous and non-repeated actions and events, and CF2, which expresses continuousness and repeatedness. We will now see what the distribution of function feature combinations looks like in the two GEFL TC subcorpora.

Figure 123 illustrates the shares of occurrence of each of the possible feature combinations in GLN and EG 2000. We note that the percentages for the two datasets are quite similar. Both subcorpora show shares of roughly 88 per cent (GLN: 87.62%, EG 2000: 88.35%) of progressives that refer to continuous and non-repeated actions or events, like those given in (183) and (184).

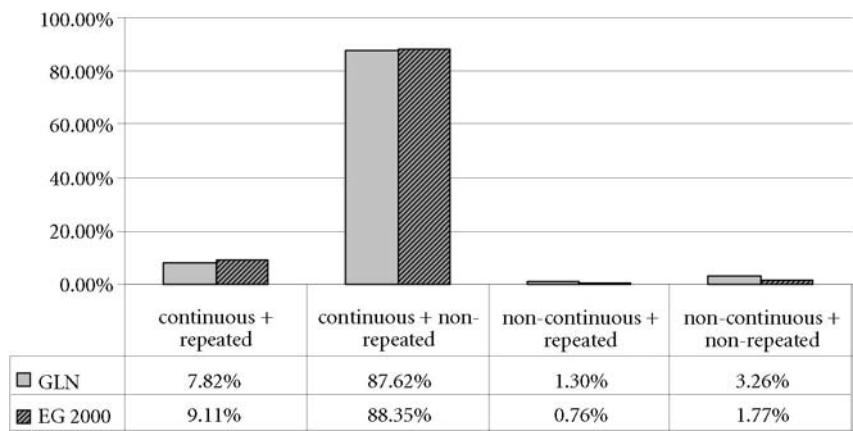


Figure 123. The distribution of central function feature combinations in GLN and EG 2000

- (183) He is making a fruit salad. He *is cutting* the fruit into little pieces. (GLN)
(184) NICK: Hey, look at that duck: it's *eating* our sandwiches! (EG 2000)

There are some instances that express continuousness and repeatedness (see e.g. (185) and (186)), but they only make up 7.82% in GLN and 9.11% in EG 2000. Much less frequent even are examples of the two non-continuous feature combinations, which in fact does not come as a surprise after we found very low values for the expression of non-continuousness earlier (cf. 5.6.2.2).

- (185) For months they hadn't had coffee at all and *had been using* burned oats themselves for a hot morning drink, (GLN)
(186) 'Those stupid kids *are playing* with fireworks again.' (EG 2000)

What the distribution in Figure 123 tells us, is that, in the examined coursebooks, progressives only have one central function: the reference to continuous and non-repeated actions or events. Non-continuousness and repeatedness are marginalised, which may create the impression that there is not more to progressives than the expression of continuousness in combination with single events or actions. It will be interesting to see in the next section whether the picture that the textbooks give of progressive functions is really that monolithic, or whether we can possibly find some reasonable shares of the above-determined additional functions in the GLN and EG 2000 datasets. We can leave out the analysis of relations between the identified central function and types of time reference, an analysis carried out for CF1 and CF2 in spoken English (cf. 4.4.4), as the large majority of GEFL TC progressives conveys the central function CF1, which means that we would get a distribution that is very similar to that illustrated in Figure 119.

Later on (cf. 5.7.10) I will analyse the progressives of which verbs express the less frequent feature combinations (continuous + repeated, non-continuous + repeated,

non-continuous + non-repeated). We will also see whether individual -ing forms show a similar function distribution as the one illustrated in Figure 122 or whether some verbs have particular preferences for a certain function.

5.6.4 Additional functions of the progressive

In the examination of BNC_spoken and BoE_brspok data we noted that progressives often express more than just “continuousness + repeatedness” or “continuousness + non-repeatedness” and that the identified central functions do not suffice to fully capture what progressive forms can actually denote. Seven additional functions were found to be expressed repeatedly in the spoken English datasets: general validity, politeness or softening, emphasis or attitude, shock or disbelief, gradual change and development, old and new habits, and framing.

When I interpreted the collected concordance lines from GLN and EG 2000, I also paid attention to these additional functions and considered whether and how often they were referred to in GEFL TC progressives. The frequencies of occurrence of each additional function in the two textbook corpora are displayed in Figure 124.

At first sight the represented shares of some additional functions, especially in EG 2000, may look impressive. However, the height of the columns in Figure 124 is rather misleading. All illustrated shares lie between 0.25 and 5.82 per cent, which means that on the whole token numbers are rather small. The visible differences between GLN and EG 2000 values are hence not as significant as they might appear. In fact, according to the chi-square test, none of the GLN/EG 2000 distributions is statistically significant.

All additional functions are similarly infrequent in both GEFL TC datasets. Comparatively frequent among these uncommon uses in EG 2000 are “general validity”, “emphasis/attitude”, and “framing”. Examples (187) and (188) refer to situations that can be seen as generally valid, while the utterances in (189) and (190) show an emphatic use of the progressive. Two textbook examples from EG 2000 which contain

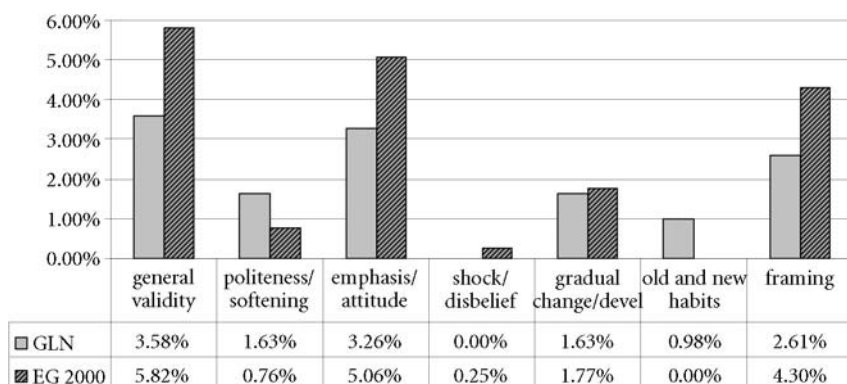


Figure 124. The distribution of additional progressive functions in GLN and EG 2000

so-called “framing” contexts are given in (191) and (192). This additional function of the progressive is also expressed in a couple of GLN concordance lines (cf. (193) and (194)). Worth noting here are the collocation with the adverbials *when* and *while*, and the preference for past progressive forms.

- (187) MARIA: There’s a crowd of tired women who *are living* from day to day. (EG 2000)
- (188) Also, space exploration *is making* the world a safer place, (EG 2000)
- (189) So if you need him, you’ve got to ask.’ *I’m asking* now, Doug. (EG 2000)
- (190) She’s having her baby today. LEWIS: She’s *been having* it all week! (EG 2000)
- (191) ‘I’ll play the CD while you’re *reading*. Then you can tell us if you like it.’ (EG 2000)
- (192) ‘*Weren’t* you *listening* when they called our flight? Let’s go!’ (EG 2000)
- (193) While he *was helping* the mother, a mountain of water broke over them both. (GLN)
- (194) While he *was wondering* what to do, he came to the Cathedral Square. (GLN)

Repeatedly used in GLN as well are the additional functions “general validity” (cf. example (195)), “politeness or softening” (cf. (196)), and “gradual change and development” (cf. (197)). I have to repeat though that all these progressive functions are relatively infrequent in the coursebook data. Progressives which refer to old or new habits (cf. (198)), or which express strong doubt or shock (cf. (199)), occur extremely rarely in GEFL TC.

- (195) ‘We don’t have time to be real teenagers. We’re either *doing* math and science all the time or getting up early to practice with the swim team. (GLN)
- (196) “Actually, I *was thinking* we should just pretend it never happened. (GLN)
- (197) If I’d fallen, I’d have broken every bone in my body - or worse. I *was getting* very tired. (GLN)
- (198) Jenny’s father *is going* to work in Nottingham now, (GLN)
- (199) RUST: Hey, hey! What *are* you *saying*? Are you suggesting these girls are not innocent...? (EG 2000)

We saw that progressive forms in the GLN and EG 2000 datasets mainly refer to continuous and non-repeated actions or events, and that they only seldomly express any of the additional functions that were identified in the examination of spoken BrNSE progressives. Probably due to the rather small token numbers found for additional functions in GLN and EG 2000, significant preferences for particular time references, i.e. for the reference to situations in the past, present, or future, could not be determined. Relations between these rare additional functions and individual verbs will be analysed and discussed in Section 5.7.11. The following section will briefly sum up the results obtained in the analysis of progressive functions in GEFL TC.

5.6.5 Summary of the findings [GEFL TC – functions]

The purpose of this Section 5.6 has been to do an analysis of the functions of progressive forms in “school” English which follows the analytic pattern of spoken BrNSE data (cf. 4.4). Basing my corpus-driven work on GEFL TC data, I intended to find out what is typically expressed by progressives in English language teaching textbooks.

First of all, I looked at the time reference expressed by each of the concordance examples from GLN and EG 2000 and determined the percentages of tokens that refer to the past, present, or future (cf. 5.6.1). I found highest shares for present time reference, followed by future time reference and then past time reference, and detected significant distributional deviations between GLN and EG 2000 concerning past and present time reference. Indeterminate cases which refer to both the present and future and express, for instance, general validity are rather rare in both GEFL TC datasets (cf. Figure 119). The analysis of the relationship between progressive tense forms and temporal orientation led to the conclusion that there are very clear form and function connections in the coursebook progressives, as summarised in Table 34.

The next analytic step was dedicated to the previously identified central function features “continuousness” and “repeatedness”. I determined very similar shares of realisations of these two features in GLN and EG 2000 and clear preferences for continuousness and non-repeatedness (cf. Figures 121 and 122). The distribution of the central feature combinations then showed that “continuous + non-repeated” actions and events are by far most frequently referred to in our textbook datasets (cf. Figure 123), which indicates that we are here only dealing with one central progressive function. Next I wanted to find out whether on the function side there is more to progressives in GEFL TC than the just-mentioned reference to single continuous events. The altogether very low shares of different additional functions of progressive forms in GLN and EG 2000 (cf. Figure 124), e.g. general validity or politeness, prove that the picture the textbooks give of progressive functions is indeed comparatively monolithic and restricted to only one central use: the expression of continuousness paired with non-repeatedness.

With respect to the investigated function phenomena, the differences between the two coursebook subcorpora, GLN and EG 2000, are far less significant than for most of the context features examined earlier (cf. 5.5). Still in order to make sure that I do not miss any important inter-textbook differences, I will treat the GLN and EG 2000 progressives of individual verbs separately when I look at verb-function relations later in parts of Section 5.7. Relations between individual verbs and context phenomena will also be dealt with separately for EG 2000 and GLN.

5.7 Verbs and progressives in GEFL TC – How lexical is EFL textbook grammar?

An important part of the analysis of progressives in spoken BrNSE above dealt with the strength of the relations between progressive constructions and individual verb forms. I found that progressive forms of different verbs behave rather differently concerning

Table 36. The 22 most frequent progressive form types in GEFL TC (in alphabetical order; cf. Table 37 for token numbers)

1. asking	6. having	11. playing	16. taking	21. wearing
2. coming	7. helping	12. reading	17. talking	22. working
3. doing	8. listening	13. saying	18. telling	
4. getting	9. looking	14. sitting	19. thinking	
5. going	10. making	15. staying	20. trying	

their preferred selection of contexts and functions, and noted that the average distributional percentages determined earlier on may be misleading as they do not necessarily tell us a lot about the actual use of all (or most) progressives but may in fact obscure important collocational and colligational tendencies of lexical items.

Having stressed the importance of paying more attention to lexical diversity within the progressive, I will now carry out the same type of verb-individual analysis as above, this time, however, based exclusively on GEFL TC data. My aim will be to discover how strong the relations between individual verbs and progressive constructions, their contexts and meanings are in the coursebooks and hence find out something about the degree of lexicalness of coursebook English grammar.

As the absolute numbers of occurrence in GLN and EG 2000 of most of the analysed -ing forms are rather small (cf. Table 37 below), I will only include in the following verb-specific analyses the 22 most frequent progressives which show token numbers between 8 (e.g. *thinking*) and 187 (*going*). Table 36 lists the selected types in alphabetical order. Those types that just occur seven times or less in the GEFL TC progressive datasets will not be dealt with in detail here. However, should there be anything interesting to note concerning the contexts and functions of these less frequent types, I will include further evidence and present relevant examples from GLN and EG 2000 in the respective section.

5.7.1 The distribution of progressive verb forms in GEFL TC

As mentioned above (cf. 5.4.2), of the 100 selected -ing forms only 67 were found in EG 2000 and just 63 occurred in GLN. Some of these types were not used in progressive constructions but formed different parts of speech, such as nouns (cf. examples (129) and (130) above), which means that a somewhat lower number of types and their respective progressive form tokens were finally added to the database and included in the context and function analysis.

Altogether 63 forms of the original set of 100 were found to occur in progressives in GEFL TC, either in GLN, or in EG 2000, or in both subcorpora. These forms are listed in Table 37, together with their token numbers. Only 49 of these 63 types showed occurrences in GLN, while 55 were found in EG 2000. As Table 37 shows, a very large number of the listed -ing forms only occur very rarely (between 1 and 8 times) in the GEFL TC datasets. Most common are progressives of the forms *going*, *doing*, *looking*, *talking*, and *coming*. *Getting*, *playing*, *sitting*, *making*, *working*, and *try-*

Table 37. The distribution of 702 progressive form tokens across 63 verb types (in order of verb type frequency), based on GEFL TC data (GLN + EG 2000)

verb form	number of tokens	verb form	number of tokens	verb form	number
1. going	187	22. wearing	8	43. bringing	2
2. doing	51	23. happening	7	44. giving	2
3. looking	41	24. learning	6	45. letting	2
4. talking	34	25. leaving	6	46. living	2
5. coming	28	26. running	5	47. showing	2
6. getting	24	27. standing	5	48. being	1
7. playing	23	28. walking	5	49. buying	1
8. sitting	22	29. watching	5	50. checking	1
9. making	21	30. wondering	5	51. cutting	1
10. working	19	31. calling	4	52. drawing	1
11. trying	18	32. eating	4	53. expecting	1
12. telling	16	33. hoping	4	54. feeling	1
13. listening	14	34. moving	4	55. following	1
14. reading	14	35. speaking	4	56. imagining	1
15. taking	13	36. using	4	57. keeping	1
16. having	12	37. writing	4	58. pulling	1
17. saying	11	38. carrying	3	59. putting	1
18. staying	9	39. holding	3	60. ringing	1
19. asking	8	40. meeting	3	61. turning	1
20. helping	8	41. selling	3	62. winning	1
21. thinking	8	42. starting	3	63. worrying	1

ing are also reasonably frequent, but all other types are only used occasionally in the GLN and EG 2000 coursebook volumes. A few textbook examples of some of the not so common progressive types are given in (200) to (204). We shall see in the following sections whether the more frequently occurring forms (the top 22 items in Table 37) show certain patterns with respect to a range of context and function phenomena.

- (200) “Yes, Leinert is the name. I’m *calling* from Green Farm. – Yes. . .” (GLN)
- (201) I mean you wouldn’t know what you *were letting* yourself in for if you didn’t know what they were saying. (GLN)
- (202) Robert: Really? Then why *are they running* away? (GLN)
- (203) ‘Yasmin! Perfect! I *was wondering* who to ask and now you’re here! I can ask you.’ (EG 2000)
- (204) Oh, *am I speaking* too fast? (EG 2000)

5.7.2 Verbs and tense form distributions [GEFL TC]

The first phenomenon I would like to investigate in this verb-specific context is the distribution of different progressive tense forms. It will be analysed which of the 22 most

frequent progressive forms most commonly occur in the PastProg, the PresProg, the PresPerfProg, and the PastPerfProg. Also, we will see in how far the average tense form shares determined earlier on (cf. Figure 100 in 5.5.1) are valid for individual verbs.

Table 38 provides a display of the percentages of different progressive tense forms for selected verbs in GLN and EG 2000. The included mini-diagrams in the right-hand table column serve to visualise the determined distributions. We note two major types of variation, one across verb types, the other one between the two textbook series. With respect to inter-verb variation, there are some significant deviations from the average values that were found for the two sets of progressive form tokens in GLN and EG 2000. Some verb forms show clearly above-average shares of PastProgs, e.g. *getting*, *having*, *sitting*, *taking*, and *talking* in GLN and EG 2000, *helping*, *saying*, *telling*, and *thinking* in GLN, and *trying* and *working* in EG 2000. It is worth noting that some of the verbs which share a preference for PastProg constructions also share certain semantic features; *saying*, *talking*, and *telling* can all be referred to as “speech verbs”. Concordance samples of PastProg-favouring forms are given in Figure 125.

Other verbs in the progressive datasets from GEFL TC have a strong preference for PresProg forms. Among these are *asking*, *going*, *making*, *playing*, *reading*, and *staying*. The concordance sample of *playing* in Figure 126 does not show a single occurrence of PastProg, PresPerfProg, or PastPerfProg forms, but exclusively instances of the present progressive. PresPerfProgs are generally infrequent in the coursebook data but comparatively common with *asking*, *having*, *looking*, *sitting*, and *trying* in EG 2000, with *working* in GLN, and in both subcorpora with *learning*, which is not one of the 22 most frequent GEFL TC verbs.

Shawna, what are you saying?" asked Mr. Zorzi. Well, er, I'm
for if you didn't know what they were saying. I didn't speak any Italian when
im to be hurt again. do you?" Brede was saying, cutting into her thoughts. "Of c
me to see if I understood what he was saying. "If you lose," he went on, "it ju
I couldn't believe what Mr. Zorzi was saying. Shawna spoke up right away. "Mr.
I'd want to know what the fellows are saying. I mean you wouldn't know what you
TED I never broke her arm. What you talking about? Let go of me ... BEN (in al
were reading. Ben and his grandma were talking. 'This is boring. Let's go and buy
ves on the new vocabulary. Others were talking in Urdu. 'Ladies,' Mrs Mir began,
o understand groups of people who were talking together. 'They would suddenly sta
o understand groups of people who were talking together. 'They would suddenly sta
n the train back to Chester Debbie was talking to Mr Martin. 'There's one thing I
anslated. The boy spoke again. He was talking about someone named Mr. Malik and
morning at breakfast the Mars Team was talking excitedly. 'Gee, Dad never gives u
leaving tomorrow!' Meanwhile Tina was talking to the electrician. 'Just look at
nger sister. You know, the guy who was talking to you earlier about baseball. Pre
I don't understand,' she said. 'I was talking to an old Chinese man at the para

Figure 125. Extracts from a GLN concordance of *saying* and an EG 2000 concordance of *talking*, illustrating high shares of PastProgs

m. Where's your saxophone? Aren't you playing it today? Sam: Of course, I am. It'
he playing? Is she playing? they playing? He is reading. She is reading. Th
doing after school on Monday? She's playing hockey. 134 c
Is he playing? Is she playing? they playing? He is reading. Sh
"Well?" Lara asked. "I can't. I'm playing in a football match in the morning.
ke to ... Jenny: Sorry, Edward! I'm playing in a hockey match after school. Edw
Mr Jones: Look, here comes Sam. He's playing his saxophone! Tommy: (clapping his
Is he playing? Is she playing? they playing? H
Nottingham. Now Jenny and Mark are playing a game on the floor together. "Liste

Figure 126. Extract from a GLN concordance of *playing*, showing exclusively PresProg forms

Table 38. The distribution of progressive tense forms for selected verbs in GLN (G) and EG 2000 (E, values in italics)⁸¹

verb form	G or E	Past Prog	Pres Prog	PresPerf Prog	PastPerf Prog	distribution graphically illustrated
		ø26.06% ø18.48%	ø70.03% ø75.95%	ø3.26% ø3.54%	ø0.33% ø0.76%	<div> <div>PastProg</div> <div>PresProg</div> <div>PresPerfProg</div> <div>PastPerfProg</div> </div>
asking	G	0.00%	100.00%	0.00%	0.00%	
	E	0.00%	80.00%	20.00%	0.00%	
coming	G	28.57%	71.43%	0.00%	0.00%	
	E	0.00%	100.00%	0.00%	0.00%	
doing	G	10.53%	78.95%	5.26%	0.00%	
	E	12.50%	75.00%	3.13%	0.00%	
getting	G	37.50%	62.50%	0.00%	0.00%	
	E	37.50%	62.50%	0.00%	0.00%	
going	G	9.89%	89.01%	1.10%	0.00%	
	E	7.29%	92.71%	0.00%	0.00%	
having	G	40.00%	60.00%	0.00%	0.00%	
	E	28.57%	57.14%	14.29%	0.00%	
helping	G	50.00%	50.00%	0.00%	0.00%	
	E	0.00%	100.00%	0.00%	0.00%	
listening	G	25.00%	75.00%	0.00%	0.00%	
	E	33.33%	66.67%	0.00%	0.00%	
looking	G	40.00%	60.00%	0.00%	0.00%	
	E	14.29%	76.19%	9.52%	0.00%	
making	G	25.00%	75.00%	0.00%	0.00%	
	E	11.11%	88.89%	0.00%	0.00%	
playing	G	0.00%	100.00%	0.00%	0.00%	
	E	7.14%	85.72%	7.14%	0.00%	
reading	G	25.00%	75.00%	0.00%	0.00%	
	E	20.00%	80.00%	0.00%	0.00%	
saying	G	66.67%	33.33%	0.00%	0.00%	
	E	0.00%	100.00%	0.00%	0.00%	
sitting	G	42.86%	57.14%	0.00%	0.00%	
	E	26.67%	46.67%	20.00%	6.67%	
staying	G	20.00%	80.00%	0.00%	0.00%	
	E	0.00%	100.00%	0.00%	0.00%	
taking	G	40.00%	60.00%	0.00%	0.00%	
	E	25.00%	75.00%	0.00%	0.00%	
talking	G	40.00%	60.00%	0.00%	0.00%	
	E	45.83%	50.00%	0.00%	4.17%	
telling	G	66.67%	33.33%	0.00%	0.00%	
	E	7.69%	92.31%	0.00%	0.00%	
thinking	G	83.33%	16.67%	0.00%	0.00%	
	E	0.00%	50.00%	50.00%	0.00%	
trying	G	25.00%	75.00%	0.00%	0.00%	
	E	40.00%	50.00%	10.00%	0.00%	
wearing	G	50.00%	50.00%	0.00%	0.00%	
	E	16.67%	83.33%	0.00%	0.00%	
working	G	28.57%	57.14%	14.29%	0.00%	
	E	41.67%	58.33%	0.00%	0.00%	

Of the selected verb forms only *sitting* and *talking* were found to occur in the PastPerfProg, and only in EG 2000 concordances. This observation leads us to the differences between the two examined textbook series. Although the previous analysis of progressive tense forms in the GLN and EG 2000 datasets led to roughly comparable results for the group of all included progressives (cf. Figure 100), we find some significant distributional discrepancies when it comes to individual verbs. As the diagrams in Table 38 indicate, some -ing forms show considerably high percentages of occurrence with a particular tense form in one GEFL TC subcorpus but much lower shares with the same tense form in the other. For instance, the PastProg shares of *coming*, *looking*, *sitting*, *telling*, and *thinking* are much higher in GLN than in EG 2000, whereas the EG 2000 concordances of *helping*, *saying*, *telling*, and *wearing* contain many more PresProgs than the respective GLN datasets.

The findings indicate that the above-specified average distribution of progressive tense forms in the GEFL subcorpora is in fact rather misleading as it neither accounts for the preferences of individual verbs nor captures the differences between GLN and EG 2000. In how far the textbook-based results are in accordance with what was observed on the basis of real spoken English will be discussed in Section 6.3.

5.7.3 Verbs and subjects [GEFL TC]

When we analysed the collocational relations between progressive forms and their subjects in GEFL TC, we found particularly high average shares of personal pronouns in subject position of progressives (cf. Figure 111 in 5.5.3).

If we now look at the percentages of personal pronoun subjects for individual verbs listed in Table 39, we see that the determined average values are not very representative of a large number of items.⁸² Many verbs show either highly above- or below-average shares of the different personal pronouns in subject position. *Getting*, *listening*, *looking*, *taking*, and *thinking*, for instance, have a clear preference for first person singular pronoun subjects (see concordance sample in Figure 127). The pronoun *you* often forms the subject of progressives with *doing* in both textbook corpora, with *having* in GLN, and with *listening*, *reading*, *thinking*, and *wearing* in EG 2000. Figure 128 shows part of the GLN concordance of *doing*, which illustrates the co-selection of this progressive form and *you* in subject position, and which nicely highlights the typical pattern “what are you doing” (see also my comments on questions in 5.7.7).

Among the third person singular pronouns found in subject position of progressives, *he* is generally most common, especially with *helping*, *telling*, *trying*, and *wearing* in GLN and with *staying* in EG 2000. Even though *she* is, on average, not a very frequently used subject, some verbs show rather high shares of co-occurrence with this pronoun, e.g. *listening* and *reading*. Only very few verbs select *it* as their subject, but those forms which do (*coming*, *getting*, *taking*, *working*) show clearly above-average shares.

Some definite verb-specific preferences and hence deviations from the average distributions can also be observed with respect to the two plural personal pronouns *we*

Table 39. The distribution of personal pronoun subjects across verb forms in GLN (G) and EG 2000 (E, values in *italics*)

verb form	G or E	subject <i>I</i>	subject <i>you</i>	subject <i>he</i>	subject <i>she</i>	subject <i>it</i>	subject <i>we</i>	subject <i>they</i>
		ø25.84% <i>ø35.13%</i>	ø20.57% <i>ø22.94%</i>	ø15.79% <i>ø10.04%</i>	ø5.26% <i>ø7.89%</i>	ø1.44% <i>ø2.51%</i>	ø16.27% <i>ø11.47%</i>	ø14.83% <i>ø10.04%</i>
asking	G	0.00%	33.33%	0.00%	0.00%	0.00%	0.00%	66.67%
	E	<i>40.00%</i>	<i>20.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
coming	G	4.76%	4.76%	9.52%	0.00%	0.00%	4.76%	14.29%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>14.29%</i>	<i>14.29%</i>	<i>0.00%</i>	<i>14.29%</i>
doing	G	5.26%	47.37%	5.26%	5.26%	0.00%	10.53%	5.26%
	E	<i>6.25%</i>	<i>53.13%</i>	<i>9.38%</i>	<i>6.25%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>6.25%</i>
getting	G	37.50%	0.00%	0.00%	0.00%	12.50%	12.50%	0.00%
	E	<i>31.25%</i>	<i>25.00%</i>	<i>0.00%</i>	<i>6.25%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
going	G	27.47%	17.58%	4.40%	2.20%	2.20%	23.08%	4.40%
	E	<i>31.25%</i>	<i>16.67%</i>	<i>4.17%</i>	<i>3.13%</i>	<i>3.13%</i>	<i>19.79%</i>	<i>4.17%</i>
having	G	0.00%	40.00%	0.00%	0.00%	0.00%	0.00%	40.00%
	E	<i>14.29%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>28.57%</i>	<i>0.00%</i>	<i>28.57%</i>	<i>14.29%</i>
helping	G	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>33.33%</i>	<i>16.67%</i>	<i>0.00%</i>	<i>16.67%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
listening	G	37.50%	12.50%	0.00%	12.50%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>33.33%</i>	<i>0.00%</i>	<i>33.33%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
looking	G	10.00%	0.00%	15.00%	10.00%	0.00%	15.00%	20.00%
	E	<i>47.62%</i>	<i>4.76%</i>	<i>4.76%</i>	<i>9.52%</i>	<i>0.00%</i>	<i>9.52%</i>	<i>4.76%</i>
making	G	16.67%	0.00%	8.33%	0.00%	0.00%	8.33%	16.67%
	E	<i>22.22%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>11.11%</i>	<i>0.00%</i>	<i>11.11%</i>	<i>0.00%</i>
playing	G	22.22%	11.11%	22.22%	22.22%	0.00%	0.00%	11.11%
	E	<i>7.14%</i>	<i>7.14%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>7.14%</i>	<i>35.71%</i>
reading	G	0.00%	0.00%	25.00%	25.00%	0.00%	0.00%	25.00%
	E	<i>40.00%</i>	<i>30.00%</i>	<i>0.00%</i>	<i>20.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
saying	G	0.00%	16.67%	16.67%	0.00%	0.00%	0.00%	16.67%
	E	<i>40.00%</i>	<i>20.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>40.00%</i>
sitting	G	0.00%	14.29%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>13.33%</i>	<i>6.67%</i>	<i>6.67%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>26.67%</i>
staying	G	0.00%	20.00%	0.00%	0.00%	0.00%	20.00%	40.00%
	E	<i>25.00%</i>	<i>25.00%</i>	<i>50.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
taking	G	60.00%	0.00%	0.00%	0.00%	0.00%	0.00%	60.00%
	E	<i>37.50%</i>	<i>12.50%</i>	<i>12.50%</i>	<i>0.00%</i>	<i>12.50%</i>	<i>0.00%</i>	<i>37.50%</i>
talking	G	0.00%	0.00%	20.00%	0.00%	0.00%	20.00%	20.00%
	E	<i>4.17%</i>	<i>20.83%</i>	<i>4.17%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>12.50%</i>
telling	G	0.00%	0.00%	33.33%	0.00%	0.00%	0.00%	33.33%
	E	<i>38.46%</i>	<i>7.69%</i>	<i>7.69%</i>	<i>15.38%</i>	<i>0.00%</i>	<i>15.38%</i>	<i>0.00%</i>
thinking	G	33.33%	16.67%	16.67%	0.00%	0.00%	0.00%	0.00%
	E	<i>50.00%</i>	<i>50.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
trying	G	25.00%	12.50%	37.50%	0.00%	0.00%	0.00%	12.50%
	E	<i>30.00%</i>	<i>0.00%</i>	<i>10.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>20.00%</i>	<i>10.00%</i>
wearing	G	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>33.33%</i>	<i>16.67%</i>	<i>16.67%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
working	G	0.00%	14.29%	14.29%	14.29%	0.00%	0.00%	0.00%
	E	<i>25.00%</i>	<i>0.00%</i>	<i>16.67%</i>	<i>8.33%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>

... my head is missing and I'm still looking for it. Oh please help me!' Jenny and coming, but she can't see it. She's looking the wrong way! Now Ben and Julia are on.' Let's go,' Katrin says. She's looking across the road. A car is coming, but onq said with a smile: 'I'm so much looking forward to Tet.' Scott asked his moth CUSTOMER 1 No, thank you. I'm just looking. ASSISTANT Good afternoon. May I he u girls? No, thank you. We're just looking. They've got great things. But look t it was St James's Park. MANNY I'm looking after it for him while he's away. No, lp you? CUSTOMER 2 Yes, please. I'm looking for my head and I must find it. The be anyone comes in, I'll tell them I'm looking for a present for my mother. ASSIST d said to the woman there, 'Hi. I'm looking for the brush Gran lost a couple of n and said, 'Sorry, Megan. I've been looking for a Dancing Lesson. I'm trying to v and they've been sitt

Figure 127. Extract from an EG 2000 concordance of *looking*, illustrating a high share of the personal pronoun *I* in subject position

David: What are you doing, Mark? Mark: I'm drawing a birthday r f the shop. "Hello, Becky. What are you doing here?" Robert asked. "I'm just taking n in the kitchen. Becky: What are you doing? Robert: I'm making a salad. Simon. eyes and ran after her. "What are you doing here?" I asked. I decided it was tin our hockey stick, too. W-what are you doing tonight, Jenny? I'm going to the youth "Excuse me," Becky said. "What are you doing?" "Me?" the man said. He stood up, b ent behind the prisoner. "What are you doing with that rope?" shouted the Sheriff. his glider. "How you doing?" the kid asked him. "Fine," Steve s turned to her. He looked her you were doing," he said. "Listen, I'm sorry straight , he wondered for a moment why he was doing this; then Joe's instructions took and watch the two boys. "What are they doing?" asks Robert. Simon gets his purse c ible bully," says Becky. "What's Simon doing with his brother?" "I don't know," s 's a good idea, Mark. Let's. What's she doing here? make a castle with water round when Becky said, "Hey, what's that man doing over there?" "Which man?" Robert aske talk about the future! What is Jenny doing after school on Monday? She's playir obert: Hello, Mrs Richards. What's Jack doing? Mrs R.: He's feeding Lenny. Lenny's walkman all the time - even when I'm doing my homework. The pop industry has or ime to be real teenagers. We're either doing math and science all the time or get i my tongue now! How long have we been doing that?" "Since you told me about your

Figure 128. Extract from a GLN concordance of *doing*, illustrating a high share of the personal pronoun *you* in subject position

and *they*. Particularly high percentages were found for the following subject-verb collocations: *we + going* (GLN and EG 2000), *we + having* (EG 2000), *we + staying* (GLN), *we + talking* (GLN), *we + trying* (EG 2000), *they + asking* (GLN), *they + having* (GLN), *they + playing* (EG 2000), *they + staying* (GLN), *they + taking* (GLN and EG 2000), and *they + telling* (GLN). As the actual token numbers for most of these collocations are rather small, it may be somewhat risky to speak of "significant co-selection patterns" at this stage. However, the findings indicate some collocational preferences of certain verbs, and prove that within a grammatical construction (in our case the progressive) individual items are presented in different ways. Again, we observed quite a number of differences between the two selected coursebook series. In the comparison of spoken English and EFL textbook corpus findings we will see in which of the two series the patterns are closer to real English.

5.7.4 Verbs and objects [GEFL TC]

Let us now turn to looking at the co-occurrence of selected progressive types and their objects. We noted earlier on (cf. 5.5.4) that, if we take a closer look at the GEFL TC progressives, there is a lot of variation among the words and phrases that appear in object position, and that the only comparatively frequent objects are nouns or noun groups introduced by *a* or *the* (e.g. *a party*, *the birds*).

If we analyse the preferences of individual verbs for certain objects, we find that there is only a small number of repeatedly occurring combinations in both GEFL TC subcorpora. Hence, a detailed statistical evaluation in form of a complex table or diagram is not required. It will be sufficient to merely list a few observed co-selection phenomena and briefly comment on them. The following verb-object patterns were found in the GLN concordances (optional items in parentheses):

- *going* + preposition + *the* + noun (group)
e.g. *to the cinema, down the hall, into the hospital, towards the pond;*
- *looking* + preposition + *the* + noun (group)
e.g. *into the forest, through the files of people I know, at the house, out of the window, at the notice board;*
- *making* + *a* + noun (group)
e.g. *a mistake, a fruit salad, a terrible noise, a cake.*

Repeatedly co-selected in the EG 2000 data were:

- *going* + preposition (mainly *to*) + *the* + noun
e.g. *to the venue, to the museum, to the cinema, on a journey;*
- *having* + *a* + noun (group)
e.g. *a problem, a Puerto Rican meal, a snack;*
- *looking* + (preposition) + *the* + noun (group)
e.g. *at the ceiling, the wrong way, for the brush;*
- *making* + *the* + noun (group)
e.g. *the video, the speech, the invitation cards for next Saturday;*
- *telling* + *him*;
- *telling* + *you*.

As we can see, the verb-object co-selection patterns vary across the textbook corpora in number and type. Individual verbs have certain collocational preferences, but these preferences differ to some extent between GLN and EG 2000. *Making*, for instance, is repeatedly followed by *a* plus a noun or noun group in GLN (cf. examples (205) and (206)), a pattern which does not occur in any of the EG 2000 progressives where *making* is typically postdetermined by *the* plus noun or noun group, as in (207) and (208).

- (205) “No, wait!” said Jenny. “We’re making *a mistake*. I think Emma could be the thief.” (GLN)
- (206) Tommy: Hm. And why are Mum and Tracey making *a cake*? (GLN)
- (207) Also, space exploration is making *the world* a safer place. (EG 2000)
- (208) ‘I’d have come if I hadn’t been so busy with my Hebrew – and *the speech* I’m making.’ (EG 2000)

The listed object patterns of *going* and *looking* were found in both GEFL TC subsections, but the repeated combinations of *having* + *a* + noun (group), *telling* + *him*, and *telling* + *you* only feature in EG 2000 (see examples (209) to (211)).

- (209) 'My parents say you can eat with us if you want. We're having a *Puerto Rican meal* and you'd be welcome.' (EG2000)
- (210) TRUNDLE: Because he's the director, Marmalade. She's telling *him*. He's telling her'. (EG 2000)
- (211) JANINE: It's because we're your friends that we're telling *you* this. (EG 2000)

In a next analytic step I will examine whether similar selectional preferences can also be found for progressives of selected verbs and prepositions.

5.7.5 Verbs and prepositions [GEFL TC]

In Section 5.5.5 I determined the average distributional shares of common prepositions in progressive constructions and noted that, naturally, the frequency of occurrence of prepositions is highly dependent on the use of the verbs they postmodify. I will now turn my attention to the co-occurrence of some common prepositions and individual verb forms and search for the most typical preposition-progressive patterns in GLN and EG 2000.

Table 40 provides the co-selectional shares of six common prepositions and the 22 most frequent verbs in GEFL TC. The most typical collocations that can be extracted from this table are, for GLN:

"coming back", "doing with", "getting at", "looking at", "looking for" (significantly less frequent than "looking at"), "staying at", "staying with", "talking about", "talking with", and "thinking about",

and for EG 2000:

"asking about", "asking for", "coming back", "looking at", "looking for" (significantly more frequent than "looking at"), "playing with", "sitting on", "staying with", "talking about", "thinking about", and "working on".

In addition, two collocations were found in both textbook corpora with the preposition *to*: "listening to" and "talking to". A selection of concordance samples which serve to illustrate some typical co-selection patterns are displayed in Figure 129.

My analysis has revealed that GLN and EG 2000 share a number of the identified progressive-preposition collocations, though not all of them. The comparison of BNC/BoE and GLN/EG 2000 data in Chapter 6 will reveal which of the repeatedly occurring patterns are actually most typical of spoken British English and whether there are more differences or more similarities between "school" English and real English.

Table 40. The distribution of frequently used prepositions across selected verb forms in GLN (G) and EG 2000 (E, values in *italics*)

verb form	G or E	preposition <i>about</i>	preposition <i>at</i>	preposition <i>back</i>	preposition <i>for</i>	preposition <i>on</i>	preposition <i>with</i>
		<i>ø9.68%</i> <i>ø13.27%</i>	<i>ø17.20%</i> <i>ø4.42%</i>	<i>ø4.42%</i> <i>ø4.30%</i>	<i>ø6.45%</i> <i>ø9.73%</i>	<i>ø4.30%</i> <i>ø14.16%</i>	<i>ø5.38%</i> <i>ø4.42%</i>
asking	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>40.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>20.00%</i>	<i>0.00%</i>	<i>0.00%</i>
coming	G	0.00%	4.76%	<i>9.52%</i>	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>14.29%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
doing	G	0.00%	0.00%	0.00%	0.00%	0.00%	<i>10.53%</i>
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>3.13%</i>
getting	G	0.00%	12.50%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>6.25%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
going	G	0.00%	0.00%	2.20%	0.00%	3.30%	1.10%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>3.13%</i>	<i>0.00%</i>	<i>5.21%</i>	<i>0.00%</i>
having	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
helping	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
listening	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
looking	G	0.00%	<i>55.00%</i>	0.00%	<i>20.00%</i>	0.00%	0.00%
	E	<i>0.00%</i>	<i>19.05%</i>	<i>0.00%</i>	<i>52.38%</i>	<i>0.00%</i>	<i>0.00%</i>
making	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
playing	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>7.14%</i>
reading	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
saying	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
sitting	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>13.33%</i>	<i>0.00%</i>
staying	G	0.00%	<i>60.00%</i>	0.00%	0.00%	0.00%	<i>20.00%</i>
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>75.00%</i>
taking	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
talking	G	<i>60.00%</i>	0.00%	0.00%	0.00%	0.00%	<i>10.00%</i>
	E	<i>37.50%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
telling	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>15.38%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
thinking	G	<i>50.00%</i>	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>50.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
trying	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
wearing	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
working	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>41.67%</i>	<i>0.00%</i>

'm making a salad. Simon. And you're listening to Radio 1. Robert: Well, no. I'm sure she was, it's just that I wasn't listening you because you're a bit older, res-
 erry?" she asked. But Jerry wasn't listening. William Weston was a famous produ-
 s must sleep in the kitchen tonight," listening to the radio. says Mrs Burton afte-
 ly. The pupil who did it is probably listening to me now. Usually this is a case
 o Radio 1. Robert: Well, no. I'm not listening to Radio 1. It's Radio Nottingham.
 Radio 1. It's Radio Nottingham. I'm listening to my mum. Becky: Your mum? But s
 A big 'Hi!' to all you kids who are listening out there. This week I want your o
 'Please help me. I'm looking for my head and I must find it. The b
 elp you? CUSTOMER 2 Yes, please. I'm looking for a present for my mother. ASSIST
 f anyone comes in, I'll tell them I'm looking for the brush Gran lost a couple of
 nd said to the woman there, 'Hi. I'm looking for a Dancing Lesson. I'm trying to
 d and said, 'Sorry, Megan. I've been looking for my friends. And they've been sit
 d. It was Hannah. 'Megan, we've been looking for you. Is everything all right?' s
 Ben and the three girls are looking at Katrin and Nick across the road.
 when I think that so many people are looking at them in the parade on Thanksgiving
 er Airport Katrin and her sister are looking for Ben and his dad. They can't see
 ill. He had misunderstandings. 'We're talking about a very red hair and freckles a
 nd they needed all the ago they were talking about trying to get space they had.
 in a refinery breathing fire. Bob was talking about his cousin Jimmy, who played
 Christoph was upset. Everyone else was talking and laughing with their exchange pa
 like that didn't interest him. He was talking about his relatives. I didn't reall
 o the Houses of Parliament. Catesby is talking to the owner of the house. Speaker
 ie chase continues. The FARMER is talking about BONNIE and CLYDE to the repor
 i has just come back to England. He is talking to Catesby, Wright, Percy and Winte
 Today SMD are talking about the Middle Ages - around 1200
 oked at him in quick surprise. "We are talking in my tongue now! How long have we

Figure 129. Extracts from GLN and EG 2000 concordances, illustrating the progressive-preposition collocations “listening to”, “looking for”, and “talking about”

5.7.6 Verbs and negation [GEFL TC]

In the contextual analysis of the whole set of GEFL TC progressives I also looked at the phenomenon of negation and found rather low shares of negated progressive forms in both EFL coursebook subcorpora (4.56% in GLN and 3.80% in EG 2000). However, as some of our previous findings indicate (cf. 4.5 and 5.7.2 to 5.7.5), such average values are not necessarily representative of all, or most, progressive types. Therefore, I will now turn to determining the shares of negation for the progressive forms of some particularly common verbs in my GEFL TC datasets.

The percentages of negated and non-negated progressives for the selected verbs are listed and graphically illustrated in Table 41. We see that only half of the verbs actually appear in negative contexts, some of them showing rather low shares of negation. Some verbs, however, show percentages of negation that lie far above the specified average values. Worth mentioning in this context are especially the forms *making* and *working*, which have comparatively high negated shares in both textbook English corpora. Above-average values were also found with *coming* and *listening* in GLN and with *asking*, *getting*, *reading*, *telling*, and *wearing* in EG 2000 (see Figure 130 for concordance samples). This last observation again indicates that there are some major distributional differences between GLN and EG 2000 concerning progressives and context features. It will be discussed later how the verb-individual findings reported on here relate to what has been said about verbs and negation in the analysis of spoken BrNSE.

Table 41. Shares of negated and non-negated progressive forms for selected verbs in GLN (G) and EG 2000 (E, values in *italics*)

verb form	G or E	negated	non- negated	distribution graphically illustrated ■ negated ■ non-negated
		ø4.56% ø3.80%	ø95.44% ø96.44%	
asking	G	0.00%	100.00%	
	E	20.00%	80.00%	
coming	G	9.52%	90.48%	
	E	0.00%	100.00%	
doing	G	0.00%	100.00%	
	E	0.00%	100.00%	
getting	G	0.00%	100.00%	
	E	18.75%	81.25%	
going	G	4.40%	95.60%	
	E	3.13%	96.88%	
having	G	0.00%	100.00%	
	E	0.00%	100.00%	
helping	G	0.00%	100.00%	
	E	0.00%	100.00%	
listening	G	37.50%	62.50%	
	E	0.00%	100.00%	
looking	G	5.00%	95.00%	
	E	0.00%	100.00%	
making	G	16.67%	83.33%	
	E	11.11%	88.89%	
playing	G	0.00%	100.00%	
	E	0.00%	100.00%	
reading	G	0.00%	100.00%	
	E	10.00%	90.00%	
saying	G	0.00%	100.00%	
	E	0.00%	100.00%	
sitting	G	0.00%	100.00%	
	E	0.00%	100.00%	
staying	G	0.00%	100.00%	
	E	0.00%	100.00%	
taking	G	0.00%	100.00%	
	E	0.00%	100.00%	
talking	G	0.00%	100.00%	
	E	0.00%	100.00%	
telling	G	0.00%	100.00%	
	E	7.69%	92.31%	
thinking	G	0.00%	100.00%	
	E	0.00%	100.00%	
trying	G	0.00%	100.00%	
	E	0.00%	100.00%	
wearing	G	0.00%	100.00%	
	E	16.67%	83.33%	
working	G	14.29%	85.71%	
	E	8.33%	91.67%	

'm making a salad. Simon. And you're listening to Radio 1. Robert: Well, no. I'm sure she was, it's just that I wasn't listening you because you're a bit older, res-
 erry?" she asked. But Jerry wasn't listening. William Weston was a famous produ-
 s must sleep in the kitchen tonight," listening to the radio. says Mrs Burton afte-
 ly. The pupil who did it is probably listening to me now. Usually this is a case
 o Radio 1. Robert: Well, no. I'm not listening to Radio 1. It's Radio Nottingham.
 Radio 1. It's Radio Nottingham. I'm listening to my mum. Becky: Your mum? But s
 A big 'Hi!' to all you kids who are listening out there. This week I want your o
 The Aberaeron school rugby team was getting ready for training. There were two re
 of my way, do you hear? You're not getting no money, not from me. Now piss off,
 That's when Stiff said, 'You're not getting to me, you're not.' But even as he sa
 little angry. He said, 'You're not getting to me. You're not.' It was when we pu
 t tractor starter handles. no, I'm getting a cramp in the leg from running. Dou
 r not. Kate, can't you see what I'm getting at?" "No group should have veto powe
 o. Everyone's in trouble, but Dad's getting the worst of it. Remember how he use
 el I know. Help. My leg cramp's getting worse, Doug. I can't run much further
 ; it's noticeable that the boys are getting wider and wider. I tried to jump a gap
 t. It's 2 o'clock. The children are getting on the school bus. 'Please, Miss Hur

Figure 130. Extracts from a GLN concordance of *listening* and an EG 2000 concordance of *getting*, illustrating a high share of negation

5.7.7 Verbs and other lexical-grammatical phenomena [GEFL TC]

Under the heading “other lexical-grammatical phenomena” I subsumed three types of constructions that may co-occur with progressive forms and that were in fact found to do so more or less frequently in spoken BrNSE: questions, if-clauses, and relative clauses. While in GEFL TC progressives are rather common in interrogative contexts, they are extremely rare in if-clauses and relative clauses in both EFL textbook subcorpora (cf. Figure 115 in 5.5.7). In the following, I will thus only deal with the relations between the progressive forms of selected verbs and questions.

Table 42 provides information on the distribution of progressive form tokens of individual verbs across interrogative and non-interrogative contexts. If we just focus on the left-hand grey-shaded parts of the displayed mini-diagrams, we notice some significant differences among verbs with respect to their shares of questions. Highest percentages can be found for *doing* in GLN and EG 2000 (see concordance extracts in Figure 131), for *asking*, *going*, *having*, and *playing* in GLN, and for *listening*, *saying*, *staying*, and *thinking* in EG 2000. The most common pattern that these items are used in in interrogative contexts in the coursebook texts is “are you V-ing”, as the concordance samples in Figures 131 and 132 nicely illustrate. A few other verbs also occur in progressives with an interrogative function (e.g. *coming*, *looking*, *telling*) but are much less frequent. The frequencies of occurrence of different progressives in questions found for spoken English and textbook English data will be compared in Section 6.3 below.

5.7.8 Verbs and adverbial specification [GEFL TC]

A final feature that was included in the contextual analysis of progressives was adverbial specification. Average shares of co-occurrences of GLN and EG 2000 progressives and different types of adverbials have been determined in 5.5.8 (see Figure 116 for a graphical display). Percentages of individual progressive types are listed, and visualised, in Table 43.

Table 42. Shares of progressive forms of selected verbs in interrogative and non-interrogative contexts in GLN (G) and EG 2000 (E, values in *italics*)




















































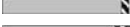



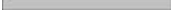


















verb form	G or E	questions	no questions	distribution graphically illustrated	
		<i>ø</i> 20.20% <i>ø</i> 20.76%	<i>ø</i> 79.80% <i>ø</i> 79.24%	■ questions	▨ no questions
asking	G	33.33%	66.67%		
	E	<i>0.00%</i>	<i>100.00%</i>		
coming	G	9.52%	90.48%		
	E	<i>0.00%</i>	<i>100.00%</i>		
doing	G	78.95%	21.05%		
	E	<i>75.00%</i>	<i>25.00%</i>		
getting	G	0.00%	100.00%		
	E	<i>6.25%</i>	<i>93.75%</i>		
going	G	26.37%	73.63%		
	E	<i>22.92%</i>	<i>77.08%</i>		
having	G	40.00%	60.00%		
	E	<i>0.00%</i>	<i>100.00%</i>		
helping	G	0.00%	100.00%		
	E	<i>16.67%</i>	<i>83.33%</i>		
listening	G	0.00%	100.00%		
	E	<i>50.00%</i>	<i>50.00%</i>		
looking	G	5.00%	95.00%		
	E	<i>4.76%</i>	<i>95.24%</i>		
making	G	8.33%	91.67%		
	E	<i>0.00%</i>	<i>100.00%</i>		
playing	G	44.44%	55.56%		
	E	<i>21.43%</i>	<i>78.57%</i>		
reading	G	0.00%	100.00%		
	E	<i>10.00%</i>	<i>90.00%</i>		
saying	G	16.67%	83.33%		
	E	<i>60.00%</i>	<i>40.00%</i>		
sitting	G	0.00%	100.00%		
	E	<i>6.67%</i>	<i>93.33%</i>		
staying	G	20.00%	80.00%		
	E	<i>50.00%</i>	<i>50.00%</i>		
taking	G	0.00%	100.00%		
	E	<i>12.50%</i>	<i>87.50%</i>		
talking	G	0.00%	100.00%		
	E	<i>20.83%</i>	<i>79.17%</i>		
telling	G	0.00%	100.00%		
	E	<i>7.69%</i>	<i>92.31%</i>		
thinking	G	16.67%	83.33%		
	E	<i>50.00%</i>	<i>50.00%</i>		
trying	G	12.50%	87.50%		
	E	<i>0.00%</i>	<i>100.00%</i>		
wearing	G	0.00%	100.00%		
	E	<i>16.67%</i>	<i>83.33%</i>		
working	G	14.29%	85.71%		
	E	<i>0.00%</i>	<i>100.00%</i>		

Table 43. Shares of different types of adverbial specification of selected verbs in GLN (G) and EG 2000 (E, values in *italics*)

verb form	G or E	time adverbial	place adverbial	other adverbial	distribution graphically illustrated ■ time adv. ▨ place adv. □ other adv.
		ø32.57% <i>ø23.54%</i>	ø2.93% <i>ø8.61%</i>	ø3.91% <i>ø3.54%</i>	
asking	G	66.67%	0.00%	0.00%	
	E	<i>40.00%</i>	<i>20.00%</i>	<i>20.00%</i>	
coming	G	33.33%	0.00%	4.76%	
	E	<i>14.29%</i>	<i>0.00%</i>	<i>0.00%</i>	
doing	G	21.05%	21.05%	0.00%	
	E	<i>12.50%</i>	<i>12.50%</i>	<i>3.13%</i>	
getting	G	37.50%	0.00%	0.00%	
	E	<i>18.75%</i>	<i>0.00%</i>	<i>0.00%</i>	
going	G	35.16%	2.20%	3.30%	
	E	<i>32.29%</i>	<i>6.25%</i>	<i>1.04%</i>	
having	G	20.00%	0.00%	0.00%	
	E	<i>42.86%</i>	<i>0.00%</i>	<i>0.00%</i>	
helping	G	50.00%	0.00%	0.00%	
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	—
listening	G	25.00%	0.00%	12.50%	
	E	<i>16.67%</i>	<i>0.00%</i>	<i>16.67%</i>	
looking	G	10.00%	0.00%	0.00%	
	E	<i>28.57%</i>	<i>4.76%</i>	<i>4.76%</i>	
making	G	16.67%	0.00%	0.00%	
	E	<i>11.11%</i>	<i>0.00%</i>	<i>0.00%</i>	
playing	G	55.56%	0.00%	0.00%	
	E	<i>42.86%</i>	<i>7.14%</i>	<i>7.14%</i>	
reading	G	25.00%	0.00%	0.00%	
	E	<i>10.00%</i>	<i>0.00%</i>	<i>0.00%</i>	
saying	G	0.00%	0.00%	0.00%	—
	E	<i>40.00%</i>	<i>0.00%</i>	<i>0.00%</i>	
sitting	G	14.29%	0.00%	0.00%	
	E	<i>26.67%</i>	<i>60.00%</i>	<i>0.00%</i>	
staying	G	20.00%	0.00%	0.00%	
	E	<i>0.00%</i>	<i>75.00%</i>	<i>0.00%</i>	
taking	G	0.00%	0.00%	0.00%	—
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	—
talking	G	30.00%	0.00%	0.00%	
	E	<i>20.83%</i>	<i>8.33%</i>	<i>8.33%</i>	
telling	G	33.33%	0.00%	0.00%	
	E	<i>7.69%</i>	<i>0.00%</i>	<i>0.00%</i>	
thinking	G	50.00%	0.00%	16.67%	
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	—
trying	G	62.50%	0.00%	0.00%	
	E	<i>20.00%</i>	<i>10.00%</i>	<i>10.00%</i>	
wearing	G	0.00%	0.00%	0.00%	—
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	—
working	G	57.14%	0.00%	42.86%	
	E	<i>16.67%</i>	<i>8.33%</i>	<i>0.00%</i>	

David: What are you doing, Mark? Mark: I'm drawing a birthday picture in the kitchen. Becky: What are you doing? Robert: I'm making a salad. Simon: I yes and ran after her. "What are you doing here?" I asked. "I decided it was time for hockey stick, too. What are you doing tonight, Jenny? I'm going to the youth Excuse me," Becky said. "What are you doing?" "Me?" the man said. He stood up, but behind the prisoner. "What are you doing with that rope?" shouted the Sheriff. "his glider. How you doing?" the kid asked him. "Fine," Steve said. "What are you doing," he said. "Listen, I'm sorry straight and your animals. See what they're doing. Are they sleeping? Or eating? Or playing? They're building a tree house with old doing?" The driver went faster and faster, doing? MRS SNOW She's sleeping, too. Listen ng Latin.' 'Yes, but what's a Roman doing here?' 'We Romans were here in Deva 20 urs! And what are you and Mr Martin doing, Miss Hunt?' 'Well, we've just been to o explain to people that Dad's just doing his job. That it's what a bank manager go ahead ... (seeing what Rust is doing) Look, please don't play with that wa ch my new video. Not now, Nick. I'm doing my homework. Ask me tomorrow. But dad' RUST I don't know. MANNY What's he doing? RUST I don't know. MANNY What you d ge. 'Ben, what are you and that girl doing?' she shouted. 'Jeremy, dear, some heri doing lying on the ground there? Having a : RUST That's terrible. What's he been doing then, girls? Flashing?' TED (angry) : if you saw it in a magazine and are doing what others do, how does it make you day with a crowd of new friends and doing what I want to do. So, hang in there,

Figure 131. Extracts from GLN and EG 2000 concordances of *doing*, illustrating high shares of questions

Robert: Are you having a good birthday, Jenny? Jenny: Well, n. They all go back to bed. "Are you having a good holiday?" asks Mrs Duncan. Ve The Taylors were having tea with Christoph. Darren felt disa turned round and smiled at they were having breakfast. Sarah. "Oh, Sarah. Don't ing after Tim. They are laughing and having fun.) 265 t happens to be my fiancée you're talking to. JOE The wrong time? What are you talking about? MERLIN Is this the year 500? thing like it. TED What are you talking about? I've done nothing ... RUST Y s? Swimming-pool? TV? What are you talking about? NAZMA Haven't you got compute TED I never broke her arm. What you talking about? Let go of me ... BEN (in alar e reading. Ben and his grandma were talking. 'This is boring. Let's go and buy s on the new vocabulary. Others were talking in Urdu. 'Ladies,' Mrs Mir began, 't

Figure 132. Extracts from a GLN concordance of *having* and an EG 2000 concordance of *talking*, illustrating high shares of questions

As the mini-diagrams included in the table show, there is a lot of inter-verb and inter-textbook-corpus variation. For many verbs highest shares are found for the co-selection of progressives and time adverbials, though, in a number of cases, above-average values only relate either to the GLN or the EG 2000 datasets. Time adverbials are particularly frequent in progressives with *asking*, *going*, and *playing* in GLN and EG 2000, with *getting*, *helping*, *thinking*, *trying*, and *working* in GLN and with *having*, *saying*, and *sitting* in EG 2000. Examples of this type of adverbial specification are given in (212) to (215).

- (212) My brother Hugh (15) and I are at a new school. We've been going there *since September*. (GLN)
- (213) The pop industry has only one aim - our money. *At the moment* I'm trying to save up for some new music, but I'm not making much progress. (GLN)
- (214) CLARK: It's his wife, remember? She's having her baby *today*. (EG 2000)
- (215) 'Very well. Australians are less serious, more easy-going. They're *always* saying "No worries, mate. She'll be right." I like that attitude.' (EG 2000)

For the two other types of adverbial specification, with place adverbials and “other” adverbials, the determined verb-individual shares are generally much lower. However, in EG 2000 the verb forms *asking*, *doing*, *sitting*, and *staying* clearly show above-average percentages of place adverbial specification (cf. (216) and (217) for examples). Place adverbials very rarely occur in GLN progressives. Only the share for *doing* is comparatively high. The picture is rather similar with respect to the so-called “other” adverbials, i.e. those items that do not belong to the “time” or “place” group. Only a small number of verb forms are found to repeatedly co-occur with adverbials such as *actually*, *only*, *if* (in the “whether” sense), or *really*. Among these verbs count *asking*, *listening*, *thinking*, and *working*. GEFL TC concordance examples which illustrate this collocation of particular progressives and “other” adverbials are given in (218) to (220).

(216) ‘Hello, Sita. What are you doing here?’ (EG 2000)

(217) Oh God, he’s sitting *over there*. (EG 2000)

(218) You ... you’re *really* asking for it. (EG 2000)

(219) It is a slow job, but he is working *hard* - much harder than his friends. (GLN)

(220) Then, after looking around to see *if* anyone was listening, she went on (EG 2000)

Again, the results obtained in the analysis of progressives and adverbial specification hint at certain co-selectional preferences of individual verbs and indicate that there is quite some lexical-grammatical variation between the two examined coursebook series. In the next three sections we will see whether similar differences between verbs, and between GLN and EG 2000 can also be found with respect to time references, central functions, and additional functions of progressives.

5.7.9 Verbs and time reference [GEFL TC]

The average distributions of different types of time reference determined for the two sets of GLN and EG 2000 progressives (cf. Figure 119 in 5.6.1) showed that most of the analysed examples from both subcorpora refer to present-time actions or events. This is in fact true for a large number of items included in the verb-specific analysis (cf. the percentages in the “present time ref.” column in Table 44).

Some verbs, however, show particularly high shares of present time reference which deviate a lot from the average values of 35.50% and 43.54%. Worth mentioning in this context are *asking*, *doing*, *helping*, *listening*, *looking*, *making*, *reading*, and *staying*. There are a couple of differences between GLN and EG 2000 concerning the shares of present time reference but they are not as significant as with the three remaining time reference types. For instance, many verb forms which show high percentages in the “past time ref.” column in Table 44 for GLN, never or very rarely refer to events in the past in EG 2000, and vice versa. It is thus rather difficult to find a pattern here and determine clearly past-time-favouring verbs. The only “clear” cases are *having*, *sit-*

Table 44. Time reference distribution for selected verbs in GLN (G) and EG 2000 (E, values in italics)

verb form	G or E	past time ref.	present time ref.	future time ref.	pres/fut time ref.	distribution graphically illustrated
		<i>ø</i> 29.32% <i>ø</i> 22.03%	<i>ø</i> 35.50% <i>ø</i> 43.54%	<i>ø</i> 31.60% <i>ø</i> 31.65%	<i>ø</i> 3.58% <i>ø</i> 2.78%	<div>■ past t. ref.</div> <div>▨ present t. ref.</div> <div>□ future t. ref.</div> <div>■ present/future t. ref.</div>
asking	G	0.00%	100.0%	0.00%	0.00%	
	E	20.00%	80.00%	0.00%	0.00%	
coming	G	28.57%	28.57%	38.10%	4.76%	
	E	0.00%	71.43%	28.57%	0.00%	
doing	G	10.53%	68.42%	10.53%	10.53%	
	E	15.63%	75.00%	6.25%	3.13%	
getting	G	37.50%	50.00%	0.00%	12.50%	
	E	31.25%	37.50%	31.25%	0.00%	
going	G	9.89%	6.59%	82.42%	1.10%	
	E	2.08%	4.17%	91.67%	2.08%	
having	G	40.00%	60.00%	0.00%	0.00%	
	E	42.86%	28.57%	28.57%	0.00%	
helping	G	50.00%	50.00%	0.00%	0.00%	
	E	0.00%	100.0%	0.00%	0.00%	
listening	G	25.00%	75.00%	0.00%	0.00%	
	E	33.33%	66.67%	0.00%	0.00%	
looking	G	40.00%	60.00%	0.00%	0.00%	
	E	23.81%	71.43%	4.76%	0.00%	
making	G	25.00%	58.33%	8.33%	8.33%	
	E	11.11%	66.67%	22.22%	0.00%	
playing	G	0.00%	55.56%	44.44%	0.00%	
	E	14.29%	64.29%	21.43%	0.00%	
reading	G	25.00%	75.00%	0.00%	0.00%	
	E	20.00%	70.00%	10.00%	0.00%	
saying	G	50.00%	16.67%	0.00%	33.33%	
	E	0.00%	60.00%	0.00%	40.00%	
sitting	G	42.86%	42.86%	0.00%	14.29%	
	E	73.33%	26.67%	0.00%	0.00%	
staying	G	20.00%	80.00%	0.00%	0.00%	
	E	0.00%	50.00%	50.00%	0.00%	
taking	G	40.00%	20.00%	40.00%	0.00%	
	E	25.00%	37.50%	37.50%	0.00%	
talking	G	40.00%	60.00%	0.00%	0.00%	
	E	50.00%	50.00%	0.00%	0.00%	
telling	G	66.67%	33.33%	0.00%	0.00%	
	E	7.69%	84.62%	7.69%	0.00%	
thinking	G	83.33%	16.67%	0.00%	0.00%	
	E	50.00%	50.00%	0.00%	0.00%	
trying	G	25.00%	75.00%	0.00%	0.00%	
	E	60.00%	40.00%	0.00%	0.00%	
wearing	G	50.00%	50.00%	0.00%	0.00%	
	E	16.67%	83.33%	0.00%	0.00%	
working	G	42.86%	14.29%	28.57%	14.29%	
	E	41.67%	50.00%	0.00%	8.33%	

mean, those lessons aren't cheap. I'm going to write to the head. It's no use ring I hope they aren't too expensive. I'm going to have a ride on the ghost train! Hot t there, Becky? Becky: Fish food. I'm going to feed the fish in our pond. David: C a supermarket not far from here. I'm going that way, too. Come with me, and I ca hat are you doing tonight, Jenny? I'm going to the youth club and I don't really w at Jenny's new house. Mr Leinert: I'm going to take Jenny to 'The Tales of Robin F ne, OK?... If my parents find out I'm going out with him, they'll go crazy! It was the sports centre at 10 o'clock, I'm going to see him. Do you want to come, too? it is the Hill of Tara,' I said. I'm going to read about it,' she said. I don't k ht Bob would see the truth. 'Kay, I'm going and I'm not coming back.' 'Well, leave Becky: Hi, Robert! We're just going to our house. We've got an idea for th ll I arrange it for you? No, I am not going to give you your money back! OK, yes, l country, but I can tell you I'm not going there until I can speak the language. than going back to Eburacum.' 'We're going to Luguvalium,' she told Juvena and t reet. 'Now,' she said to her, 'we're going to walk up Northgate Street and then ning now?' asked Mrs McCann. 'We're going to the museum now, Jenny said. Mrs M ! Soon Debbie was tired again. 'I'm going back to bed!' she said and went upsta plan to get there?' Glen asked. 'I'm going to fly. What else?' 'What about the t sons on Saturdays,' said Emma. 'I'm going to watch the football match on Saturd , Debbie went home with Jenny. 'I'm going to sleep at Jenny's house tonight,' s morning. At 12.30 Debbie said, 'I'm going to eat my sandwiches now.' 'Me too,' t. 'Well, you two,' Sally said, 'I'm going to Alison's party. Oh ... don't forge trying to get back to the wall. 'I'm going to fall!' he screamed. 'You'll be all tin didn't. 'Oh no,' she said. 'He's going to sing now, and it's all your fault e as she speaks to Lynn Kid, you're going to have your picture up there ... Po

Figure 133. Extracts from GLN and EG 2000 concordances of *going*, illustrating high shares of future time reference

ting, talking, thinking, and working, as they have high above-average shares of past time reference in both GEFL TC subcorpora.

With respect to individual verbs and future time reference, there is also some inter-corpus variation, though the differences across verbs are even more significant. While verbs such as *coming*, *going*, *playing*, *staying*, and *taking* show values of up to 91 per cent, many other forms are never used in future time reference contexts. *Going* is probably the most obvious case. The so-called “*going to future*” (cf. also 4.5.9) is very frequently used in both coursebook series, as the concordance extracts in Figure 133 illustrate, and this is certainly the main reason for the extremely high percentages in the “future time ref.” column. Like *going* in the “*going to future*” examples, some of the other future-time-favouring verb forms, *coming*, *playing*, *staying*, and *talking*, also express intentions of speakers to do something at a later point in time but, compared to the *going* examples, these intentions appear to be more definite and fixed, rather like plans (cf. the examples in (221) to (223)).

- (221) Alex: Great that you're *coming* to visit us for the summer, Daniel! (GLN)
- (222) I've never been to New York City but I'm going next week. I'm *staying* with my pen-friend in Manhattan. (EG 2000)
- (223) NADJA: Niels is cooking Sunday lunch this weekend. And I'm *playing* volleyball. (EG 2000)

Finally, the distribution of “indeterminate” time reference cases, i.e. such instances which refer to the present and/or future, is rather unequal across verbs and restricted to very few forms. Only *doing*, *saying*, and *working* show above-average shares of progressives with present/future time reference, but then again the absolute numbers of occurrence are still rather small. This means that in GEFL TC just a few progressives refer to something that can be valid in general, at the time of speaking and at a later

point in time. How these time reference findings compare to the results from the spoken English analysis will be seen later in Chapter 6. Let us now look at some relations between progressive functions and selected verb forms.

5.7.10 Verbs and central functions of the progressive [GEFL TC]

Having identified two central function features of progressives in the spoken BrNSE datasets (continuousness and repeatedness), I specified the distributions of these features and of their four possible combinations in the coursebook English corpora and found that one feature combination, namely “continuous + non-repeated”, is by far the most frequent of the four in GLN and EG 2000.

So far, we have only observed that non-continuousness and repeatedness are marginalised in the textbooks since the feature combinations “continuous + repeated”, “non-continuous + repeated”, and “non-continuous + non-repeated” show rather low shares in the GLN and EG 2000 progressives, but we have not looked further into these less frequent feature combinations and their lexical realisations. I would now like to analyse which verbs, or to be more precise, the progressives of which verbs express these feature combinations. Also, we will see in our verb-specific approach whether there is a lot of variation among verbs concerning the shares of “continuous + non-repeated” progressives or whether the distribution displayed in Figure 123 can be called representative of most of the items analysed.







Table 45 lists the percentages of the four mentioned feature combinations in GLN and EG 2000 for our 22 selected verbs. The bar charts included in the table serve to illustrate the verb-specific distributions. A first thing that strikes us is that, except for *asking* and *looking*, all verbs exclusively refer to continuous actions or events. Besides, for a large number of verbs the percentage of “continuous + non-repeated” progressives even amounts to 100 per cent (e.g. *having*, *listening*, *reading*, *staying*). This means that, at least with respect to this most frequent feature combination, the distribution shown in Figure 123 certainly is representative of most verbs. However, there is quite some variation with respect to the expression of “continuous + repeated” actions or events. Some of the selected verbs, e.g. *doing*, *getting*, *saying*, and *working*, show highly above-average shares of this feature combination, while for many others the share is 0% (e.g. for *having*, *helping*, *listening*, *looking*, and *making*, to mention only a few forms). Also, some of the verbs (*playing*, *taking*, *talking*) only express continuousness and repeatedness in EG 2000, others only do so in GLN (*coming*, *telling*, *trying*).

The picture is rather simple when it comes to the two non-continuous feature combinations. Of the 22 most frequent progressive forms in GEFL TC only *asking* is used to express non-continuousness and repeatedness, as in example (224). In fact, as was true for the BNC_spoken and BoE_brspeak data (cf. Table 29), *asking* always refers to non-continuous actions or events in GEFL TC; in GLN non-continuousness is more frequently combined with repeatedness, in EG 2000 predominantly with non-repeatedness. The only other form that occurs in non-continuous and non-repeated contexts, rather rarely though, is *looking* (cf. example (225)).

Table 45. The distribution of verb form tokens across the four combinations of the central function features “continuously” and “repeatedness” for selected verbs in GLN (G) and EG 2000 (E, values in italics)

verb form	G or E	cont. + repeated	cont. + non- repeated	non- cont. + repeated	non- cont. + non-rep.	distribution graphically illustrated
						<div><div></div> cont.+rep. <div></div> cont.+non-rep. <div></div> non-cont.+rep. <div></div> non-cont+non-rep.</div>
		ø7.82% ø9.11%	ø88.62% ø88.35%	ø1.30% ø0.76%	ø3.26% ø1.77%	
asking	G	0.00%	0.00%	66.67%	33.33%	<div><div></div><div></div></div>
	E	0.00%	0.00%	20.00%	80.00%	<div><div></div><div></div></div>
coming	G	9.52%	90.48%	0.00%	0.00%	<div><div></div><div></div></div>
	E	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
doing	G	15.79%	84.21%	0.00%	0.00%	<div><div></div><div></div></div>
	E	6.25%	93.75%	0.00%	0.00%	<div><div></div><div></div></div>
getting	G	25.00%	75.00%	0.00%	0.00%	<div><div></div><div></div></div>
	E	25.00%	75.00%	0.00%	0.00%	<div><div></div><div></div></div>
going	G	5.49%	94.51%	0.00%	0.00%	<div><div></div><div></div></div>
	E	8.33%	91.67%	0.00%	0.00%	<div><div></div><div></div></div>
having	G	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
	E	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
helping	G	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
	E	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
listening	G	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
	E	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
looking	G	0.00%	95.00%	0.00%	5.00%	<div><div></div><div></div></div>
	E	0.00%	95.24%	0.00%	4.76%	<div><div></div><div></div></div>
making	G	0.00%	91.67%	0.00%	8.33%	<div><div></div><div></div></div>
	E	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
playing	G	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
	E	14.29%	85.71%	0.00%	0.00%	<div><div></div><div></div></div>
reading	G	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
	E	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
saying	G	33.33%	66.67%	0.00%	0.00%	<div><div></div><div></div></div>
	E	40.00%	60.00%	0.00%	0.00%	<div><div></div><div></div></div>
sitting	G	14.29%	85.71%	0.00%	0.00%	<div><div></div><div></div></div>
	E	6.67%	93.33%	0.00%	0.00%	<div><div></div><div></div></div>
staying	G	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
	E	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
taking	G	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
	E	12.50%	87.50%	0.00%	0.00%	<div><div></div><div></div></div>
talking	G	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
	E	8.33%	91.67%	0.00%	0.00%	<div><div></div><div></div></div>
telling	G	33.33%	66.67%	0.00%	0.00%	<div><div></div><div></div></div>
	E	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
thinking	G	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>
	E	0.00%	100.0%	0.00%	0.00%	<div><div></div><div></div></div>

Table 45. (continued)

trying	G	12.50%	87.50%	0.00%	0.00%	
	E	0.00%	100.0%	0.00%	0.00%	
wearing	G	0.00%	100.0%	0.00%	0.00%	
	E	0.00%	100.0%	0.00%	0.00%	
working	G	42.86%	57.14%	0.00%	0.00%	
	E	50.00%	50.00%	0.00%	0.00%	

- (224) Guy Fawkes: You're *asking* questions again. Let's just say we want them to keep quiet. (GLN)
- (225) Deirdre Fitzgerald *was looking* up at Kate standing in front of the librarian's desk (EG 2000)

On the whole, we get less variation in the distribution of these central function feature combinations in GEFL TC than we found for most of the examined context phenomena. A larger number of verb-specific preferences was only observed for the combination “continuous + repeated”. These findings certainly support what we said earlier on, namely that, in the examined textbooks, most progressives only have one central function: the reference to continuous and non-repeated actions or events.

5.7.11 Verbs and additional functions of the progressive [GEFL TC]

We noted above that all additional functions that were found in the BNC/BoE progressives are rather infrequent in the GEFL TC datasets. A detailed analysis of the potential relations between individual verbs and additional functions may hence appear redundant. However, to enable a systematic comparison of spoken English and coursebook English findings, I decided to carry out the same kind of verb-specific analysis as I did for the BNC/BoE data. Given the significant inter-verb differences found for progressives and their functions in spoken English, I anticipated similar results for GLN and EG 2000. Besides, I have not made any statements about existing connections between verbs and additional functions so far.

The frequencies with which the 22 most common verbs in GEFL TC refer to the identified additional functions are listed in Table 46 (above-average shares have been shaded grey in the table). There is not much to say about the particularly rare functions “politeness or softening”, “gradual change and development”, and “old and new habits”. Politeness is only expressed by *thinking* in GLN (cf. example (226)). The only two “gradual change verbs” are *getting* and *coming* (GLN only; cf. example (227)), and just *going* and *staying* (both in GLN) refer to (temporary) habitual actions (cf. example (228)).

- (226) “Actually, I *was thinking* we should just pretend it never happened. It's not worth ruining a friendship over.” (GLN)

Table 46. Frequencies of occurrence of additional progressive functions across verb forms in GLN (G) and EG 2000 (E, values in *italics*)

verb form	G or E	general validity <i>ø3.58%</i> <i>ø5.82%</i>	politeness/ softening <i>ø1.63%</i> <i>ø0.76%</i>	emphasis/ shock <i>ø3.26%</i> <i>ø5.31%</i>	gradual change <i>ø1.63%</i> <i>ø1.77%</i>	old/new habits <i>ø0.98%</i> <i>ø0.00%</i>	framing <i>ø2.61%</i> <i>ø4.30%</i>
asking	G	0.00%	0.00%	66.67%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>60.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
coming	G	4.76%	0.00%	0.00%	9.52%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
doing	G	10.53%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>3.13%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
getting	G	12.50%	0.00%	0.00%	25.00%	0.00%	0.00%
	E	<i>18.75%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>37.50%</i>	<i>0.00%</i>	<i>12.50%</i>
going	G	0.00%	0.00%	1.10%	0.00%	2.20%	1.10%
	E	<i>4.17%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
having	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
helping	G	0.00%	0.00%	0.00%	0.00%	0.00%	50.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
listening	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>16.67%</i>
looking	G	0.00%	0.00%	0.00%	0.00%	0.00%	5.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>4.76%</i>
making	G	0.00%	0.00%	25.00%	0.00%	0.00%	0.00%
	E	<i>11.11%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
playing	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>7.14%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
reading	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>10.00%</i>
saying	G	33.33%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>40.00%</i>	<i>0.00%</i>	<i>100.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
sitting	G	14.29%	0.00%	0.00%	0.00%	0.00%	28.57%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>20.00%</i>
staying	G	0.00%	0.00%	0.00%	0.00%	20.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
taking	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>25.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>12.50%</i>
talking	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>8.33%</i>	<i>0.00%</i>	<i>16.67%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>4.17%</i>
telling	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>7.69%</i>	<i>0.00%</i>	<i>23.08%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
thinking	G	0.00%	33.33%	16.67%	0.00%	0.00%	16.67%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>50.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
trying	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>10.00%</i>
wearing	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>
working	G	28.57%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	<i>8.33%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>0.00%</i>	<i>8.33%</i>

(227) Bread is toasted and your favourite music fills the air. The intelligent home *is coming* to life. (GLN)

(228) “*Are you staying* at Tulloch House?” (GLN)

The additional functions “general validity”, “emphasis, attitude, or shock”, and “framing” are expressed by a larger number of different verbs, some of which show rather high percentages. *Getting*, *saying*, and *working*, for instance, are relatively common in general validity contexts, such as those exemplified in (229) and (230). Above-average shares of this additional function are also found with *doing* and *sitting* in GLN and with *making*, *playing*, *talking*, and *telling* in EG 2000. In the EG 2000 concordance data, *saying* not only commonly expresses “general validity”, all occurrences of the form also have an emphatic function (cf. example (230)). Other “emphasis or attitude” verbs are *asking*, *making*, *taking*, *telling*, and *thinking* (cf. example (231)), which all show very low token numbers. The largest number of verbs can be found to occur in examples of the “framing” kind as described in 4.4.5 and 5.6.4 above. This function appears to be rather widely spread across verb types. Particularly high shares of framing contexts are found for *helping*, *sitting*, and *thinking* in GLN and for *getting*, *listening*, *reading*, *sitting*, *taking*, and *trying* in EG 2000 (cf. (232) and (233) for typical textbook examples).

(229) MARIA: There’s a crowd of tired women who *are living* from day to day. (EG 2000)

(230) They’re always *saying* “No worries, mate. She’ll be right.” (EG 2000)

(231) TED: What *are you talking* about? I’ve done nothing ... (EG 2000)

(232) While he *was helping* the mother, a mountain of water broke over them both. (GLN)

(233) ‘*Weren’t you listening* when they called our flight?’ (EG 2000)

The just described types of repeated co-selection of individual verbs and additional functions are not supposed to convey the wrong impression that these functions are particularly well-represented in the coursebooks. To stress this fact again, the overall shares are very low for all additional functions (cf. Figure 124), and the foremost function of progressives in GLN and EG 2000 is to refer to single continued events, but there certainly is some intra-verb variation in this field which must not be neglected. Before taking a look at the order in which progressive forms and their functions are introduced in the textbooks, I will now briefly summarise the findings of the GEFL TC “verbs and progressives” analysis.

5.7.12 Summary of the findings [GEFL TC – verbs and progressives]

In the previous sections we looked into some important relations between individual verbs and typical contexts and functions of progressives in the EFL textbook corpus data. Among the co-selection types we analysed were that of verbs and different tense forms (PastProg, PresProg, etc.), verbs and subjects, verbs and negation, verbs and

interrogative contexts, verbs and adverbial specification, verbs and time reference, and verbs and central and additional progressive functions.

The results are rather similar for almost all examined phenomena. We found a lot of distributional variation across verbs and across corpora (GLN vs. EG 2000). This means that, for example, the determined average shares of time reference types of the progressive datasets (cf. 5.6.1) are not representative of the majority of verbs, but that individual verbs show largely different shares and may have a preference for one particular type of time reference. It also means that, for the same verb, percentages may deviate substantially between GLN and EG 2000. With all context features included in the verb-specific analysis we observed certain collocational preferences of individual verbs and deviations from the average values. An example was the selection of personal pronoun subjects which differed strongly from verb form to verb form (cf. 5.7.3). Only with the central function feature combination “continuous + non-repeated” we found that the very high shares determined for the entire sets of GLN and EG 2000 progressives are representative of most verbs (cf. 5.7.10). There is, however, quite some inter-verb (and inter-corpus) variation with respect to the other repeatedly occurring feature combination “continuous + repeated”, and for the identified additional functions (cf. 5.7.10 and 5.7.11).

Given these findings and to answer the question posed in the headline of 5.7 “How lexical is EFL textbook grammar?”, we can say that, like the grammar of spoken English (cf. 4.5), the grammar in the textbooks, or at least one important part of it, shows a considerable degree of lexicalness too. However, in the GEFL TC subcorpora, lexical-grammatical patterns mainly occur with the investigated context phenomena and less significantly with the function features. We will see in Chapter 6 in how far the two analysed types of English, spoken BrNSE and “school” English, differ regarding their lexicalness of the grammatical construction “progressive”, i.e. it will be analysed whether the most typical lexical-grammatical patterns determined for the BNC_spoken and BoE_brspok data could also be found in GLN and EG 2000, or whether the verb-progressive relations that surfaced in the coursebooks differ strongly from those in real spoken BrNSE.

5.8 Progressive progression (?) – When and how are progressives introduced in the textbooks?

To account for the use of progressives in an English language teaching context, I have treated a large number of texts from twelve coursebooks that are widely used in the German EFL classroom as a corpus. On the basis of this corpus, GEFL TC, I have carried out a detailed analysis of progressive forms, functions, and contexts. This analysis of progressives in context will now be complemented by an examination of the material we find in the coursebooks in addition to the texts which form GEFL TC, i.e. exercises, grammar sections and so-called “Language Summaries” in the GLN and EG 2000 volumes, and the grammar books and booklets that are used alongside the text-

books. I am interested in the grammatical progression in the two series and in the amount and kind of explicit information that the coursebook authors provide on the use of progressives in English.

In the following sections we will see where in the six volumes of GLN and EG 2000 the different progressive forms are introduced, and describe at what stage in the courses we find information about their functions and preferred contexts. In 5.8.3 the findings based on both coursebook series will be summarised and discussed in the light of some of the results obtained in the BNC/BoE data analysis.

5.8.1 The progression in *Learning English Green Line New* (volumes 1–6)

As was to be expected, the first progressive tense form that is introduced in volume 1 of the *Green Line New* series (Aston et al. 1995) is the present progressive (PresProg). PresProg forms are first used in Step B of Unit 4 in a short text which describes what four children shown in two pictures are occupied with, and in a dialogue between three friends. Part of this dialogue is given in (234).

- (234) Becky: What *are* you *doing*?
 Robert: I'm *making* a salad.
 Simon: And you're *listening* to Radio 1.
 Robert: Well, no. I'm not *listening* to Radio 1. It's Radio Nottingham. I'm *listening* to my mum.

In example (234) we find contracted and non-contracted, negated and non-negated PresProgs. The two short texts are backed up by five exercises to help learners practice the new forms. Grammar sections at the end of the coursebook provide information on the use of the grammatical structures introduced in the GLN units. We learn there that the PresProg is used to express that someone is in the process of doing something (e.g. *She's reading*.) or to refer to an ongoing action (e.g. *It's raining*. cf. Aston et al. 1995: 136).

What is perhaps not expected, particularly if we remember what we said earlier about the general frequencies of occurrence of progressive and simple tense forms in spoken and written English (cf. 3.5.1), is the introduction of the simple present in Unit 6, Step B of GLN, vol. 1, i.e. two units later than its progressive counterpart. This common tense form sequence problem has been discussed by a number of scholars. Mohan (1976:41), for instance, asks the question “Which tense should beginners of English learn first – the present simple or the present continuous?” and notes that “teachers normally teach the continuous tense first because the course they are using introduces it first.”, which is in fact what we found in *Green Line New*. Pürschel (1981:89), Scheffer (1975:111), and Turner (1981:194) also comment on the stage of introduction of progressive and non-progressive tense forms. They all question or criticise this commonly used sequence, and Turner (1981:196) mentions “frequency of occurrence in [...] actual usage” as an argument for a different “non-progressive first” progression.

It is indeed questionable whether the textbooks have got their priorities right concerning the introduction of tense forms. Textbook authors might argue, from a contrastive perspective, that learners ought to be presented with the progressive at a very early stage in the course, since this form is new to German native-speakers and “typically English”. However, this early focus on a new “typically English” construction may include the danger of overuse of the form in learner production, which is what Johansson and Stavestrand (1987) and Virtanen (1997) found in their analyses of progressives in Finnish, Norwegian, and Swedish learner English. Another counter-argument against introducing the progressive before the non-progressive form is the complexity of the former construction. Forms like “I go” and “she writes” should be less difficult for beginners to learn than “I am going” and “she is writing”. Based on these considerations and on the findings of corpus-based studies that have determined the frequencies of progressive and simple forms in English (esp. Biber et al. 1999), I would hence suggest to change the sequence of tense form introduction in the coursebooks and present the simple present earlier in the course than the present progressive.

In Unit 7, Step B of GLN volume 1, the two introduced forms (PresProg and simple present) and their functions are contrasted with the help of one short text and two exercises. A paragraph in the grammar section repeats their basic functions and provides information on collocations with time adverbials. We learn that PresProgs are often specified by adverbials such as *today*, *now*, *just now*, or *this morning*, while *always*, *often*, *sometimes*, or *every day* are typical collocates of the simple present (cf. also *GrammatischesBeiheft* 1 + 2, Lampater & Pasch 1997: 27). Unit 10, Step B introduces the so-called “going to future” in one short text and four exercises as a form that is used to talk about intentions or planned actions.

GLN volume 2 starts with a brief revision of the *going to* future in Unit 1, Step C and a review of the simple present vs. PresProg contrast in Unit 3, Step B, before it moves on to introducing the next progressive form type, the past progressive (PastProg, Unit 9, Step B). Examples of the PastProg are given in two short narrative texts which mainly consist of dialogue (see (245) for an extract). The form is then practiced in four exercises and dealt with in the grammar section of the book. Concerning the function of the PastProg, the information we get is rather general (reference to ongoing actions in the past), and it seems that a focus lies on the “framing” use; and in fact all examples given in the grammar section express framing contexts. In all examples, the PastProg collocates either with *while* or with *when* (cf. also Lampater & Pasch 1997: 31).

- (245) First she filmed her dad while he *was going* into the bathroom. “Ugh, put that thing down, Sarah,” said her dad and closed the bathroom door. When Sarah went downstairs, her mother and Kim *were sitting* in the kitchen.

New in volume 3 of the GLN series are the use of the PresProg with future meaning, and the present perfect progressive (PresPerfProg). One short text and three exercises in Unit 2, Step B and one paragraph in the grammar section focus on the function

of the PresProg, usually in collocation with future time adverbials such as *tomorrow*, *next week*, or *on Tuesday*, to express fixed plans or arrangements. A picture of a diary is supposed to clarify this particular use. The PresPerfProg is first used in one short text and five exercises in Step B of Unit 3. Two brief extracts from this text are given in (246). Significant in these and other PresPerfProg examples is the co-occurrence with *since* and *for*, a collocation that is also stressed in the discussion of the form in the grammar section. As for the function of the PresPerfProg, we get to know that it serves to express durativity and refers to some action that started in the past and has not stopped yet. We find examples with the verb forms *waiting*, *raining*, *living*, and *learning* (cf. also Lampater & Pasch 1998:16).

- (246) We live in North Wales. But really we come from the South. We've only *been living* here since last August. [...] I've *been learning* Welsh for three years but I still find it very difficult!

The only section in volume 4 that explicitly deals with progressives is Unit 1, Step C which gives a revision of “the tenses”. PresProgs, PastProgs, and PresPerfProgs are used in one text and two exercises. In the related grammar section at the end of the coursebook we find an overview of tense forms but do not get any information on their functions or use. The past perfect progressive (PastPerfProg) does not appear in the text or exercises in Unit 1, but is included in the grammar section overview.

In volumes 5 and 6 of GLN, “Topics” and “Modules” replace the “Units” and “Steps”. Each topic consists of three modules, the Text Module (working with texts), the Word Module (vocabulary work), and the Language Module (grammar revision and new grammar topics). The Language Module in Topic 1 reviews the simple-progressive contrast, this time, however, not only with reference to present tense forms but past and present perfect forms too. We find some short text samples in speech bubbles (for an example see (247)) and two exercises on the topic. One of the related grammar sections focuses on the notions of stative vs. dynamic verbs. We learn that the latter (e.g. *run*, *walk*, *play*, *swim*) can be used with simple or progressive forms, while the former (e.g. *be*, *have*, *think*, *like*) do not normally occur in the progressive, only in some exceptional cases when they express an activity. Also discussed in the grammar sections of volume 5, though not in any of the Topics or Modules, are different ways of referring to the future, including the forms and functions of the *going to* future (for plans and for actions in the near future), and the PresProg with future time reference (for definite arrangements, always in collocation with a time adverbial). We find another overview here of “the tenses”, again only of forms and without any comments on contexts or functions.

- (247) A big 'Hi!' to all you kids who *are listening* out there. This week I want your opinions on the world of pop. OK – I'm *waiting* for your call right now.

Topic 1 in GLN volume 6 introduces the emphatic use of the PresProg in collocation with *always* (in one exercise) and provides another revision of different tense forms (in five exercises), including the PresProg, PastProg, and PresPerfProg. This time, we

also find information on the functions of the tense forms (cf. Table 47 for quotes from the coursebook). The PresProg with *always* is also covered in the grammar section (cf. (248) for two examples). We learn that the form can function to express a negative emotional attitude and that *always*, in this context, means “again and again”.

- (248) You're *always* asking for money. I sometimes wonder where it all goes! –
Parents! They're *always* complaining!

The Language Module of Topic 2 in the same volume presents a “revision” of the Past-PerfProg (in two exercises), which is somewhat surprising since this form had not been properly introduced before but only mentioned in some previous overviews on tense forms. About the PastPerfProg we learn that the form refers to continued actions or events that happened before a certain point in time, that it usually collocates with *since* or *for*, and that it is restricted to dynamic verbs. It is interesting in this context to see that the few PastPerfProg form tokens in our database come from volume 6 of GLN and not from any of the other textbook volumes. However, in GLN volume 4 there is one occurrence of “had been planning”, a form which is not included in our list of 100 high-frequency verbs.

5.8.2 The progression in *English G 2000 A* (volumes 1–6)

The grammatical progression in *English G 2000 A* (EG 2000) is largely similar to that in GLN, at least with respect to progressive form types. Volume 1 (Unit 3) starts with a comparatively detailed introduction of the PresProg. We find eight short texts and seven exercises that are dedicated to this new tense form. Example (249) presents part of one of the texts from Unit 3 (note: Marmalade is a cat and Trundle a tortoise). The “Language Summary” at the back of the textbook comments on form and function of the PresProg and refers to the existence of an equivalent form in some German dialects (cf. 5.1). We learn that the form is used to express that someone is in the process of doing something, or to refer to an action which is still in progress at the time of speaking.

- (249) MARMALADE: You *aren't packing*, Trundle!
TRUNDLE: No, Marmalade, I'm not *packing*. Sally and Jenny *are packing*. I'm *sitting* here and *writing* a letter.
MARMALADE: Boring! Bye bye!

In Unit 5 of the same volume, there is an initial comparison of the PresProg and the simple present, which is also introduced in that unit (no. 5), like in GLN two units after the progressive form (cf. the discussion on the sequence of progressive and non-progressive tense forms in 5.8.1). Unit 7 discusses the contrast between the two forms in some more detail by means of six short and two longer texts and four exercises (see (250) for a relevant EG 2000 example). The related Language Summary section notes the frequent collocation of the PresProg with time adverbials such as *now*, *at the*

moment, today, or this morning, and introduces the notion of stative verbs, which are not normally used in the progressive (cf. Schwarz 1997: 151).

- (250) MISS HUNT: OK then, go and find your animals. See what they're *doing*. Are they *sleeping*? Or *eating*? Or *playing*? Where and when do they usually sleep, play and eat?

Also introduced in volume 1 of EG 2000 in four short texts and five exercises (Unit 8) is the *going to* future. Its form and functions (expression of intentions and future plans) are presented in a Language Summary section, where we even find a hint on the grammaticalization of the form. The coursebook authors note that, in this construction, "going" has got nothing to do with moving from A to B.

The only progressive form that is new in EG 2000 volume 2 is the PastProg. PastProgs are first used in four short texts, one longer text, and three exercises in Unit 8 of the volume (cf. example (251) below). In the texts, the PastProg mainly occurs in questions or in framing contexts, a function which is also referred to in the Language Summary at the back of the book. We learn there that the PastProg often describes an action which is in progress when another action starts, and that the form basically expresses that an action is in progress at a particular moment in the past.

- (251) 'Have you seen an older woman and a young boy?' 'Was the lady *wearing* a yellow dress?' 'Yes, she was.' 'They left five minutes ago.' 'Where *were* they *going*? Did they say?' 'Well, the woman said something about video games...'

Unit 1 of volume 3 starts with a revision of the PresProg and then introduces the PresPerfProg as a new tense form that is used to express an ongoing action which started in the past. The form appears in three texts (cf. 252) for a short dialogue extract) and is the topic of two exercises. Like in GLN, the collocation of the PresPerfProg with *since* and *for* is stressed in the Language Summary of EG 2000. Also like in GLN, there is a unit in volume 3 of the EG 2000 series (Unit 5) that deals with the use of PresProgs to refer to fixed plans and future arrangements. The PresProg with future meaning features in one text and two exercises and in a short section of the Language Summary. The grammar booklet that accompanies EG 2000 volumes 3 and 4 (*Grammatikheft A3/4*, Schwarz 2000) covers that use of the PresProg too and exemplifies it with the brief telephone exchange given in (253).

- (252) MEGAN: *Have you been waiting* long, Dad?
MR OWEN: Yes, for half an hour. Are you very hungry?
MEGAN: No, it's OK, Dad. I didn't get home till 4.30. And I've *been writing* my diary since then.
- (253) "What *are you doing* next Saturday?"
"I'm *meeting* Bert."

Volume 4 of EG 2000 first provides a review of the PresProg with future meaning (Unit 1) and then introduces the PastPerfProg (Unit 2), however only the form without

giving any explicit information on its function and use. In the Language Summary of that unit we only learn that “the past perfect has got a progressive form too” (translated from Schwarz 1999: 126), and find the example displayed in (254) below. Two short texts and two exercises in Unit 3 of the same volume focus on a revision of tenses. The related overview in the Language Summary section of the book includes the forms and functions of all previously introduced progressives, i.e. PresProg, PastProg, PresPerfProg, and PastPerfProg. It is noted that PastPerfProgs refer to actions that continued up to a certain point in the past. Again, the verb form used in the illustrative example is *waiting*.

(254) The letter came on Friday. Amy *had been waiting* for it since Monday.

Several units in EG 2000 volumes 5 and 6 serve to further exemplify the different progressive forms in different contexts and to revise their functions. In addition to the functions mentioned before, Unit 5 in volume 5 deals with the use of the PresProg with *always* in order to express that something happens again and again, or too often. The same unit provides information on the relation between aspect and state and activity verbs. In the Language Summary it is said that “[s]tate verbs aren’t usually used in the progressive form” and that “[a]ctivity verbs can be in the simple form [...] or the progressive form” (Schwarz 2001: 139). Exceptions to these rules, concerning verbs such as *have*, *think*, *see*, *look*, and *feel* which can express both states and activities, are also mentioned. Unit 3 in volume 6 reviews those progressive forms that refer to the past; it provides three exercises and comments on the use of the PastProg, the PresPerfProg, and the PastPerfProg, repeating the functions given earlier on in the course. Ways of talking about the future are summarised and revised in Unit 4 in one exercise in the grammar practice section.

There is no Language Summary in EG 2000 volume 6, but a separate grammar book is available (*Cornelsen English Grammar*, CEG, Fleischhack & Schwarz 2001) which is supposed to be used alongside the coursebook. This grammar basically sums up the contents of all Language Summary sections of volumes 1 to 5 in a clear and systematic way and provides further examples and form/function overviews. In the chapter on “The tenses of the full verbs” we find such systematic overviews for each progressive form, usually in contrast with its non-progressive counterpart. In addition to the functions given for the PresProg in the coursebooks, the CEG mentions temporary repeated actions that can be expressed by this form and provides the two sentences in (255) to exemplify this use.

(255) Sally *is working* at a restaurant during her holidays.
John *is walking* to work till the bus strike is over.

Similar examples (cf. (256)) are also found in the *Learning English Grundgrammatik* (LEG, Ungerer et al. 2001) which accompanies the GLN series. However, the comments on functions only focus on temporariness and do not mention the repeated occurrence of the described actions.

- (256) Our local football team **is playing** very well this season.
 My uncle **is writing** a book about golf. (emphasis in original)

I will now provide a comparative overview of the grammatical progression in GLN and EG 2000 and briefly discuss how the presentation of progressive forms, functions, and contexts in the grammar sections of the coursebooks relates to the corpus-driven findings obtained in my analysis of progressives in spoken BrNSE. A much more detailed discussion of the similarities and differences between GLN and EG 2000 data on the one hand and BNC_spoken and BoE_brspok data on the other can be found in Chapter 6. The analysis in Chapter 6 will further evaluate and compare the results of Chapters 4 and 5.

5.8.3 Summary and discussion

In his discussion of the treatment of the present perfect in German EFL teaching materials, Schlüter (2002a:328) notes substantial differences in the presentation of the construction in two coursebook series *English G 2000 A*, which is also used in my study, and *Password Green*, another relatively recent series published by Klett Verlag, which is comparable in many respects to GLN but incorporates different new teaching methods. With many of our context and function phenomena we also found rather different results for the analysed GLN and EG 2000 progressives (cf. Sections 5.5 to 5.7). Concerning the sequence of introduction of progressive tense forms, however, we could not observe any significant discrepancies. As Table 47 indicates, the progression in both series is: PresProg (vol. 1) → PastProg (vol. 2) → PresPerfProg (vol. 3) → PastPerfProg (vol. 4). The table provides a comparative overview of the treatment of progressives in the two EFL coursebook series and summarises what has been said about the grammatical progression in GLN and EG 2000 in the previous two sections.

A look at the GLN- and EG 2000-columns in Table 47 shows that the sequence of progressive-related topics in both series is strikingly similar and that it runs very much in parallel. Not only are the same forms introduced roughly at the same stage in the course, the textbooks also provide largely comparable information on the functions of these forms and their preferred adverbial collocation. Sometimes different, though, is the detail in which certain phenomena are presented, practiced, and exemplified. On the whole, EG 2000 contains a larger number of short texts and exercises that deal with progressives than GLN. On the other hand, we find more revision sections in GLN which do not introduce any new features but review what has been said before and usually offer one or two additional exercises (cf. esp. vol. 2).

If we now compare the presentation of progressives in the two coursebook series with some of our corpus-driven results obtained from the BNC/BoE analysis, we can first of all say that, judged by frequencies of occurrence, the sequence of introduction of the progressive forms in school grammar certainly makes sense. The progressive forms, PresProg, PastProg, PresPerfProg, and PastPerfProg, are introduced in order of frequency of occurrence in spoken BrNSE. When we look at what the coursebooks

Table 47. The introduction and presentation of progressives in GLN and EG 2000

textbook volume	<i>Green Line New (GLN)</i>	<i>English G 2000 A (EG 2000)</i>
vol. 1 (Aston et al. 1995; Schwarz ed. 1997)	<p>Unit 4, Step B: PresProg (new), 2 short texts, 5 exercises; grammar section: PresProg examples, form, function (someone is in the process of doing something; action still in progress; German-English contrast: no equivalent form in German)</p> <p>[Unit 6, Step B: Simple Present (new)]</p> <p>Unit 7, Step B: Simple Present and PresProg in contrast, 1 short text, 2 exercises; grammar section: PresProg examples (with time adverbials), function (action in progress, not finished), collocation with time adv., e.g. <i>today, now, just now, this morning</i></p> <p>Unit 10, Step B: <i>going to</i> future, 1 short text, 4 exercises, grammar section: examples, form, function (talk about intentions or planned actions)</p>	<p>Unit 3: PresProg (new), 8 short texts, 7 exercises, language summary: PresProg examples, form (short/long forms), function (someone is in the process of doing something; action still in progress; German equivalent in some dialects)</p> <p>[Unit 5: Simple Present (new)]</p> <p>Unit 5: Simple Present and PresProg – an initial comparison, 1 text, 3 exercises; language summary: function (action in progress at the moment of speaking)</p> <p>Unit 7: Simple Present and PresProg in contrast, 6 short and 2 longer texts, 4 exercises, language summary: PresProg examples, function (action in progress, not finished), collocation with time adv., e.g. <i>now, at the moment, today, this morning</i>, stative verbs (e.g. <i>be, believe, hear, know, see</i>) not normally used in the progressive</p> <p>Unit 8: <i>going to</i> future, 4 short texts and 1 longer text, 5 exercises, language summary: form, function (intentions and future plans), hint on grammaticalization (<i>going has got</i> nothing to do with moving from A to B)</p>
vol. 2 (Ashford et al. 1996; Schwarz ed. 1998a)	<p>Unit 1, Step C: <i>going to</i> future (revision), 1 exercise;</p> <p>Unit 3, Step B: Simple Present and PresProg in contrast (revision), 1 text, 1 exercise;</p> <p>Unit 4, Step B: PresProg (revision), 1 short text, 1 exercise, no additional information;</p> <p>Unit 9, Step B: PastProg (new), 2 short texts, 4 exercises, grammar section: PastProg examples, all expressing the framing function; function (ongoing actions in the past; something unfinished in progress, new action starts), form, collocation with <i>while, when</i></p>	<p>Unit 8: PastProg (new), 4 short texts and 1 long text, 3 exercises, language summary: PastProg examples, function (action in progress at a particular moment in the past; framing: the PastProg often describes an action which is in progress when another action starts)</p>

Table 47. (continued)

vol. 3 (Ashford et al. 1997; Schwarz ed. 1999)	<p>Workshop A, Step 5 (“Practice Page”): PresProg (revision), 1 exercise Unit 2, Step B: PresProg with future meaning, 1 short text, 3 exercises; grammar section: exam- ples, all with future time adverbial; func- tion (to express fixed plans or arrange- ments; picture of a diary), collocation with time adverbials, e.g. <i>tomorrow</i>, <i>next week</i>, <i>on Tuesday</i> Unit 3, Step B: PresPerfProg (new), 1 short text, 5 ex- ercises; grammar section: form, func- tion (durative: something started in the past and has not stopped yet), exam- ples (<i>waiting</i>, <i>raining</i>, <i>living</i>, <i>learning</i>), German translation with the adverbial <i>schon</i>, collocation with <i>since</i> and <i>for</i></p>	<p>Unit 1: PresProg (revision), 1 exercise, language summary: example, function (action in progress, unfinished); PresPerfProg (new), 2 short texts and 1 long text, 2 ex- ercises, language summary: form, func- tion (express an ongoing action which started in the past), collocation with <i>since</i> and <i>for</i>, not used with stative verbs Unit 3: PastProg (revision), 1 text, 1 exercise Unit 5: PresProg with future meaning, 1 text, 2 exercises, language summary: examples, function (for fixed plans and future ar- rangements)</p>
vol. 4 (Ashford et al. 1998; Schwarz ed. 1999)	<p>Unit 1, Step C: tenses (revision), 1 text, 2 exercises; grammar section: overview on tense forms (including PastPerfProg), no in- formation on functions and use</p>	<p>Unit 1: PresProg with future meaning (revi- sion), 1 exercise, language summary: ex- amples, function (fixed plans or arrange- ments); Unit 2: PastPerfProg (new), 1 text, language summary: 1 example, no information on functions and use; Unit 3: tenses (revision), 2 short texts, 2 exer- cises, language summary: overview of forms and functions, PresProg (actions in progress at the moment of speak- ing, future plans), PastProg (actions in progress at a certain moment in the past), PresPerfProg (ongoing actions that started in the past), PastPerfProg (actions that continued up to a certain point in the past)</p>
vol. 5 (Ashford et al. 1999; Schwarz ed. 2001)	<p>Topic 1, Language Module: Simple and progressive in contrast (past, present, present perfect), text in speech bubbles, 2 exercises; grammar section: dynamic vs. stative verbs, use of dynamic verbs (e.g. <i>run</i>, <i>walk</i>, <i>play</i>, <i>swim</i>) with</p>	<p>Unit 5: progressives (revision), language sum- mary: forms and functions as in vol. 3 (Unit 3), in addition: PresProg with <i>al- ways</i> to express that something happens again and again (or too often); aspect</p>

Table 47. (continued)

	<p>simple/progressive forms, use of static verbs (e.g. <i>be, have, think, like</i>) not normally with progressive forms, only if they express an activity;</p> <p>Exclusively in grammar sections: ways of expressing future time (revision), forms and functions, <i>going to</i> future (plans, near future), PresProg (definite arrangements, collocation with time adv.); tenses (revision), overview of forms only</p>	<p>and state verbs (e.g. <i>believe, know, like, mean, want</i>) vs. activity verbs (e.g. <i>play, buy, write</i>), 1 long text, 5 exercises, language summary: “[s]tate verbs aren’t usually used in the progressive form”, “[a]ctivity verbs can be in the simple form . . . or the progressive form” p. 139; note on exceptions: some verbs, such as <i>have, think, see, look, feel</i>, can express states and activities and be used in the progressive and simple form;</p>
<p>vol. 6 (Ashford et al. 2000); Schwarz ed. 2002)</p>	<p>Topic 1, Language Module:</p> <p>PresProg + <i>always</i>, 1 exercise, grammar section: PresProg with emphatic usage, examples all with <i>always</i>, function (to express a negative emotional attitude), meaning of <i>always</i> in this context: “again and again”;</p> <p>PresProg, PastProg, PresPerfProg (revision), 5 exercises, grammar section: review of tenses, forms and functions, PresProg (“describes what is happening at the moment of speaking and is not yet over”, “also used for activities going on over a longer period of time”, “also describes definite arrangements that have been made for the future” p. 110); PastProg (“stresses that an event in the past was happening <i>around a certain point in time</i>. It gives background to something else in the past.” p. 110, emphasis in original), PresPerfProg (“stresses that an event that started in the past has gone right up to the present (and may even continue)” p. 110), PastPerfProg (“expresses that an event began before a point of time in the past and continued up to that point” p. 111);</p> <p>Topic 2, Language Module:</p> <p>PastPerfProg (revision), 2 exercises, grammar section: examples (with <i>complaining, standing</i>), function (events happened before a point in time, continued activities), collocation with <i>since</i> and <i>for</i>, restriction: only used with dynamic verbs</p>	<p>Unit 3:</p> <p>PastProg, PresPerfProg, and PastPerfProg (revision), 3 exercises in grammar practice section, comments on use: PastProg for actions “in progress at a time in the past”, PresPerfProg for “an action that began in the past and has continued up to the present”, PastPerfProg for “an action that continued up to a time in the past” p. 95;</p> <p>Unit 4:</p> <p>talking about the future (revision), 1 exercise in grammar practice section, comments on use: <i>going to</i> future “for an intention, for a prediction based on the present situation”; PresProg “for an arrangement” p. 97;</p> <p>there are no language summary sections in vol. 6, instead there is a separate grammar book to use alongside the book (<i>Cornelsen English Grammar</i>, Fleischhack & Schwarz 2001);</p>

and school grammars say about progressive functions, though, we get a rather simplified picture that deviates quite a bit from actual usage. With respect to all progressive forms, GLN and EG 2000 focus on the function feature of continuousness and stress that *an action is in progress at a certain point in time*. Very much marginalised in the books is the second central function feature we found to be frequently present in real spoken English progressives, namely repeatedness. None of the grammar or Language Summary sections in GLN and EG 2000 deals with that function. Only in CEG (Fleischhack & Schwarz 2001:56) we find a reference to the use of PresProgs for temporary repeated actions (see (255) above for examples). LEG (Ungerer et al. 2001:56) gives similar examples (see (256)) but does not mention repeatedness at all.

Apart from continuousness and the “action in progress” function, the coursebooks explicitly refer to only two of our seven identified additional functions (cf. 4.4.5): “framing” and “emphasis/attitude”. The “emphasis/attitude” function, however, appears to be restricted to the use of PresProg forms in collocation with the adverbial *always*. We noted above (5.8.1 and 5.8.2) that GLN (vol. 6) and EG 2000 (vol. 5) introduce this pattern as a means to express a negative emotional attitude (cf. also Table 47). What they do not cover is the more frequent general emphatic function of progressives that we discussed earlier on in the function analysis of BNC and BoE data (cf. 4.4.5). On the other hand, the “framing” function, which was rather rare in our spoken English material, is given a lot of attention in the coursebook grammar sections. The explanations we find in volume 2 of GLN and EG 2000 on the function of the PastProg might create the impression that the form is always used in framing contexts which is not at all the case (cf. 4.5). Still on the function side, we find that those additional functions that were found to be comparatively frequent in spoken BrNSE do not appear in textbook grammar. I think it would be important to cover the functions “general validity”, “politeness and softening”, and “gradual change”, at least in one of the tense form revision units in volumes 5 or 6 of GLN and EG 2000 and thus widen the meaning spectrum of progressives in their coursebook presentation.

Concerning time reference and taking into account the high percentage of BNC/BoE corpus examples which express future time (cf. 4.4.1), we may argue for an earlier introduction of the PresProg with future meaning. However, what I would consider much more important than an earlier introduction of this particular form-function relation is the inclusion of an enhanced lexical-grammatical perspective, not just related to verbs that most frequently occur in future time reference contexts but also related to all other progressive forms and functions. The only reasonably detailed verb-related information we find in the coursebooks is on the stative-dynamic verb distinction in the context of contrasting simple and progressive forms. There is not much differentiation among progressive tense forms with the exception of LEG, the grammar book that backs up the GLN series. It is noted in LEG that the PresPerfProg is particularly frequent with verbs such as *lie, sit, stand, wait, rain, shine, and snow* (cf. Ungerer et al. 2001:58). In EG 2000 and with respect to the other progressive forms in GLN, all we can do is infer from the examples used in the grammar sections and grammar booklets which verbs might be typical in which constructions. It would certainly

be good to get more information on verb-specific preferences for different progressive forms, functions, and contexts.

Also, there could be more information in the textbooks on co-selection phenomena. The GLN and EG 2000 grammar and Language Summary sections include some remarks on progressive-adverbial collocation but do not refer to any of the other context features that were examined in the present analysis of progressives. While the collocation with time adverbials even tends to be overemphasised, other co-selection patterns, e.g. of progressives and negation, progressives and subjects, or progressives and questions, are not mentioned at all.

We can already see from the discussion of the grammatical progression in the coursebooks that the presentation of progressives in EFL teaching materials is not fully in accordance with their use in natural English, and that we might want to change a few things in the coursebooks and school grammars in order to bring the English we teach more fully into accordance with real-life language use. In the next chapter it will be discussed in some more detail in what respects the use of progressives in “school” English differs from its use in real spoken English. We will see which features in GLN and EG 2000 data are similar to what occurs in BNC_spoken and BoE_brspok, and where we find significant deviations from the actual use of progressives. Chapter 7 will then further evaluate the results from the BNC/BoE-GEFL TC data comparison and deal with a number of possible pedagogical implications of our corpus-driven analysis.

CHAPTER 6

Progressives in real spoken English and in “school” English

A comparison

The previous two chapters have dealt in some detail with the presentation of progressives in EFL textbooks (Ch. 5) and with their use in spoken English (Ch. 4). A number of context and function phenomena have been included in the analysis to ensure a holistic approach to the progressive in use, rather than treating it as an isolated grammatical structure. The analysis highlighted many typical progressive-related patterns and co-selection phenomena, and revealed that the connections between individual lexical items and the selected context and function features are much stronger than most existing grammatical descriptions of the progressive indicate. This implied that a lot more attention in language analysis and description still needs to be paid to the interrelation of lexis and grammar and that probably more emphasis has to be put on lexis.

I will now take my investigation of progressives in real spoken English and “school” English one important step further in that I compare what I found in BNC_spoken and BoE_brspok data with the results of my coursebook corpus analysis. This comparison of corpus findings (BNC/BoE vs. GLN/EG 2000) is supposed to provide an answer to the question “How much attention do EFL textbooks pay to the facts of real English?” In the following sections, the presentation of progressives in the teaching materials will hence be evaluated in the light of spoken BrNSE corpus evidence.

The discussion of similarities and differences in the behaviour of progressives in “school” English and real English will include both average and verb-specific distributions. However, taking into account the above-mentioned findings on the lexical grammar of progressive forms, it will not only be on a general but also on a verb-by-verb basis. I am hoping to show which aspects in the presentation of progressives in the coursebooks differ most significantly from actual usage so that suggestions can be made later on that basis concerning possible corpus-informed adjustments in language teaching.

6.1 Progressives and context phenomena

Having identified the most typical co-selection patterns of different types of progressives in spoken BrNSE and EFL textbook English, the next question that I will deal with is “Do the contexts of BNC and BoE progressives correspond to those found in the GLN and EG 2000 datasets?”

Included in the comparative context analysis are the same phenomena that were examined in the respective parts of Chapters 4 and 5 (cf. the list of features in Table 7 in 4.2.4). Accordingly, the following sections will discuss the co-occurrence of progressives and different tense forms, *TO BE* form contractions, subjects, objects, prepositions, negation, questions, *if*-clauses, relative clauses, and time/place/other adverbials and compare the respective results presented in 4.5 with those discussed in 5.7.

6.1.1 Distribution of different tense forms

The first phenomenon we looked at in the context analysis of progressives was the distribution of PresProg, PastProg, PresPerfProg, and PastPerfProg forms. The shares of these four forms were determined for the progressive datasets from BNC_spoken, BoE_brspok, GLN, and EG 2000 and separately for individual verbs in the same corpora.

We noted that most frequent by far in all examined corpora were present progressives, followed by past progressives. Comparatively very rare in the coursebooks and in spoken English are PresPerfProg, PastPerfProg, and fragmentary forms, as in *What you wearing tomorrow night?* or *I ai n't telling you no more* (both from BNC_spoken). Figure 134 provides a combination of Figures 8 and 100 and shows that, for the groups of progressives from BNC_spoken, BoE_brspok, GLN, and EG 2000, the percentages of different tense form types are on the whole rather similar. Only the values that go back to EG 2000 differ to some extent from the percentages based on the other three datasets. In the EG 2000 subcorpus of GEFL TC we find a particularly high share of PastProg forms.

More significant, however, are the differences that become apparent when we look at the actual realisations of the PresProg, i.e. the distribution of forms of the verb *to BE* within this tense form. If we compare the pie charts in Figures 9, 10, 101, and 102, we note some clear discrepancies between spoken English and “school” English data. While in both spoken BrNSE corpora *'re V-ing* is by far the most frequent PresProg pattern (with shares of 35.43% and 33.86%), GLN and EG 2000 clearly favour the construction *are V-ing*. The form *'s V-ing* is also much more common in BNC/BoE than in GLN, though not in EG 2000. So it appears that contracted PresProg forms are perhaps not given enough attention in the textbook series, particularly in GLN. This finding gets support on a more general level, as the bar charts in Figure 135 illustrate. We see that the shares of long and short forms in GLN deviate significantly from the percentages in the three other corpora. With respect to verb form contraction, progressives are obviously somewhat misrepresented in this coursebook series.

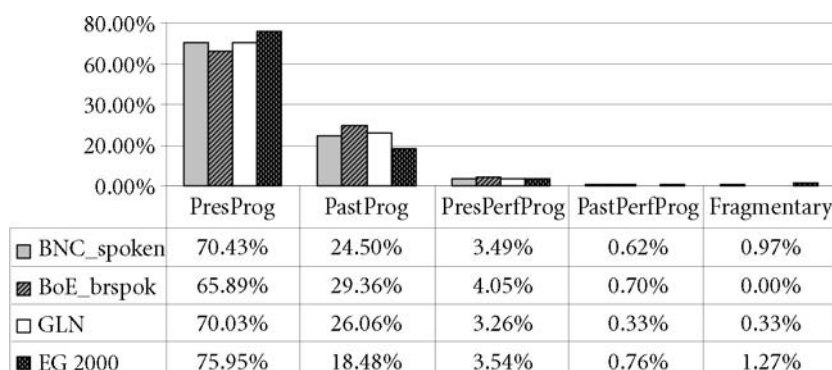


Figure 134. The distribution of different tense forms in BNC_spoken, BoE_brspok, GLN, and EG 2000

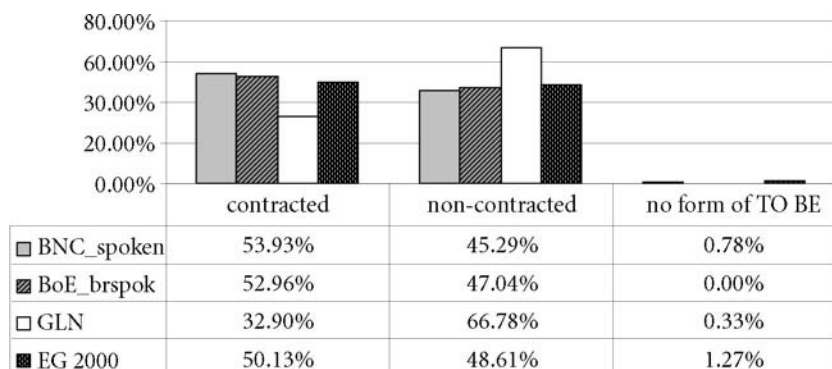



















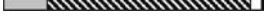




Figure 135. Contracted and non-contracted progressive forms in BNC_spoken, BoE_brspok, GLN, and EG 2000

A look at the verb-specific tense form distributions uncovers further differences, but also similarities, between the presentation of progressives in the teaching materials and their actual use. Table 48 displays the tense form shares in BNC/BoE, GLN, and EG 2000 for some selected verbs. We basically find an over-representation of PresProgs with the displayed verbs in the textbooks and an under-representation of the past time reference forms, i.e. PastProg, PresPerfProg, and PastPerfProg (note, however, the values for *trying*, which show an opposite trend). For many others of the more frequently occurring verbs in GEFL TC, the tense form distribution is comparable to that determined for the spoken English datasets.

On the whole we can say that there are a number of similarities between textbook and real spoken English data when it comes to progressive tense form distributions. However, there is a tendency in “school” English (especially in the language used in EG 2000) to neglect PastProgs in favour of the PresProg. Also, we found significantly

Table 48. The distribution of progressive tense forms for selected verbs in BNC/BoE (B), GLN (G), and EG 2000 (E)

verb form	B, G, or E	PastProg	PresProg	PresPerf Prog	PastPerf Prog	distribution graphically illustrated
						
asking	B	33.11%	61.59%	3.97%	1.32%	
	G	0.00%	100.00%	0.00%	0.00%	
	E	0.00%	80.00%	20.00%	0.00%	
going	B	15.38%	84.13%	0.48%	0.00%	
	G	9.89%	89.01%	1.10%	0.00%	
	E	7.29%	92.71%	0.00%	0.00%	
helping	B	30.59%	60.00%	9.41%	0.00%	
	G	50.00%	50.00%	0.00%	0.00%	
	E	0.00%	100.00%	0.00%	0.00%	
playing	B	19.09%	70.00%	7.27%	3.64%	
	G	0.00%	100.00%	0.00%	0.00%	
	E	7.14%	85.72%	7.14%	0.00%	
reading	B	36.11%	51.39%	11.11%	1.39%	
	G	25.00%	75.00%	0.00%	0.00%	
	E	20.00%	80.00%	0.00%	0.00%	
staying	B	31.85%	65.93%	1.48%	0.74%	
	G	20.00%	80.00%	0.00%	0.00%	
	E	0.00%	100.00%	0.00%	0.00%	
trying	B	16.85%	78.26%	4.35%	0.54%	
	G	25.00%	75.00%	0.00%	0.00%	
	E	40.00%	50.00%	10.00%	0.00%	

lower shares in GEFL TC than in BNC/BoE for the contracted patterns *'re V-ing* and *'s V-ing*. In addition there is a general trend in GLN to largely under-represent short forms (cf. Figure 135).

6.1.2 Progressives and subjects

Another feature examined in the context analysis was the co-selection of progressives and items in subject position. Most common by far in both types of English (“school” English and real spoken English) were the personal pronouns *I, you, he, she, it, we,* and *they*. These pronouns showed largely different frequencies in the analysed datasets, with *I, you,* and *we* topping the list (cf. Figures 23 and 110). The distributions of personal pronoun subjects in BNC_spoken, BoE_brspok, GLN, and EG 2000 are illustrated again in Figure 136.

When we compare the shares of personal pronoun subjects in GLN with the BNC and BoE values, we notice three significant deviations. The pronoun *he* is considerably

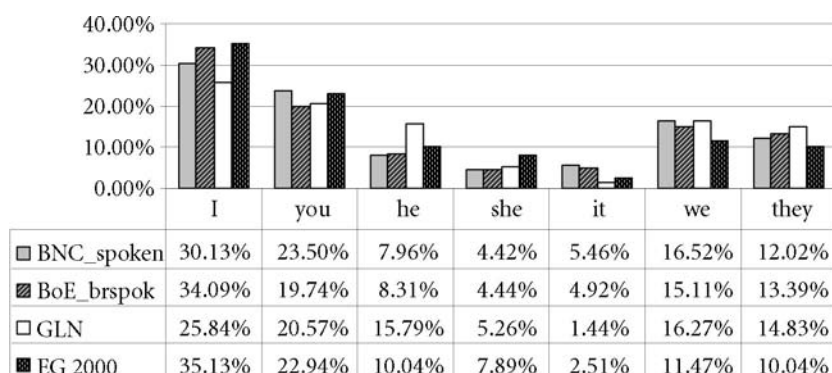


Figure 136. The distribution of personal pronouns as subjects of progressives in BNC_spoken, BoE_brspok, GLN, and EG 2000

overused in the GLN coursebook data, whereas *I* and *it* occur much less frequently in subject position of GLN progressives than in the spoken English data. Some differences also exist between spoken BrNSE and the second textbook series. In EG 2000 *she* is found rather often in subject position of progressives. On the other hand, comparatively infrequent here are the pronouns *we* and *they*.

Another clear contextual difference between GEFL TC and BNC/BoE progressives relates to the use of personal names as subjects. As Figures 24 and 111 above have shown (cf. Sections 4.3.3 and 5.5.3), names only form the subject of 2.80 and 2.44 per cent of the BNC_spoken and BoE_brspok concordance sets, but occur in subject position of 12.15 and 14.33 per cent in EG 2000 and GLN progressives. In general, and obviously due to the restricted vocabulary in the coursebooks, we find less variation among GLN and EG 2000 subjects than in spoken BrNSE, and there is certainly nothing criticisable about that. We might, however, criticise the existing mismatches concerning several personal pronouns and the significant overuse of personal names in subject position of the coursebook progressives.

If we now turn our attention to the relations between verbs and the distribution of personal pronoun subjects, an area in which we found a lot of variation in the analysed datasets (cf. 4.5.3 and 5.7.3), we see that the differences continue. The largely deviating percentages in the columns of Table 49 demonstrate that spoken English and “school” English progressives do not have a lot in common when it comes to subject-verb type patterns. To mention but one example, the shares of *you* as the subject of *doing*-progressives in GLN and EG 2000 data are unusually high while the first person pronoun subjects *I* and *we* are very much under-represented in the coursebook concordances of the same verb form. We can hence say that textbook authors and native speakers of English, to a large extent, favour different items in subject position of progressive constructions, or rather they mainly favour the same items, i.e. personal pronouns, but with differing degrees of frequency.

Table 49. The distribution of personal pronoun subjects across selected verbs in BNC/BoE (B), GLN (G), and EG 2000 (E)

verb form	B, G, or E	subject <i>I</i>	subject <i>you</i>	subject <i>he</i>	subject <i>she</i>	subject <i>it</i>	subject <i>we</i>	subject <i>they</i>
asking	B	29.80%	21.85%	5.96%	3.31%	0.00%	9.93%	8.61%
	G	0.00%	33.33%	0.00%	0.00%	0.00%	0.00%	66.67%
	E	40.00%	20.00%	0.00%	0.00%	0.00%	0.00%	0.00%
doing	B	14.55%	22.42%	7.27%	0.61%	1.82%	18.18%	11.52%
	G	5.26%	47.37%	5.26%	5.26%	0.00%	10.53%	5.26%
	E	6.25%	53.13%	9.38%	6.25%	0.00%	0.00%	6.25%
having	B	11.43%	20.00%	5.71%	4.29%	1.43%	25.71%	11.43%
	G	0.00%	40.00%	0.00%	0.00%	0.00%	0.00%	40.00%
	E	14.29%	0.00%	0.00%	28.57%	0.00%	28.57%	14.29%
looking	B	21.92%	28.77%	4.11%	0.68%	1.37%	28.08%	7.53%
	G	10.00%	0.00%	15.00%	10.00%	0.00%	15.00%	20.00%
	E	47.62%	4.76%	4.76%	9.52%	0.00%	9.52%	4.76%
saying	B	32.03%	25.49%	4.58%	1.96%	2.61%	7.84%	6.54%
	G	0.00%	16.67%	16.67%	0.00%	0.00%	0.00%	16.67%
	E	40.00%	20.00%	0.00%	0.00%	0.00%	0.00%	40.00%
taking	B	20.00%	13.33%	8.33%	4.17%	3.33%	13.33%	11.67%
	G	60.00%	0.00%	0.00%	0.00%	0.00%	0.00%	60.00%
	E	37.50%	12.50%	12.50%	0.00%	12.50%	0.00%	37.50%
talking	B	19.15%	20.21%	4.26%	2.13%	0.00%	36.70%	8.51%
	G	0.00%	0.00%	20.00%	0.00%	0.00%	20.00%	20.00%
	E	4.17%	20.83%	4.17%	0.00%	0.00%	0.00%	12.50%
wearing	B	29.53%	20.81%	14.09%	10.07%	0.00%	2.01%	11.41%
	G	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
	E	0.00%	33.33%	16.67%	16.67%	0.00%	0.00%	0.00%

6.1.3 Progressives and objects

The high degree of variation that we found among items in object position of progressive forms was shared by all examined datasets. Our search for typical co-occurrence patterns based on the entire sets of progressives from BNC_spoken, BoE_brspok, GLN, and EG 2000, i.e. all verb forms taken together, was not particularly fruitful and only highlighted two rather common object types: *the* + noun (group) and *a* + noun (group).

A comparison of spoken English and “school” English findings now shows that these two types of objects are considerably more frequent in GLN than in BNC/BoE (cf. Figures 25 and 112 in 4.5.4 and 5.7.4). In EG 2000, the share of *the* + noun (group) in object position is very similar to the BNC_spoken and BoE_brspok values but, as in GLN, *a* + noun (group) is much more common in this coursebook series than in spoken English.

As for the verb-individual findings on the co-selection of progressives and objects, we note further differences on the “repeated patterns” side between BNC/BoE and GLN/EG 2000 data but we also find that EG 2000 is much closer to real-life language patterns than GLN. As our lists of typical verb-object collocations in 5.7.4 showed, the GLN series only features three patterns (*going* + preposition + *the* + noun (group), *looking* + preposition + *the* + noun (group), and *making* + preposition + *a* + noun (group)), the first one of which is actually very rare in the BNC/BoE data (cf. Table 23 in 4.5.4). Common patterns that were found in spoken English, such as *having* + *a* + noun (group), *telling* + *him*, or *telling* + *you*, are not repeatedly used in the GLN texts. The three patterns that I have just mentioned do, however, occur several times in EG 2000, in addition to the other object-verb collocation types that were found in GLN. This does not mean that EG 2000 mirrors actual usage, only that with regard to progressives and their objects it covers more of the typical patterns found in spoken BrNSE.

A number of real English patterns are missing in both coursebook series. Sometimes the reason for this non-occurrence in GLN and EG 2000 is clear: we do not find any repeated instances of *accepting* + *the* + noun (group), *bothering* + *me*, *liking* + *it*, *providing* + *a* + noun (group), or *setting* + *the* + noun (group), because these verb forms are either very rare or do not occur at all in GEFL TC. However, with *asking* + *you*, *listening* + *to* + *the* + noun (group), *reading* + *it*, *talking* + *the* + noun (group), and *telling* + *me*, I do not see any apparent reason why these collocations should not be presented in the textbooks and why other contexts (e.g. *asking* + *questions*, *listening* + *to* + *Ben*, *telling* + *you*, *telling* + *Ben*) should dominate instead.

6.1.4 Progressives and prepositions

The above analysis of progressives and prepositions in spoken English and in “school” English supported the assumption that prepositions are not typical features of “the progressive” as such but that their repeated occurrence is strongly lexically determined. Let us therefore just focus on the comparison of some verb-specific collocation phenomena of progressive forms and prepositions such as *at*, *about*, *for*, or *with*.

A look at the lists of common progressive-preposition collocations in GLN, EG 2000, and BNC/BoE (cf. 4.5.5 and 5.7.5) reveals that some of the GLN and EG 2000 patterns are also frequent in spoken BrNSE (e.g. *looking at*, *looking for*, *listening to*, *talking about*, *talking to*, and *thinking about*), whereas other collocations that are repeatedly used in the textbooks do not, or at least not frequently, occur in the datasets from BNC_spoken and BoE_brspek. Worth mentioning in this latter group are *coming back*, *getting at*, *talking with*, and *working on*. While these collocations appear relatively more often in “school” English than in real spoken English, some other typical spoken English patterns, especially *agreeing with*, *dealing with*, *picking up*, *ringing up*, *setting up*, and *sorting out*, are completely missing or very rare in the coursebook data.

As for an assessment of the two textbook series, we found again that EG 2000 shows fewer deviations from real English than GLN and that it includes a larger num-

ber of the progressive-preposition patterns that frequently appear in the BNC/BoE data. This observation confirms the findings of an earlier study on the use of *looking at* and *looking for* in GEFL TC and BNC/BoE. In this smaller-scale study I also found that the EG 2000 values are closer to the frequencies determined for the spoken BrNSE corpora than the GLN values (cf. Römer Forthcoming).

6.1.5 Progressives and negation

In Section 5.5.6 I briefly commented on the difference between the average shares of negated progressives in BNC_spoken and BoE_brspok on the one hand and GLN and EG 2000 on the other. I found that, with 3.80 and 4.56 per cent, the values of negation are significantly lower in the coursebooks than in the spoken BrNSE data, where shares of 7.98 and 8.59 per cent had been determined. Figure 137 provides a combination of Figures 27 and 114 and illustrates the just-described difference across corpus types.

A possible explanation for these rather low percentages of progressive negation in the textbooks could be that, although I have selected only spoken-type material to base my analysis on (cf. 5.3.1), this kind of “school” English is closer to writing than to actual speech. In several written English language corpora the shares of negation were found to be considerably lower than in spoken data (cf. Smith 2002; Mindt 2000).

Further negation-related dissimilarities between “school” English and real spoken English arise from the comparison of verb-specific values. Table 50 illustrates the distribution of negated and non-negated progressives across GLN, EG 2000, and BNC/BoE for some individual verbs and highlights a couple of differences between coursebook and spoken English data, but also among the two textbook series (cf. my discussion in 5.7.6). In most cases, as was to be expected from the average figures in GLN and EG 2000, the percentages of negated progressives are significantly lower in the textbooks than in BNC/BoE data (cf. e.g. *doing*, *going*, *having*, or *saying*).

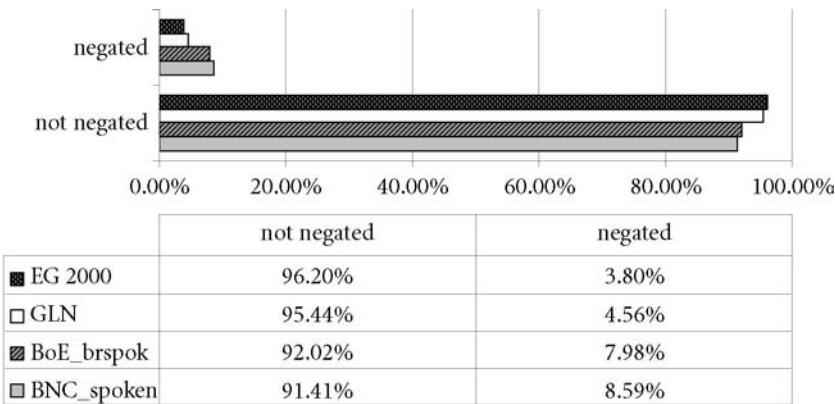
































Figure 137. Shares of negated and non-negated progressives in the BNC_spoken, BoE_brspok, GLN, and EG 2000 datasets

Table 50. Shares of negated and non-negated progressive forms for selected verbs in BNC/BoE (B), GLN (G), and EG 2000 (E)

verb form	B, G, or E	negated	non-negated	distribution graphically illustrated
				■ negated ■ non-negated
asking	B	3.97%	96.03%	
	G	0.00%	100.00%	
	E	20.00%	80.00%	
doing	B	7.27%	92.73%	
	G	0.00%	100.00%	
	E	0.00%	100.00%	
getting	B	9.72%	90.28%	
	G	0.00%	100.00%	
	E	18.75%	81.25%	
going	B	13.94%	86.06%	
	G	4.40%	95.60%	
	E	3.13%	96.88%	
having	B	12.86%	87.14%	
	G	0.00%	100.00%	
	E	0.00%	100.00%	
helping	B	5.88%	94.12%	
	G	0.00%	100.00%	
	E	0.00%	100.00%	
listening	B	16.03%	83.97%	
	G	37.50%	62.50%	
	E	0.00%	100.00%	
making	B	3.85%	96.15%	
	G	16.67%	83.33%	
	E	11.11%	88.89%	
playing	B	10.00%	90.00%	
	G	0.00%	100.00%	
	E	0.00%	100.00%	
saying	B	14.38%	85.62%	
	G	0.00%	100.00%	
	E	0.00%	100.00%	

For many verbs there are no negated instances at all in GLN and EG 2000, even for some of the verbs which showed above-average values of negation in spoken English corpus data. On the other hand, we find very high percentages of negation with *asking*, *getting*, and *making* in EG 2000, and with *listening* and *making* in GLN. These values of 11.11% to 37.50% clearly deviate from the much smaller figures that we determined for the respective verbs in the BNC/BoE progressives. We can hence say that there are some significant inter-corpus differences, not only concerning average

shares of negated progressives, but also among some of the more commonly occurring individual verb forms.

6.1.6 Progressives and other lexical-grammatical phenomena

As discussed in Section 4.3.7, progressives in spoken BrNSE were found to occur to different degrees of frequency in questions, if-clauses, and relative clauses. I will now compare the co-selection shares determined for these phenomena in the BNC/BoE data with the corresponding values from the GEFL TC analysis.

With respect to average percentages, the differences between “school” English and spoken English are significant for all three context features. Figure 138 shows that, while a much larger amount of GLN and EG 2000 progressives occur in interrogative contexts than they normally do in real speech, the attested co-selection patterns of progressives and if-clauses, and of progressives and relative clauses are very much under-represented in the textbooks. Due to these very low token numbers of the last two features in GEFL TC, our verb-specific comparative analysis will only deal with questions.

As the mini-diagrams included in Table 51 illustrate, several verbs show rather different shares of progressives in interrogative and non-interrogative contexts in GLN, EG 2000, and BNC/BoE. The percentages given for most of the selected verbs echo the general trend of the observed overuse of questions in coursebook progressives. Particularly significant in this context are the distributions of *doing*, *playing*, and *thinking*. Rather surprisingly low, however, are shares of questions in the *looking* datasets from GLN and EG 2000, as this particular verb form shows an above-average occurrence in interrogative contexts in actual spoken English. Not covered in the coursebooks either, most probably again due to generally small token numbers of the respective verb forms, is the co-selection of questions and progressives with *bringing*, *eating*, *feeling*,

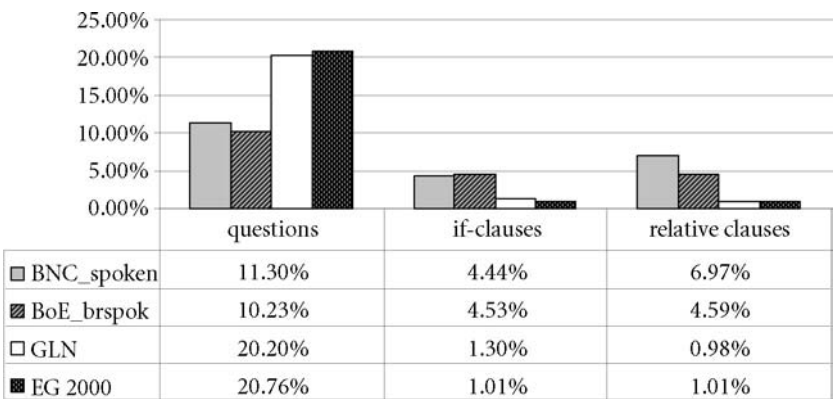


























Figure 138. Shares of questions, if-clauses, and relative clauses in the BNC_spoken, BoE_brspok, GLN, and EG 2000 datasets

Table 51. Shares of progressive forms of selected verbs in interrogative and non-interrogative contexts in BNC/BoE (B), GLN (G), and EG 2000 (E)

verb form	B, G, or E	questions	no questions	distribution graphically illustrated	
				□ questions	▨ no questions
doing	B	16.36%	83.64%		
	G	78.95%	21.05%		
	E	75.00%	25.00%		
going	B	14.90%	85.10%		
	G	26.37%	73.63%		
	E	22.92%	77.08%		
having	B	7.14%	92.86%		
	G	40.00%	60.00%		
	E	0.00%	100.00%		
listening	B	15.38%	84.62%		
	G	0.00%	100.00%		
	E	50.00%	50.00%		
looking	B	13.70%	86.30%		
	G	5.00%	95.00%		
	E	4.76%	95.24%		
playing	B	19.09%	80.91%		
	G	44.44%	55.56%		
	E	21.43%	78.57%		
talking	B	10.11%	89.89%		
	G	0.00%	100.00%		
	E	20.83%	79.17%		
thinking	B	4.96%	95.04%		
	G	16.67%	83.33%		
	E	50.00%	50.00%		

living, remembering, using, and winning, which is typical of spoken BrNSE (cf. Table 26 in 4.5.7).

6.1.7 Progressives and adverbial specification

A final feature that I covered in the context analyses based on “school” English and spoken English data was the co-occurrence of progressive forms and adverbials.

Figure 139 visualises the percentages of adverbial specification of progressives in the four examined datasets. We find quite some deviation among the individual values for all types of adverbials, “time”, “place”, and “others”. On average, time adverbials are clearly overused in the GLN data, while EG 2000 shows a percentage which is similar to the one in BNC_spoken. The shares of place adverbials are rather dissimilar in all four corpora, ranging from 1.41% in BNC_spoken to 8.61% in EG 2000. Only with

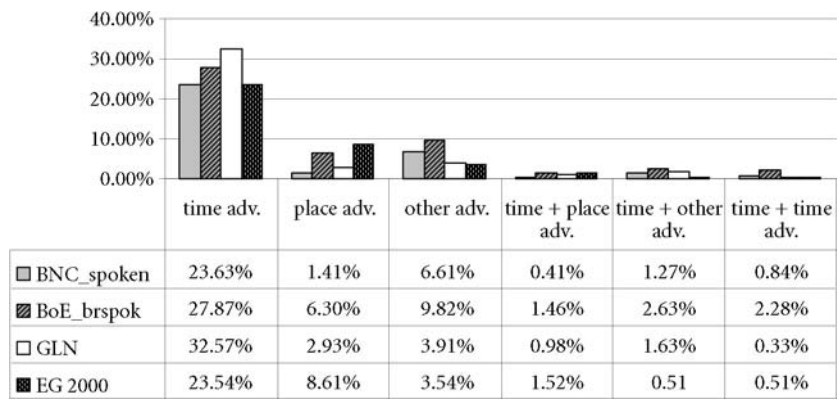


Figure 139. Shares of adverbial specification in the BNC_spoken, BoE_brspok, GLN, and EG 2000 datasets

the so-called “other adverbials” we find a clear trend of under-representation in the coursebook progressives.

Since earlier on we identified a number of verb-specific co-selection patterns, especially of progressives and time adverbials, it is perhaps more profitable to look at individual items now instead of further discussing differences in average distributions. The shares of adverbial specification of selected verbs in GLN, EG 2000, and BNC/BoE data are given and graphically illustrated in Table 52. As the bar charts indicate, there are some significant inter-corpus differences in the verb-specific distribution of time adverbials. Whereas some verbs, such as *asking*, *going*, and *playing*, occur relatively much more often with an adverbial of time in the coursebook progressives than in BNC/BoE data, other forms, e.g. *doing* and *making*, show comparatively low percentages of time adverbial specification in GLN and EG 2000, or none at all as in the case of *taking* and *wearing*. Of course we need to keep in mind that the absolute numbers of occurrence of some progressive types in GEFL TC are rather small, but I would argue that this cannot be taken as an explanation of all the distributional differences illustrated in Table 52 and that a couple of adjustments in the coursebooks might be appropriate.

Dissimilarities between spoken and “school” English can also be found concerning the specification of progressives of different verbs with adverbials of place, such as *here* or *there*. With a number of verbs, e.g. *asking*, *doing*, *sitting*, or *trying*, we observe a significant overuse of place adverbial collocation in EG 2000 (cf. Table 52), a finding which is in accordance with the average shares displayed in Figure 139. However, things are even worse in the GLN data where we find hardly any instances of place adverbial specification of progressive forms at all. Even though both coursebook series misrepresent the actual co-occurrence patterns of progressives and adverbials of place, I consider it better to provide a few instances too many (as in EG 2000) rather than none at all, as is the case with *playing*, *sitting*, and other forms in GLN. Similar

Table 52. Shares of different types of adverbial specification of selected verbs in BNC/BoE (B), GLN (G) and EG 2000 (E)

verb form	B, G, or E	time adverbial	place adverbial	other adverbial	distribution graphically illustrated
					<div> <div></div> time adv. <div></div> place adv. <div></div> other adv. </div>
asking	B	23.18%	0.00%	7.28%	
	G	66.67%	0.00%	0.00%	
	E	40.00%	20.00%	20.00%	
doing	B	24.24%	4.24%	4.85%	
	G	21.05%	21.05%	0.00%	
	E	12.50%	12.50%	3.13%	
getting	B	21.53%	0.00%	9.72%	
	G	37.50%	0.00%	0.00%	
	E	18.75%	0.00%	0.00%	
going	B	19.23%	0.00%	5.77%	
	G	35.16%	2.20%	3.30%	
	E	32.29%	6.25%	1.04%	
making	B	21.15%	3.85%	6.73%	
	G	16.67%	0.00%	0.00%	
	E	11.11%	0.00%	0.00%	
playing	B	31.82%	6.36%	1.82%	
	G	55.56%	0.00%	0.00%	
	E	42.86%	7.14%	7.14%	
sitting	B	24.00%	36.80%	3.20%	
	G	14.29%	0.00%	0.00%	
	E	26.67%	60.00%	0.00%	
taking	B	26.67%	1.67%	5.00%	
	G	0.00%	0.00%	0.00%	–
	E	0.00%	0.00%	0.00%	–
trying	B	22.83%	0.54%	7.61%	
	G	62.50%	0.00%	0.00%	
	E	20.00%	10.00%	10.00%	
wearing	B	22.82%	0.00%	6.71%	
	G	0.00%	0.00%	0.00%	–
	E	0.00%	0.00%	0.00%	–
working	B	37.93%	6.03%	8.62%	
	G	57.14%	0.00%	42.86%	
	E	16.67%	8.33%	0.00%	

observations can be made on “other” adverbials and their collocation with individual progressives in the different corpora. While in EG 2000 and BNC/BoE there is some overlap among the verbs that are specified by adverbials such as *actually*, *if* (in the “whether” sense), or *really* (see e.g. the distributions for *doing*, *playing*, and *trying* in Table 52), we find a rather deviant picture when we compare BNC/BoE and GLN pro-

gressives. In GLN, *asking, doing, getting, making, playing, trying, and wearing* (among others) never co-occur with adverbials of the “other” group, although these verbs repeatedly do so in our real spoken English datasets. On the other hand, the relative frequency of adverbial co-selection with *working* progressives is comparatively high in this coursebook series (see (257) and (258) for examples from GLN).

(257) Joe was working *slowly*, as if he had second thoughts. (GLN)

(258) It is a slow job, but he is working *hard* – much harder than his friends. (GLN)

On the whole, this comparison of GEFL TC and spoken BrNSE corpus data shows that there are certainly a few things concerning the co-selection of progressives and adverbials in the examined EFL textbooks, especially in GLN, that we might consider to change if we want to represent actual language use more appropriately in language teaching. Before summing up some central findings on the contexts of progressives in use and progressives in the books, I shall now take a look at the function side and see whether progressive forms express and refer to comparable things in “school” English and real spoken English.

6.2 Progressives and function phenomena

In analogy with Section 6.1 which dealt with the contexts of progressives in “school” English versus real spoken English, I will now carry out a comparative function analysis and address the question “Do the functions of BNC and BoE progressives correspond to those found in the GLN and EG 2000 datasets?”

The phenomena covered in this comparison are the same that were discussed in the function analysis parts of Chapters 4 and 5, i.e. in Sections 4.4 and 5.6 (cf. the list of features in Table 6 in 4.2.4). The following sections will hence deal with progressives and time reference, progressives and the central function features “repeatedness” and “continuousness”, and with progressives and the identified additional functions “general validity”, “politeness or softening”, “emphasis or attitude”, “shock or disbelief”, “gradual change and development”, “old habits”, “new habits”, and “framing”. Having found a few distributional similarities but predominantly differences between the preferred *contexts* of progressives in the coursebooks and the spoken BrNSE datasets, it will be interesting to see what the situation is like on the *function* side, and whether there are strong discrepancies, or rather strong similarities, in what progressive forms typically express in real-life language use and in EFL textbooks. As mentioned earlier on, I will include both average and verb-specific distributions in my discussion.

6.2.1 Progressives and time reference

In 4.4.1 and 5.6.1 we determined how many of the progressives from BNC_spoken and BoE_brspok, and from GLN and EG 2000 refer to actions or events in the past, present, or future. The resulting distributions that are based on spoken BrNSE and EFL coursebook English respectively are combined and contrasted in Figure 140.

If we compare the percentages which go back to the BNC/BoE data with the GLN and EG 2000 values, we notice some considerable distributional differences. Rather similar are only the shares of past and present time reference progressives in the two spoken English corpora and in GLN. All other coursebook values deviate significantly from the patterns of actual usage. In EG 2000 we observe a preference for progressives with present time orientation and at the same time an under-representation of forms which refer to past actions or events.

Even more significant, though, are the differences concerning future and present/future (“indeterminate”) time reference shares. With 31.60 and 31.65 per cent of all examined concordance lines, GLN and EG 2000 progressives do relatively much more often express futurity than the examples from BNC_spoken and BoE_brspok where values of only 18.55% and 15.68% were found. Also, the percentages of indeterminate cases, i.e. of progressives which express present and/or future time orientation or refer to generally valid actions, are very different in the four datasets. While this is a rather common type of time reference in spoken English (values of 15.99% and 14.38%), it is only rarely expressed in the coursebooks (values of 3.58% and 2.78%). We can thus say that, on average, the time reference choices made in “school” English differ a lot from those made by native speakers of British English. Let us now see whether this also applies for verb-specific distributions or whether, with respect to individual verbs, there are more similarities than differences between BNC/BoE and GLN/EG 2000 progressives.

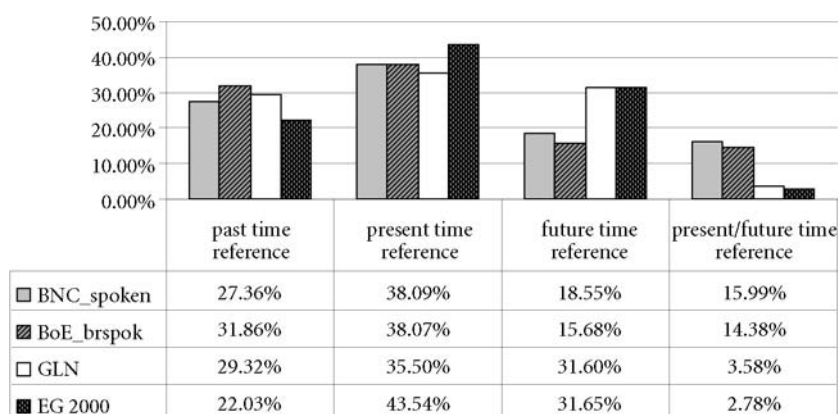


Figure 140. The distribution of time references of progressives in BNC_spoken, BoE_brspok, GLN, and EG 2000

Table 53 illustrates the distribution of time reference types for selected verbs in BNC/BoE, GLN, and EG 2000. The grey-shaded leftmost parts of the included mini-diagrams represent the shares of past time orientation that is expressed in the respective datasets. We notice that, although the average values in BNC_spoken, BoE_brspok, and GLN were rather similar, many of the verb-individual shares differ a lot between BNC/BoE and GLN. This is particularly true for *asking*, *doing*, *playing*, *reading*, *staying*, *thinking*, and *wearing*, where we either find considerably higher percentages of past time reference for GLN or for BNC/BoE progressives. *Asking* and *playing*, for instance, which refer to a past time action or event in roughly 40% of the cases in the spoken English datasets, never do so in GLN. On the other hand, *thinking* and *wearing* show much higher past time reference shares in GLN than in BNC/BoE. The verb-specific differences are even more significant when we look at EG 2000 data where the percentages of past time reference deviate a lot from the determined BNC/BoE values for all selected verbs except *getting*, *listening*, *looking*, and *taking*. It is, for instance, not clear why none of the EG 2000 progressives with *coming*, *helping*, *saying*, and *staying* refer to the past when we consider that these verbs often show this time orientation in spoken BrNSE.





With respect to the shares of progressives that refer to present time actions or events, both coursebook series differ from real-life usage too. Again, we find instances of considerable over- and under-representation with a large number of verb forms, e.g. comparatively high present time reference shares with *asking*, *doing*, *having*, *playing*, *reading*, *staying*, and *trying* in GLN and with *helping*, *saying*, *telling*, and *wearing* in EG 2000. Further differences can be observed when we look at future and present/future time reference. As was to be expected, the latter type is only expressed by a small number of verbs in the coursebooks, and when it is, the percentages are generally rather low (cf. e.g. *doing*, *going*, *making*, and *working*). *Having*, *helping*, *looking*, *reading*, *taking*, and *working* which, among other forms, frequently refer to temporally “indeterminate” events in spoken English, never express this time reference type in GLN and EG 2000. Even though we might expect higher values for future time reference throughout in the GLN and EG 2000 data (given the comparatively very high average shares, cf. Figure 140), the related percentages are higher in the coursebooks than in BNC/BoE only with some verbs (e.g. *going* and *playing*) and lower with several others (e.g. *coming* and *sitting*).

The comparison of verb-specific time reference distributions across corpora provides supportive evidence for the differences discussed with respect to average values, but it also shows that the picture is more complex than Figure 140 indicates. Since there are a number of verb-related deviations in GLN and EG 2000 both from average values and from the shares which go back to BNC/BoE progressives, we need to take into account the behaviour of individual verb forms and their preferences for certain time reference types. On the whole, we can say that the time reference distribution of progressives in the examined coursebooks is far from being a mirror image of the distribution in spoken BrNSE and that a number of changes would be required to bring “school” English more into accordance with actual language usage.

Table 53. Time reference distribution for selected verbs in BNC/BoE (B), GLN (G) and EG 2000 (E)

verb form	B, G, or E	past time ref.	present time ref.	future time ref.	pres/fut time ref.	distribution graphically illustrated
						<div> <div></div> past t. ref. <div></div> present t. ref. <div></div> future t. ref. <div></div> pres/fut t. ref. </div>
asking	B	39.74%	53.64%	1.99%	4.64%	
	G	0.00%	100.00%	0.00%	0.00%	
	E	20.00%	80.00%	0.00%	0.00%	
coming	B	25.00%	13.00%	56.00%	6.00%	
	G	28.57%	28.57%	38.10%	4.76%	
	E	0.00%	71.43%	28.57%	0.00%	
doing	B	22.42%	43.64%	15.15%	18.79%	
	G	10.53%	68.42%	10.53%	10.53%	
	E	15.63%	75.00%	6.25%	3.13%	
getting	B	26.39%	38.89%	16.67%	18.06%	
	G	37.50%	50.00%	0.00%	12.50%	
	E	31.25%	37.50%	31.25%	0.00%	
going	B	13.46%	5.29%	77.40%	3.85%	
	G	9.89%	6.59%	82.42%	1.10%	
	E	2.08%	4.17%	91.67%	2.08%	
having	B	28.57%	21.43%	30.00%	20.00%	
	G	40.00%	60.00%	0.00%	0.00%	
	E	42.86%	28.57%	28.57%	0.00%	
helping	B	38.82%	14.12%	11.76%	35.29%	
	G	50.00%	50.00%	0.00%	0.00%	
	E	0.00%	100.00%	0.00%	0.00%	
listening	B	32.05%	54.49%	1.28%	12.18%	
	G	25.00%	75.00%	0.00%	0.00%	
	E	33.33%	66.67%	0.00%	0.00%	
looking	B	27.40%	50.68%	4.11%	17.81%	
	G	40.00%	60.00%	0.00%	0.00%	
	E	23.81%	71.43%	4.76%	0.00%	
making	B	26.92%	43.27%	12.50%	17.31%	
	G	25.00%	58.33%	8.33%	8.33%	
	E	11.11%	66.67%	22.22%	0.00%	
playing	B	40.00%	28.18%	17.27%	14.55%	
	G	0.00%	55.56%	44.44%	0.00%	
	E	14.29%	64.29%	21.43%	0.00%	
reading	B	48.61%	20.83%	6.94%	23.61%	
	G	25.00%	75.00%	0.00%	0.00%	
	E	20.00%	70.00%	10.00%	0.00%	

Table 53. (continued)

saying	B	34.64%	50.33%	1.31%	13.73%	
	G	50.00%	16.67%	0.00%	33.33%	
	E	0.00%	60.00%	0.00%	40.00%	
sitting	B	47.20%	25.60%	14.40%	12.80%	
	G	42.86%	42.86%	0.00%	14.29%	
	E	73.33%	26.67%	0.00%	0.00%	
staying	B	31.85%	10.37%	47.41%	10.37%	
	G	20.00%	80.00%	0.00%	0.00%	
	E	0.00%	50.00%	50.00%	0.00%	
taking	B	24.17%	17.50%	37.50%	20.83%	
	G	40.00%	20.00%	40.00%	0.00%	
	E	25.00%	37.50%	37.50%	0.00%	
talking	B	35.11%	48.94%	4.26%	11.70%	
	G	40.00%	60.00%	0.00%	0.00%	
	E	50.00%	50.00%	0.00%	0.00%	
telling	B	52.50%	23.13%	11.25%	13.13%	
	G	66.67%	33.33%	0.00%	0.00%	
	E	7.69%	84.62%	7.69%	0.00%	
thinking	B	41.84%	45.39%	2.13%	10.64%	
	G	83.33%	16.67%	0.00%	0.00%	
	E	50.00%	50.00%	0.00%	0.00%	
trying	B	21.74%	54.89%	7.07%	16.30%	
	G	25.00%	75.00%	0.00%	0.00%	
	E	60.00%	40.00%	0.00%	0.00%	
wearing	B	34.23%	26.17%	13.42%	26.17%	
	G	50.00%	50.00%	0.00%	0.00%	
	E	16.67%	83.33%	0.00%	0.00%	
working	B	47.41%	27.59%	6.90%	18.10%	
	G	42.86%	14.29%	28.57%	14.29%	
	E	41.67%	50.00%	0.00%	8.33%	

6.2.2 Progressives and central functions

The perhaps most important part of our function analysis of progressives in spoken English and then later in “school” English was the determination of a central function or central functions of the forms. A careful examination of thousands of BNC_spoken and BoE_brspok concordance lines led to two function features that served well to capture what our progressives expressed: continuousness and repeatedness. The realisation of these features in the BNC and BoE datasets and the analysis of all possible feature combinations brought to light two central functions of progressives in spoken BrNSE, namely (a) the expression of continuous and non-repeated actions or events, and (b) the expression of continuous and repeated actions or events. Before we compare the shares of these central functions in BNC/BoE and GLN/EG 2000, let

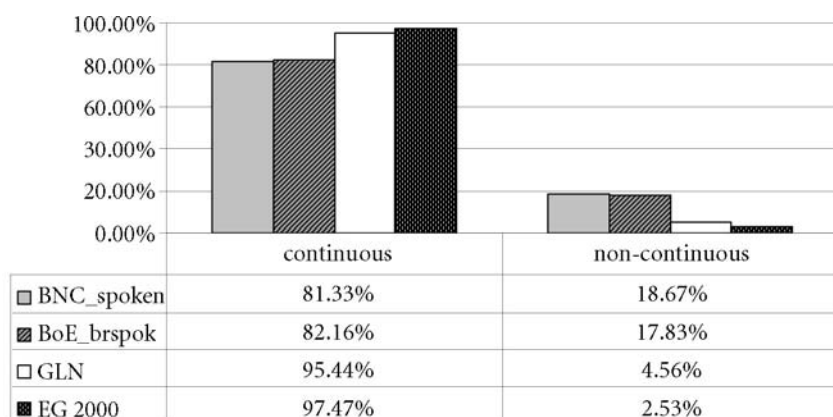


Figure 141. The distribution of progressives referring to continuous and non-continuous actions or events in BNC_spoken, BoE_brspok, GLN, and EG 2000

us first treat the function features separately and take a quick look at the distribution of progressives which refer to continuous and non-continuous, and to repeated and non-repeated actions or events in the four datasets.

In 4.4.2 above we found that although the labels “the continuous” or “continuous aspect” are sometimes used to refer to what we call “progressives”, according to my definition of “continuousness”, roughly 18 per cent of the progressive forms from BNC_spoken and BoE_brspok express non-continuousness. This finding is illustrated once again in Figure 141, which displays the average shares of continuous and non-continuous actions or events in the analysed progressive datasets. Figure 141 shows that, in contrast to BNC and BoE, non-continuousness is a very rare feature in the GLN and EG 2000 progressives where we observed values of only 4.56 and 2.53 per cent (cf. Section 5.6.2). Even more significant are the differences between spoken English and “school” English data when we turn our attention to the second function feature that occurred to be central in the semantic analysis of the BNC_spoken and BoE_brspok progressives (cf. 4.4.2): repeatedness.

As the bar charts in Figure 142 indicate, progressives which refer to repeated actions or events are considerably more frequent in spoken BrNSE than in the course-books. With 9.12 and 9.87 per cent the shares in GLN and EG 2000 are comparatively very low. This means that the textbooks neglect the expression of repeatedness and in turn emphasise the use of progressive forms to refer to non-repeated, single events. For this evident imbalance between “school” English and real spoken English, Figures 143 and 144 provide supportive corpus evidence from BNC_spoken, GLN, and EG 2000. The BNC_spoken concordance sample of *doing* in Figure 143 contains a number of lines in which repeatedness is expressed, whereas the situation is rather different in both textbook concordances. In the *doing* concordances from GLN and EG 2000 the predominant pattern is “(what) are you doing?”, a question which refers to single, continuous actions or events.

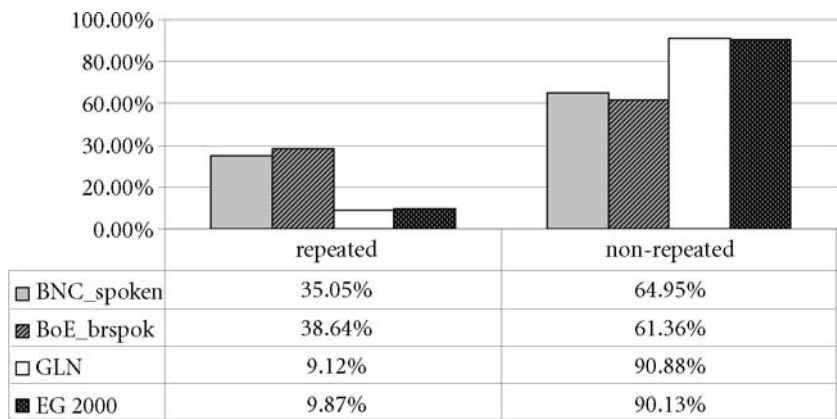


Figure 142. The distribution of progressives referring to repeated and non-repeated actions or events in BNC_spoken, BoE_brspok, GLN, and EG 2000

dominic Kelly. You are once again doing it completely and utterly wrong. at the g, Oh they 're all rushing ahead and doing it and I 'm getting left I know, behind. the sort of thing archaeologists are doing all the time. That 's what we 're trying We have welcomed what the Board are doing, and we 've encouraged them to find new lled Epping Forest and Redbridge are doing, erm, they 've got some lobbying worksh safety conscious. And erm they are doing whatever they can do just to keep people , had rang me and said that they are doing something in Venice and they could in fac re here. There 's things that we are doing like I 've been told by a couple of peop something like, please help, we are doing linguistic research, we want to make rec president. Here are a couple who are doing very awkward cooking over an open fire b g. But in general the people who are doing some of those things, a lot of it 's pro ding and Sheffield and whoever are doing the course Yeah. they should have things at that Well Dennis er Roger 's been doing for it for Leeds North West have n't you through over the year 's I 've been doing this and they say, whether it 's because tury, its something that we 've been doing you know since the beginning of time with e physics and stuff that we 've been doing behind now and just spend one lecture loo . And those of yo you who 've been doing it for some time will now realise how eas

Figure 143. Sample of a BNC_spoken concordance of *doing*

Given these results on continuousness and repeatedness, the distributions found for the four possible feature combinations in the examined sets of spoken English and “school” English progressives are hardly surprising. Figure 145 displays significant differences between BNC_spoken and BoE_brspok on the one hand and GLN and EG 2000 on the other for all functions, in particular for the identified central functions “CF1” (continuous + non-repeated actions/events) and “CF2” (continuous + repeated actions/events; cf. 4.4.3). The figure indicates that only CF1 can be regarded a central function in the coursebooks, which means a considerable simplification and misrepresentation of what progressives in spoken BrNSE actually express. Let us now compare the shares of the four combinations of the features “continuousness” and “repeatedness” for some selected verbs, again across different corpora.

Earlier on we identified some typical relations between the central function feature combinations and individual verbs in the BNC/BoE data (cf. 4.5.10). On the basis of verb-specific function distributions we were able to clearly distinguish “continuous verbs”, e.g. *expecting*, *hoping*, and *wondering*, from “non-continuous verbs”, e.g. *asking*, *calling*, and *paying*. Such relations did not become apparent when we looked at the

you feel it coming, stop what you're doing," the school counselor had advised. 'Hello, Jenny. What are you doing? Oh, you're reading. JENNY I'm reading a really good book. DEBBIE Hi, Jenny. What are you doing? DEBBIE I'm phoning you! JENNY Very interesting. My friend Hey! What are you doing? Run, Sita! What? Stop that girl! MRS SNOW Fine. JENNY What are you doing? MRS SNOW I'm making dinner. What at came out. 'Hello, Sita. What are you doing here?' asked Debbie. 'Hi, Debbie! I LAYLA What are you doing on Saturday morning, Tim? Can you play bandad,' Sheena asked, 'what were you doing when you had your accident on that c / magazine. 'Oh, Stiff, what were you doing with it?' 'Didn't know it was there, The others move over. RUST What you doing there? BEN Just ... lying down a bit parts to search Rust. RUST What you doing? RUDI (to the rest of the gang) Clear j? RUST I don't know. MANNY What you doing? BEN Sorry? MANNY Are you with him? on TV tonight?' Ray said he was doing a school project and the officer was questions. I wanted to know who was doing these terrible things. POLICE Who did Oh, he's fine. Mum, what's Trundle doing? MRS SNOW He's sleeping. JENNY And find your animals. See what they're doing. Are they sleeping? Or eating? Or playing? What are they doing? They're building a tree house with ol David: What are you doing, Mark? Mark: I'm drawing a birthday the shop. "Hello, Becky. What are you doing here?" Robert asked. "I'm just taking n in the kitchen. Becky: What are you doing? Robert: I'm making a salad. Simon: eyes and ran after her. "What are you doing here?" I asked. I decided it was t our hockey stick, too. W-what are you doing tonight, Jenny? I'm going to the you "Excuse me," Becky said. "What are you doing?" "Me?" the man said. He stood up, t nt behind the prisoner. "What are you doing with that rope?" shouted the Sheriff. his glider. "How you doing?" the kid asked him. "Fine," Steve s turned to her. He looked her you were doing," he said. "Listen, I'm sorry straight he wondered for a moment why he was doing this; then Joe's instructions took nd watch the two boys. "What are they doing?" asks Robert. Simon gets his purse ble bully," says Becky. "What's Simon doing with his brother?" "I don't know," s a good idea, Mark. Let's what's she doing here? make a castle with water pour hen Becky said, "Hey, what's that man doing over there?" "Which man?" Robert ask talk about the future! What is Jenny doing after school on Monday? She's play bert: Hello, Mrs Richards. What's Jack doing? Mrs R.: He's feeding Lenny. Lenny's

Figure 144. Samples of EG 2000 (above) and GLN concordances of *doing*

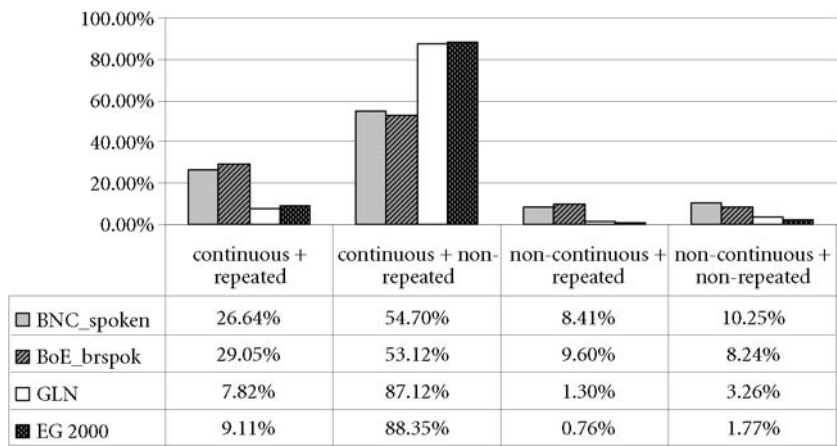


Figure 145. The distribution of central function feature combinations in BNC_spoken, BoE_brspok, GLN, and EG 2000

progressives from the two GEFL TC subcorpora. In GLN and EG 2000 the picture was found to be far less complex than in BNC/BoE, and almost all verbs expressed continuousness, mainly in combination with non-repeatedness. By means of clarification, the respective graphic illustrations of these function distributions in the analysed corpora are compared in Table 54.

Table 54. The distribution of verb form tokens across the four combinations of the central function features “continuousness” and “repeatedness” for selected verbs in BNC/BoE (B), GLN (G) and EG 2000 (E)

verb form	B, G, or E	cont. + rep.	cont. + non-rep.	non- cont. + rep.	non- cont. + non-rep.	distribution graphically illustrated			
						■ cont.+rep.	▨ cont.+non-rep.	░ non-cont.+rep.	■ non-cont.+non-rep.
asking	B	0.00%	0.00%	32.45%	67.55%				
	G	0.00%	0.00%	66.67%	33.33%				
	E	0.00%	0.00%	20.00%	80.00%				
coming	B	36.00%	64.00%	0.00%	0.00%				
	G	9.52%	90.48%	0.00%	0.00%				
	E	0.00%	100.0%	0.00%	0.00%				
doing	B	59.39%	40.61%	0.00%	0.00%				
	G	15.79%	84.21%	0.00%	0.00%				
	E	6.25%	93.75%	0.00%	0.00%				
getting	B	40.28%	59.72%	0.00%	0.00%				
	G	25.00%	75.00%	0.00%	0.00%				
	E	25.00%	75.00%	0.00%	0.00%				
going	B	23.56%	76.44%	0.00%	0.00%				
	G	5.49%	94.51%	0.00%	0.00%				
	E	8.33%	91.67%	0.00%	0.00%				
having	B	44.29%	55.71%	0.00%	0.00%				
	G	0.00%	100.0%	0.00%	0.00%				
	E	0.00%	100.0%	0.00%	0.00%				
helping	B	76.47%	23.53%	0.00%	0.00%				
	G	0.00%	100.0%	0.00%	0.00%				
	E	0.00%	100.0%	0.00%	0.00%				
listening	B	12.18%	87.82%	0.00%	0.00%				
	G	0.00%	100.0%	0.00%	0.00%				
	E	0.00%	100.0%	0.00%	0.00%				
looking	B	27.40%	56.85%	5.48%	10.27%				
	G	0.00%	95.00%	0.00%	5.00%				
	E	0.00%	95.24%	0.00%	4.76%				
making	B	36.54%	39.42%	6.73%	17.31%				
	G	0.00%	91.67%	0.00%	8.33%				
	E	0.00%	100.0%	0.00%	0.00%				
playing	B	30.00%	70.00%	0.00%	0.00%				
	G	0.00%	100.0%	0.00%	0.00%				
	E	14.29%	85.71%	0.00%	0.00%				
reading	B	34.72%	65.28%	0.00%	0.00%				
	G	0.00%	100.0%	0.00%	0.00%				
	E	0.00%	100.0%	0.00%	0.00%				

Table 54. (continued)

saying	B	24.84%	75.16%	0.00%	0.00%	
	G	33.33%	66.67%	0.00%	0.00%	
	E	40.00%	60.00%	0.00%	0.00%	
sitting	B	15.20%	80.00%	0.00%	4.80%	
	G	14.29%	85.71%	0.00%	0.00%	
	E	6.67%	93.33%	0.00%	0.00%	
staying	B	11.11%	88.89%	0.00%	0.00%	
	G	0.00%	100.0%	0.00%	0.00%	
	E	0.00%	100.0%	0.00%	0.00%	
taking	B	22.50%	37.50%	27.50%	12.50%	
	G	0.00%	100.0%	0.00%	0.00%	
	E	12.50%	87.50%	0.00%	0.00%	
talking	B	15.96%	84.04%	0.00%	0.00%	
	G	0.00%	100.0%	0.00%	0.00%	
	E	8.33%	91.67%	0.00%	0.00%	
telling	B	26.88%	73.13%	0.00%	0.00%	
	G	33.33%	66.67%	0.00%	0.00%	
	E	0.00%	100.0%	0.00%	0.00%	
thinking	B	15.60%	84.40%	0.00%	0.00%	
	G	0.00%	100.0%	0.00%	0.00%	
	E	0.00%	100.0%	0.00%	0.00%	
trying	B	30.43%	69.57%	0.00%	0.00%	
	G	12.50%	87.50%	0.00%	0.00%	
	E	0.00%	100.0%	0.00%	0.00%	
wearing	B	39.60%	60.40%	0.00%	0.00%	
	G	0.00%	100.0%	0.00%	0.00%	
	E	0.00%	100.0%	0.00%	0.00%	
working	B	77.59%	22.41%	0.00%	0.00%	
	G	42.86%	57.14%	0.00%	0.00%	
	E	50.00%	50.00%	0.00%	0.00%	

The first verb form in the list, *asking*, presents the aforementioned exception from the predominant “continuousness pattern” in the coursebook data. Just like the concordance lines from BNC/BoE, all GEFL TC examples express non-continuousness, though GLN clearly prefers “repeated” contexts (cf. e.g. (259)), while the combination “non-continuous + non-repeated” strongly prevails in EG 2000, as in (260).

(259) People are beginning to ask questions. Daryl: Really? What kind of questions *are they asking*? (GLN)

(260) You’re probably wondering, Doug, why I’m not *asking* anyone around here for help. (EG 2000)

If we look at the distributions for the remaining 21 verbs in Table 54, we notice that the major difference between BNC/BoE and coursebook data relates to the feature combination “continuous + repeated”. Especially with respect to the verb forms *coming*, *doing*, *having*, *helping*, *looking*, *playing*, *wearing*, and *working*, progressives that refer to continuous and repeated actions or events are largely under-represented in GLN and EG 2000. Instead of presenting this function which is rather common in spoken BrNSE, the textbooks mainly use progressives that express non-repeatedness, as the GEFL TC examples in (261) to (263) show.

- (261) “Let’s hope Mr Smith comes soon,” says Jack. “He’s *coming*! That’s his car now.” (GLN)
- (262) SAMUEL: Niels is cooking Sunday lunch this weekend. Nadja *is playing* volleyball, and I’m going to the cinema. (EG 2000)
- (263) “Did you notice his clothes?” asked Mr. Chung. “He *was wearing* red and black. Those are the colors of that gang, the Jackals. They hate people like us.” (GLN)

Also worth pointing out are the distributional differences that apply to *making* and *taking*. In our BNC_spoken and BoE_brspek datasets, though not in GLN and EG 2000, these forms often express non-continuousness. Progressives like those in (264) and (265) do not occur in any of the coursebooks. As we have been able to show, there are some significant deviations from actual language use in EFL textbook language concerning the central functions of progressive forms, both related to average and to verb-specific values. Next I will investigate whether the distributions of additional progressive functions also show significant discrepancies across spoken BrE and coursebook corpora.

- (264) the too little sugar where your person has been stabilized and they’re *taking* insulin injections. (BNC_spoken)
- (265) so what you’re doing is you’re *taking* an average tax you’re # removing an average tax of twenty five per cent (BoE_brspek)

6.2.3 Progressives and additional functions

In order to highlight how the average values determined for additional functions in GLN and EG 2000 compare to the frequencies of occurrence in BNC_spoken and BoE_brspek, the results obtained in the related analyses of spoken English and coursebook data (cf. 4.4.5 and 5.6.4) are combined in one diagram, Figure 146.

Figure 146 clearly indicates a few significant inter-corpus differences concerning the distribution of some additional progressive functions. Most apparent are the differences in the case of “general validity”, which was found to be the most frequent additional function in BNC_spoken and BoE_brspek. The shares of 3.58% in GLN and 5.82% in EG 2000 lie far below the values we determined for this function in

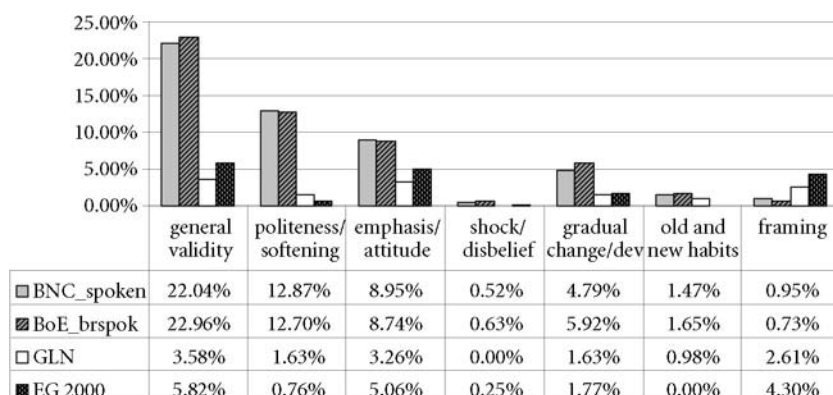


Figure 146. The distribution of additional progressive functions in BNC_spoken, BoE_brspok, GLN, and EG 2000

the spoken English datasets (>22%). Also strongly under-represented in the textbooks are the second, third, and fourth most common additional functions of progressives in BNC/BoE, namely “politeness or softening”, “emphasis or attitude”, and “gradual change and development”. While the percentages for these functions range from 0.76% to 5.06% in GLN and EG 2000, they lie between 4.79% and 13.87% in BNC/BoE.

The only additional function which shows higher percentages in the coursebook progressives than in spoken BrNSE is “framing”. It was already mentioned above that this function is given a lot of attention in the grammar sections of GLN and EG 2000 and that the explanations and examples provided in parts of the textbook series might convey the impression that the PastProg is always used in framing contexts (cf. 5.8.3). However, when we look at the use of PastProgs in GLN and EG 2000, we note that this is not the case. Nevertheless we may consider it questionable why the framing function is emphasised in school grammars and repeatedly used in the coursebook texts when it is actually very uncommon in real spoken English.

If we now turn our attention to the connections between additional progressive functions and individual verbs, we find further differences between BNC/BoE and GEFL TC data. Table 55 combines relevant values from Tables 30 and 46 (cf. 4.5.11 and 5.7.11). Perhaps most significant are the differences among the percentages listed in the “general validity” column of Table 55. If we compare the values for BNC/BoE with those for GLN and EG 2000, we see immediately that, with the exception of *saying*, all verb forms show much higher shares of this additional function in spoken BrNSE than in coursebook English. The differences are particularly significant in the case of *having*, *helping*, *reading*, *taking*, *trying*, and *wearing*, verbs which often express general validity in BNC/BoE progressives but never do so in the respective data from the GEFL TC subcorpora.

Due to the generally rather low frequencies of occurrence of these functions in the GLN and EG 2000 datasets, there is not a great deal to say about “politeness

or softening”, “gradual change and development”, and “old/new habits”. Worth noting, however, are the attested preferences of the forms *asking*, *saying*, *telling*, *thinking*, and *trying* for the “politeness” function. With the exception of *thinking* which expresses this function in a small number of GLN progressives, these preferences are unfortunately not mirrored in the coursebooks. One shared pattern of textbook and spoken English data can be found with *asking*, though, where the shares of the “emphasis” function are considerable in all corpora. Apart from this similarity, we observe a few distributional differences concerning the expression of emphasis or attitude and find an over-representation of that function with *making* and *thinking* in GLN and with *saying*, *taking*, *talking*, and *thinking* in EG 2000, but on the other hand an under-representation with *telling* in both coursebook corpora.

As the very low corresponding values from BNC/BoE show, “framing” is expressed comparatively too often by GLN progressives with *helping*, *looking*, *sitting*, and *thinking* and by EG 2000 progressives with *getting*, *listening*, *looking*, *reading*, *sitting*, *taking*, *talking*, *trying*, and *working*. In this context, a word of caution appears appropriate. As can be guessed from some of the percentages given in Table 55 for verb-function combination, frequencies of occurrence of these specific co-selection types are often rather small. If we only looked at individual verbs from our list in isolation, it would hence probably not be entirely safe to make claims about misrepresentations in the coursebooks. However, as we could see especially in the case of “general validity” and “framing”, we often observe the same kind of corpus-internal differences with a number of verbs. I think that such “cumulative evidence” certainly shows us important tendencies which hint at some deviant patterns of language usage. I have to admit, though, that some of the results reported on in previous sections of this chapter might be more representative and that some of the differences found with respect to contexts, time references, and central functions may be more significant than the results related to additional functions. In the following section, the most central findings of the comparison of spoken English and “school” English data will be summarised and discussed.

6.3 Summary of the findings: Progressives in use vs. progressives in the books

Recently, a number of corpus linguistic researchers have shown an interest in comparative analyses of some central lexical-grammatical features in natural language corpora and in language teaching materials (cf. Barlow 1996; Bland 1988; Conrad 2000, 2004; Grabowski & Mindt 1994, 1995; Gilmore 2004; Haase 1995; Ljung 1990; Lorenz 2002; Mindt 1992, 1996; Römer 2004a, 2004b; Schlüter 2002a; Tesch 1990; Wong 2002). No matter what language features they investigated, all of them observed considerable mismatches between real native-speaker English as captured in different corpora, and “school” English as presented in EFL/ESL coursebooks or in learner’s grammars and usage books. It was repeatedly found that in these reference works for learners “very

Table 55. Frequencies of occurrence of additional progressive functions across verb forms in BNC/BoE (B), GLN (G), and EG 2000 (E)

verb form	B, G, or E	general validity	politeness/softening	emphasis/shock	gradual change	old and new habits	framing
asking	B	7.28%	27.15%	49.67%	0.00%	0.00%	0.00%
	G	0.00%	0.00%	66.67%	0.00%	0.00%	0.00%
	E	0.00%	0.00%	60.00%	0.00%	0.00%	0.00%
coming	B	12.00%	1.00%	0.00%	8.00%	1.00%	1.00%
	G	4.76%	0.00%	0.00%	9.52%	0.00%	0.00%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
doing	B	28.48%	0.00%	0.61%	0.00%	0.00%	0.61%
	G	10.53%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	3.13%	0.00%	0.00%	0.00%	0.00%	0.00%
getting	B	22.22%	2.78%	0.69%	43.74%	0.00%	0.00%
	G	12.50%	0.00%	0.00%	25.00%	0.00%	0.00%
	E	18.75%	0.00%	0.00%	37.50%	0.00%	12.50%
going	B	16.83%	2.88%	1.92%	1.92%	0.00%	1.44%
	G	0.00%	0.00%	1.10%	0.00%	2.20%	1.10%
	E	4.17%	0.00%	0.00%	0.00%	0.00%	0.00%
having	B	20.00%	1.43%	0.00%	0.00%	0.00%	1.43%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
helping	B	40.00%	2.35%	1.18%	1.18%	0.00%	0.00%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	50.00%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
listening	B	8.97%	5.77%	10.90%	0.00%	0.00%	1.28%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	16.67%
looking	B	16.44%	0.00%	1.37%	0.68%	1.37%	0.00%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	5.00%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	4.76%
making	B	20.19%	15.38%	7.69%	7.69%	0.00%	0.96%
	G	0.00%	0.00%	25.00%	0.00%	0.00%	0.00%
	E	11.11%	0.00%	0.00%	0.00%	0.00%	0.00%
playing	B	17.27%	0.00%	0.91%	0.91%	0.00%	1.82%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	7.14%	0.00%	0.00%	0.00%	0.00%	0.00%
reading	B	26.39%	0.00%	4.17%	0.00%	0.00%	4.17%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	10.00%
saying	B	18.30%	58.17%	9.15%	0.00%	0.00%	0.00%
	G	33.33%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	40.00%	0.00%	100.00%	0.00%	0.00%	0.00%

Table 55. (continued)

sitting	B	17.60%	0.00%	0.00%	0.00%	0.00%	2.40%
	G	14.29%	0.00%	0.00%	0.00%	0.00%	28.57%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	20.00%
staying	B	11.85%	0.00%	2.22%	0.00%	14.81%	0.74%
	G	0.00%	0.00%	0.00%	0.00%	20.00%	0.00%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
taking	B	23.33%	0.83%	7.50%	9.17%	0.00%	0.00%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	0.00%	0.00%	25.00%	0.00%	0.00%	12.50%
talking	B	14.36%	0.00%	3.72%	0.00%	0.00%	0.53%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	8.33%	0.00%	16.67%	0.00%	0.00%	4.17%
telling	B	16.88%	30.00%	31.25%	0.00%	0.00%	0.63%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	7.69%	0.00%	23.08%	0.00%	0.00%	0.00%
thinking	B	15.60%	54.61%	3.55%	0.00%	0.00%	0.00%
	G	0.00%	33.33%	16.67%	0.00%	0.00%	16.67%
	E	0.00%	0.00%	50.00%	0.00%	0.00%	0.00%
trying	B	25.54%	15.22%	5.98%	1.63%	0.00%	0.54%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	10.00%
wearing	B	32.31%	0.00%	7.38%	0.00%	0.00%	0.00%
	G	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
working	B	55.17%	1.72%	1.72%	0.00%	0.00%	0.00%
	G	28.57%	0.00%	0.00%	0.00%	0.00%	0.00%
	E	8.33%	0.00%	0.00%	0.00%	0.00%	8.33%

often the rule does not reflect the evidence of language use” (Tognini-Bonelli 2001: 10) and that

a piece of English, put forward as a model of a kind of English and probably written for the purpose, does not show the same features as are found in appropriate selections from a corpus. (Sinclair 2004b: 6–7)

Conrad (2004), for instance, observed that the treatment of the linking adverbial *though* in several best-selling ESL textbooks differs a lot from its actual usage in different language registers (conversation and academic prose) and that the descriptions in the pedagogic materials are misleading. Very similar observations have been made by Mindt, Schlüter, and myself, so it can be described as a central finding of such comparative studies that, as very appropriately stated by Mauranen (2004: 90), “what is taught as functional language use is not necessarily in agreement with what is frequent in the language, or even appears at all.”

The results of our large-scale corpus-driven analysis of progressives in spoken BrNSE and “school” English hit the same note. By comparing findings based on data from BNC_spoken and BoE_brspok with those obtained in an EFL coursebook corpus analysis, we were able to show that progressive forms show largely different average and verb-specific distributions of a number of function and context features in “school” English and real spoken English. In other words, both types of English favour different presentation patterns of progressives. Here is a list of the most significant pattern deviations. Most of the listed deviations apply to both coursebook series, GLN and EG 2000, very few may only apply to either GLN or EG 2000 (see 6.1 and 6.2 for information on inter-coursebook differences).

On the context side:

1. Tense form distribution (for details cf. 6.1.1)
 - over-representation of PresProgs with a number of verbs, e.g. with *asking*, *going*, *playing*, and *staying*
 - under-representation of PastProgs and PresPerfProgs with a number of verbs, e.g. with *asking*, *going*, *playing*, and *staying*
2. Tense form contractions (for details cf. 6.1.1)
 - general under-representation of short forms of TO BE
 - under-representation of *'re V-ing* and *'s V-ing*
 - over-representation of *are V-ing*
3. Co-selection of progressives and subjects (for details cf. 6.1.2)
 - overuse of the pronouns *he* and *she* in subject position of progressives
 - underuse of *I*, *it*, *we*, and *they* as subjects
 - overuse of personal names as subjects
 - many verb-specific deviations from BNC/BoE concerning the shares of different pronouns in subject position, e.g. overuse of *you + doing* and underuse of *I/we + doing* in the coursebook data
4. Co-selection of progressives and objects (for details cf. 6.1.3)
 - over-representation of *the + noun (group)* and *a + noun (group)* as objects of progressives
 - over-representation of *a + noun (group)* in object position
 - misrepresentation of several verb-object patterns, e.g. overuse of *going + preposition + the + noun (group)*, underuse of *having + a + noun (group)*, and *telling + you*
 - missing of some typical verb-object patterns, e.g. *asking + you*, *bothering + me*, *liking + it*, *listening + to + the + noun (group)*, *telling + me*

5. Co-selection of progressives and prepositions (for details cf. 6.1.4)
 - overuse of a couple of verb-preposition collocations, e.g. *coming back*, *getting at*, *looking at*, and *working on*
 - underuse (or missing) of some typical BNC/BoE verb-preposition collocations, e.g. *agreeing with*, *dealing with*, *ringing up*, and *setting up*
6. Progressives in negative and interrogative contexts, if-clauses, and relative clauses (for details cf. 6.1.5 and 6.1.6)
 - under-representation of negated progressives
 - several verb-related deviations from spoken English with respect to negation, e.g. underuse of negation with *doing*, *going*, *having*, and *saying*; overuse of negation with *asking*, *getting*, *listening*, and *making*
 - over-representation of interrogative contexts
 - under-representation of progressive-if-clause and progressive-relative-clause co-selection
 - a number of verb-specific discrepancies concerning shares of questions, e.g. overuse of interrogative contexts with *doing*, *playing*, and *thinking*; underuse with *looking*; missing co-selection of progressives and questions with *bringing*, *eating*, *feeling*, *living*, *remembering*, *using*, and *winning*
7. Co-selection of progressives and adverbials (for details cf. 6.1.7)
 - overuse of time adverbials
 - underuse of “other” adverbials
 - several inter-corpus differences in the verb-related distribution of adverbials, e.g. overuse of time adverbials with *asking*, *going*, and *playing*; underuse of time adverbials with *doing*, *making*, and *wearing*; overuse of place adverbial collocation with *asking*, *doing*, *sitting*, and *trying*

On the function side:

1. Time reference distribution (for details cf. 6.2.1)
 - general over-representation of progressives with present time orientation
 - under-representation of progressives with past time orientation
 - overuse of future time reference progressives
 - underuse of progressives with present/future (“indeterminate”) time orientation
 - a lot of verb-individual differences with respect to time reference distributions between BNC/BoE and GEFL TC data, e.g. under-representation of past time reference with *asking*, *coming*, *helping*, *playing*, *saying*, and *staying*; over-representation of past time reference with *sitting*, *thinking*, *trying*, and *wearing*; overuse of present time orientation with *asking*, *doing*, *having*, *playing*, *reading*, *saying*, *telling*, and *wearing*; under-representation of “indeterminate” time reference with a large number of verbs (e.g. *helping*, *looking*, and *work-*

ing); overuse of future time reference with *going* and *playing*, underuse with *coming* and *sitting*

2. Distribution of central functions (for details cf. 6.2.2)

- general under-representation of the central function features “non-continuousness” and, especially, “repeatedness”
- underuse of the central function “CF2” (continuous + repeated actions/events), in particular with the verb forms *coming*, *doing*, *having*, *helping*, *looking*, *playing*, *wearing*, and *working*
- over-representation of non-repeatedness with the aforementioned verbs
- under-representation of non-continuousness with *making* and *taking*

3. Distribution of additional functions (for details cf. 6.2.3)

- underuse of the additional functions “general validity”, “politeness or softening”, “emphasis or attitude”, and “gradual change and development”
- overuse of the additional function “framing”
- underuse (or missing) of “general validity” progressives, especially with *having*, *helping*, *reading*, *taking*, *trying*, and *wearing*
- underuse (or missing) of “politeness” progressives, especially with *asking*, *saying*, *telling*, *thinking*, and *trying*
- over-representation of the “emphasis” function with *making*, *saying*, *taking*, *talking*, and *thinking*; under-representation with *telling*
- overuse of “framing” progressives especially with *getting*, *helping*, *listening*, *looking*, *reading*, *sitting*, *taking*, *talking*, *thinking*, *trying*, and *working*.

This long list of attested differences between “school” English and real spoken English documents that language descriptions and presentations in teaching materials may differ considerably from the facts we find in corpora. Evidently, in the ELT classroom, the language which is taught is not the same as the language which is used by a very large number of native speakers. Put differently, pedagogical descriptions are rather inadequate and the classroom does not quite meet reality. We now need to consider how we could react to these findings and how we could (or should?) respond to the observed inadequacies related to language presentation in pedagogical materials. Chapter 7 will deal with these questions and discuss some pedagogical implications of this corpus-driven study.

CHAPTER 7

Pedagogical implications

True facts, textbooks, teaching

I have just pointed out that, in the use and presentation of progressive forms, there exist a large number of discrepancies between coursebook and corpus, i.e. between what I call “school” English and real spoken English. My comparative corpus-driven analysis of large amounts of data uncovered some considerable inadequacies in pedagogical description and, especially with respect to the functions of progressives, showed a rather simplified picture in the teaching materials.

The question is now in what way we could, and how we should in fact, respond to these observations. I would propose that there are at least three options how to react:

1. Ignore the findings.
2. Make drastic changes in the textbooks so that they present language exactly as it is typically used by native speakers.
3. Make moderate changes in the textbooks and pay attention to the most salient facts we find in natural language corpora.

Option no. 1 would imply that, despite the mismatches we found, we would leave everything as it is, not accept the corpus evidence, and not change anything in the teaching materials. I think, however, that the attested differences between BNC/BoE and GEFL TC data are too significant to be ignored. Also, as my account of this linguistic approach indicated (cf. 2.1), working *corpus-driven* means taking the data seriously, which in turn means that we must not ignore the evidence (cf. Tognini-Bonelli 1996a: 113). As appropriately stated by Stubbs (2000: 17), “[n]ew forms of data have to be carefully assessed and interpreted, but it would be absurd to ignore them.” This is fully in line with the position statement given by the “father” of corpus-driven linguistics, John Sinclair. His statement includes the postulates “[a]ccept the evidence” and “[r]eflect the evidence” (Sinclair 1991: 4), which leads us to the proposed options no. 2 and no. 3.

If we want to reflect the evidence, certain changes in language descriptions (in our case in pedagogical descriptions) are required. For the time being, I consider it rather difficult to decide how drastic these changes ought to be. Unless further comparative empirical studies on progressives (and other language phenomena) in “school” English and real English have been carried out to back up our findings, it is perhaps not entirely safe to call for a radical revision of the textbooks and an exact depiction of real English in teaching materials. At present, my preferred option would hence be the third one. I

would suggest that a number of changes be made, but these changes should be careful and well-considered. As research in the field of corpus-driven linguistics advances, its influence on textbooks and other (pedagogical) reference works will certainly become more substantial.

Perhaps less difficult to deal with than the “How drastic...?” question is the question about the type of changes that ought to be made. The present corpus-driven study has highlighted a number of aspects related to the use and teaching of progressives which we need to pay more attention to if we want to reflect corpus evidence and adjust pedagogical language descriptions to actual usage. I think that such an adjustment could increase economy and efficiency in, and lead to an improvement of language teaching (I suspect that a lot of precious time is wasted in the ELT classroom by teaching not so common language patterns and rather infrequent meanings of items). What is more, it might make life easier for the language learner. Like Glisan & Drescher (1993:25) I believe that “the grammar in today’s textbooks should be revised to reflect more accurately the target language as it is used currently in native speaker speech” (cf. also Turner 1981:200).

What I am going to do in the following is respond to the major mismatches found between English in use and English in German ELT settings. I will also raise some more general points that I consider relevant at the interface of corpus linguistics and language teaching, and deal with such issues of authenticity and communicative competence. On the basis of the discussion of some pedagogical implications of my corpus-driven analyses, I will then try to develop a new concept of teaching progressives, incorporating my empirical evidence.

7.1 Corpus-driven linguistics and language teaching

“What does corpus-driven linguistics have to offer to language teaching?” The increasing interest of researchers and practitioners in pedagogical corpus applications, as documented by the growing number of publications in the field⁸³ and by the growing number of people who attend specialist conferences, such as the biennial Teaching and Language Corpora conference (TaLC), or workshops and courses dedicated to corpora use in language teaching, certainly indicates that the answer to this question can only be “A lot.” Corpus-driven work, which takes corpus data seriously, provides us with much invaluable information on the nature of language. It shows us how language is typically used in natural discourse and, among other things, reveals what items are frequent in which text types, which words are likely to occur in combination, or what meaning is the most common of several senses an item may have. Naturally, these points are – or should be – of major interest to everyone whose aims are to understand the language better and to help learners master it with less difficulty and greater confidence. The more learners know about how a language works, the more confident they can be about its use.

Given these obvious strengths of corpus-driven research in a pedagogical context, it is surprising that, as Mindt (1996:232) observes, “teaching materials and especially textbooks for teaching English as a foreign language” have apparently not been very much affected by the results of such research up to now. And indeed, the findings of my comparative analysis show that, even several years after Mindt made this observation, real-life language with its typical usage patterns has not quite entered the coursebooks. With large corpora, sophisticated software packages, and the analytical techniques of corpus linguistics at hand, I do not see what might hinder us from making for better quality information in the coursebooks and what may keep us, in Mindt’s (1996:247) words, from “bring[ing] textbooks for teaching English as a foreign language into closer correspondence with actual English.” To quote Aarts (2000:34) who also stresses the strengths of corpora in language description, “there are no excuses any more; we are now in a position to write grammars based on empirical fact.” Corpus-driven linguistics certainly has the potential to make a significant contribution to language pedagogy and to strongly affect teaching materials in a positive sense – we just have to let it.

7.2 The case for more authenticity in the classroom

In 1984 at an international conference on “Progress in English Studies”, John Sinclair asked the question “Why are so many of our model sentences, dialogues and texts [in teaching materials] clearly artificial?” (Sinclair 1985:254) Today, about 20 years on, this question still preoccupies a large number of scholars in corpus and applied linguistics.

The issue of whether to use artificial, invented or natural, authentic language examples in coursebooks and reference works provides plenty of food for discussion among linguists (see in particular the recent controversies in the journals *Applied Linguistics* and *ELT Journal*; Carter 1998; G. Cook 1998, 2001, 2003; V. Cook 2002; Stubbs 2001a; Widdowson 1996, 2000 2001; cf. also Breen 1985 for an earlier article on that topic).

As early as 1935, Firth expressed doubts about the use and usefulness of invented language in reference materials and observed that many of the examples in grammar books are “just nonsense” from a semantic point of view (1957:24). To illustrate his point, he quotes the sentence “I have not seen your father’s pen, but I have read the book of your uncle’s gardener.” Comparatively odd utterances or exchanges can also be found in the most recently published coursebooks for ELT, including the two series my “school” English analysis is based on. The two short dialogues in (266) and (267), taken from GLN and EG 2000, may serve as illustrative evidence.⁸⁴

(266) Becky: What are you doing?

Robert: I’m making a salad.

Simon: And you’re listening to Radio 1. Robert: Well, no. I’m not listening to Radio 1. It’s Radio Nottingham. I’m listening to my mum.

Becky: Your mum? But she isn't at home.

Robert: No, she isn't. (GLN)

(267) MR SNOW: Hi, Sal.

SALLY: Hi, Dad. Is Tracy waiting in the car?

MR SNOW: No, she isn't. She's at home. Where's your sister? Is she packing?

SALLY: No, she isn't.

MR SNOW: Is she reading?

SALLY: Yes, she is.

MR SNOW: Oh no. (EG 2000)

This type of English with its somewhat artificial sound is put forward by the course-books as some kind of simplified standard English, a model for learner language reception and production. A very similar observation has been made by de Beaugrande. He quotes examples he “found in real textbooks” (e.g. “I hear with my ears.” and “I like a bedroom with green walls.”) and argues that

[i]nsofar as this variety of standard English is not found in the everyday language use of native speakers, I propose to label it *non-authentic English*, and to question whether exposure to it, however prolonged, can be realistically expected to build fluency for the *authentic English* used by native speakers.

(de Beaugrande 2001a: 117; italics in original)

In another unpublished article, de Beaugrande (2001b) further discusses this point and states that in an ELT setting we may “hinder the development of fluency by excluding data samples that fluent native speakers actually say.” This hindrance of fluency in learner production is precisely the main problem I see with the dominant use of non-authentic examples like those given in (266) and (267) in the classroom. I would assume that fluency in authentic English can be much more effectively improved by presenting learners with attested examples taken from authentic English, and not with artificial examples that have been invented by materials writers.

Cook, who does not seem to be very much in favour of authenticity in the classroom, plays down the “attested vs. invented sentences” problem. He claims that “[t]he utterances in attested data have also been invented, though for communication rather than illustration.” (Cook 2001: 376), thus putting both types of examples on one level. What is crucial here, however, is the reason Cook gives for “inventing” utterances in natural discourse: “for communication”. It certainly makes a difference whether language is “invented” by a competent speaker in a real-life communicative context or by a textbook author in the context of writing ELT materials. While the purpose of the former is communication, the purpose of the latter is the illustration of a particular language phenomenon. In other words, authentic examples arise from a specific communicative context and usually serve a pragmatic function while non-authentic examples serve to exemplify the grammatical structure dealt with in the respective textbook unit.⁸⁵ It is only natural that the outcome in both “language production” contexts is rather different and that the language invented in the second context will differ a lot

from the language used in a communicative situation. Like Willis (2003:224) I am suspicious of invented textbook sentences and see “a serious danger” in the circumstance that “specifically designed texts will show the language not as it is, but as the course writers imagine it to be or would like it to be.” As McCarthy (1988:49) fittingly points out, “blackboard sentences [...] have no licence to be taught as examples of *natural* use to learners” (italics in original; on naturalness see also Sinclair 1984).

The good news is that, with a growing number of corpora around, “there is no longer any need to invent example sentences in the time-honoured way.” (Francis 1993: 138) Instead of the constructed dialogues in (266) and (267), we might as well use corpus examples that have actually occurred in real conversations among real people. Such authentic examples, like those given in (268) to (271), could serve the same purpose as their concocted textbook counterparts, namely to exemplify the use of present progressives in affirmative and interrogative contexts. If we present learners with real instances of those progressive forms that are found to be most common in questions (e.g. *bringing*, *feeling*, *happening*, and *listening*; cf. 4.5.7 for other verb forms with high shares of questions), they will be better prepared for the English they are likely to encounter in real life. Besides, as Willis (2003:224) notes in this context, “[r]eal language provides a refreshing link between the classroom and the world outside”.

- (268) Hello Barbara! Hello! Oh are you bringing Bernard? I ’m bringing Bernard.
That ’s your privilege today then? Oh yep! (BNC_spoken)
- (269) How are you feeling now? (BNC_spoken)
- (270) What ’s happening now, does anybody know? (BNC_spoken)
- (271) Are you listening <M01> Yes I’m hanging in your every word get to the
point. (BoE_brspok)

What applied linguists have criticised with reference to proposals of using corpus examples in teaching materials is that such examples are inevitably severed from their authentic context and thus cannot be authentic for the learner in the language teaching classroom (cf. e.g. Widdowson 1991, 1998, 2001). It is true that authentic examples are decontextualised when we take them from a corpus to include them in a course-book. Obviously, we cannot transfer the whole context of an actual conversation into the classroom (on this issue see also Mauranen 2004:91), but it should be possible for learners to recontextualise examples, after all, as McCarthy recently noted, learners are used to doing that since they have to recontextualise things all the time in their everyday life, e.g. while watching soap operas on TV.⁸⁶ I suspect that it is harder for a learner to authenticate contextualised invented examples than to (re-)contextualise authentic ones. On this issue, I fully agree with Sinclair (1991:5) who calls it an “absurd notion that invented examples can actually represent the language better than real ones” and would subscribe to his well-known precept for language teaching “[p]resent real examples only” which is followed by the entirely plausible statement “[l]anguage cannot be invented, it can only be captured.” (Sinclair 1997:31)

Unlike people who object to using real “messy” data in language teaching (cf. e.g. Cook 1998; Hedge 2000; Owen 1996; Widdowson 2000), I think that an idealised “tidy” type of English should not be the preferred option. Pupils and students have to learn how to deal with “messiness” and “untidiness” in language, because in real-life communicative contexts they are unlikely to encounter the kind of well-formed sentences that their EFL coursebooks present them with. It might be argued that real examples taken straight from a corpus are too messy to be presented to the learners as models. Messiness in an authentic utterance can of course be controlled by the materials writer before the utterance is included in a coursebook. However, when we look at concordances of items we would like to introduce, we always find a number of concordance lines that may need no editing at all. Besides, I consider it rather risky to make changes in the corpus examples, since any change implies moving away from the naturalness of real discourse. A model is probably most effective when it clearly displays the feature it is modelling. It is the responsibility of the coursebook writer to make pedagogically appropriate choices from any of the rich collections of authentic data that are available. There are many features of an instance that can make it unsuitable for a pedagogical purpose; an example may contain difficult vocabulary, it may be sexist, or generally politically incorrect.⁸⁷ Among these features, the real-time features of natural discourse (e.g. repetition, hesitation, false starts, repair strategies) are certainly not prominent.

Since I also believe that “[i]n a classroom, corpus data is light years ahead of invented examples in authenticity” (Mauranen 2004:94), I would again like to stress the importance of banishing invented sentences from the coursebooks and suggest to present learners with real examples of “used” language instead (cf. also Römer 2004a, 2004b). What we need in the coursebooks is more than just constructed texts which exemplify the rules given in the adjacent grammar boxes. We need samples of actual discourse which show learners how language really works; in Sinclair’s terms, we need “examples that can be trusted because they have been used in real communication” (Cobuild 1990:vii).

7.3 Improving communicative competence: Teach the typical

[...] for most learners the pay off [of learning a second or foreign language] will be the ability to communicate with other speakers or writers of the language.
(Lock 1996:3)

The previous section on authenticity has already briefly dealt with communication and the importance of enabling language learners to participate in real-life communicative contexts. It is the task of successful language teaching (cf. e.g. Rutherford 1987:147), as also highlighted in the German curricular guidelines,⁸⁸ to help learners improve their communicative competence in the target language, i.e. to develop strategies of expressing (or encoding) their own communicative intentions and of understanding

(or decoding) the messages of their communication partners. In a nutshell, the aim is to make learners better communicators.

The concept of “communicative competence” was introduced by Dell Hymes in 1972 and has since then been discussed and further developed by Hymes himself (cf. Hymes 1972, 1992) and several other applied linguistic researchers (cf. Allen & Widowson 1974; Piepho 1974; Candlin 1976; Breen & Candlin 1980; Canale & Swain 1980, 1988; Canale 1983; Savignon 1983, 1990; Bachmann 1990). As a critical reaction to Chomsky’s competence vs. performance model proposed in *Aspects of the Theory of Syntax* (Chomsky 1965), Hymes points out that this view is largely over-simplified since it only analyses and describes what is formally possible in the language. What we need though, Hymes (1972:273) argues, to fully capture language and to understand how we can communicate effectively, is “a broader theory”. Not only do we need to know whether something is *possible*, but also whether and to what degree it is psycholinguistically *feasible*, contextually *appropriate*, and actually *performed* (cf. Hymes 1972:281; italics in original). Hymes (1972:282) states that

[t]here is an important sense in which a normal member of a community has knowledge with respect to all these aspects of the communicative systems available to him [or her]. He [or she] will interpret or assess the conduct of others and himself [or herself] in ways that reflect a knowledge of each (possible, feasible, appropriate), done (if so, how often).

Communicative competence in Hymes’s sense hence consists of four parts (and abilities on the speaker’s side) which are related to the four proposed questions: grammatical, psycholinguistic, sociocultural, and de facto. Let us now only focus on the last part which is perhaps the most important of the four, the “de facto” part, i.e. the knowledge of what actually occurs in language behaviour.

As observed by Stubbs (2001b:13), “our communicative competence relies on knowledge of what is expected or typical.” For the identification of what is “expected or typical” in the language, corpora and corpus-analytic tools are of course invaluable. They can, as I have just shown for the use of progressives, highlight the central contexts and functions of linguistic structures. As Aston (1997:51) rightly notes, “[...] corpora can play a useful role in the acquisition and restructuring of the schematic knowledge underlying communicative competence.” Following Stubbs’s observation on communicative competence and typicality, I would thus like to argue that learners can become communicatively more competent, i.e. they can become better communicators more easily, if they are taught the most typical things about language phenomena first and if they are confronted with the central features before they are exposed to the less common, marginal features. There probably is not much use in teaching learners things about a language which are not typical of its real use. This applies to features on the context and on the function side of linguistic items or structures. In the next two sections, I will further elaborate on the issue of “teaching the typical” and discuss what this means for the functions and contexts that language items, in our case progressives, are and ought to be presented in.

7.4 Teaching progressives in natural contexts

In this section I will argue that, in order to make learners communicatively more competent, they ought to be presented with items in the contexts in which they typically occur in natural language. As Hunston and Francis (2000:268) state,

[i]f the learner wishes to sound 'natural', 'idiomatic', or 'native-like', [...], he or she needs to use the collocations, the phraseologies and the patterns of English that native speakers automatically choose.

They go on by saying that

[...] patterns are essential to fluency as well as to accuracy. It is an unfortunate learner who has to think of every next word separately when uttering a sentence. (Hunston & Francis 2000:271; see also Hunston 2002b:167; Mindt 1985:174; Langacker 2000a:2)

I fully support this view. Also, much in the same vein, Barlow (2003:7) observes that

[...] knowledge of language lies in knowing how and when to use items from a large stock of word combinations, sometimes referred to as chunks or pre-fabricated units

and criticises that “[...] made up sentences, such as those used in coursebooks, tend not to contain collocations.” (Barlow 2003:7)

Collocations, i.e. the common co-occurrence patterns of items in real discourse (cf. 2.2.1), are certainly worth paying more attention to in language teaching (cf. also Nesselhauf 2005). At present, as my corpus-driven analysis has shown, the collocation and context patterns of progressives in coursebook English still differ considerably from those in natural spoken English. The examined textbook progressives are contextualised and not just presented randomly, but they do not occur in the contexts in which they most typically occur in real English. This, however, is crucial when we talk about “contextualisation” of words and phrases (cf. Lewis 1993:103; see also Lewis 2000). Like Lewis (1993:119) I regard collocation as the provider of a “most powerful organisational principle for language teaching”, and like Sinclair (1997:38) I ask myself “[w]hy [...] strong structural patterns [are] not recorded in grammars [and EFL textbooks]”. I would therefore like to make a couple of suggestions on what could be changed about the context patterns of progressives in teaching materials. These suggestions refer back to the list of differences which was the result of the comparison of corpus and coursebook findings (cf. 6.3). It needs to be stressed again that, for the time being and as long as no further comparable studies on coursebook English have been carried out, the changes suggested should not be understood as radical but as careful and gradual.

One helpful change to come closer to the collocational proportions found in real English would be to use more progressive examples that contain short forms of TO BE, in particular the patterns *'re V-ing* and *'s V-ing*, and fewer that contain the long form *are V-ing*. With a number of verbs (especially *asking*, *going*, *playing*, and *staying*)

it may be better to focus more on PastProgs and PresPerfProgs, as they are the preferred tense forms in authentic English, and less on present progressives. With respect to co-occurrences of progressives and subjects, I would suggest to incorporate more examples with the personal pronouns *I*, *it*, *we*, and *they* and fewer with *he*, *she*, and personal names in subject position. Some verb-specific patterns that are over-represented in the coursebooks, e.g. *you + doing*, should perhaps be used less frequently. We could also do something about the overuse of some verb-object collocations, for example by replacing a couple of instances of *going + preposition + the + noun (group)* by the underused patterns *asking + you*, *bothering + me*, *having + a + noun (group)*, *listening + to + the + noun (group)*, or *telling + you*.

Other changes may apply to the distribution of verb-preposition collocations. Some combinations, such as *agreeing with*, *dealing with*, *ringing up*, and *setting up* could be used more often, whereas less emphasis could be put on others, e.g. *coming back*, *getting at*, and *working on*. A stronger focus should perhaps be on negated progressives, since they are largely under-represented in the coursebooks, in particular with the verb forms *doing*, *going*, *having*, and *saying*. More negated progressives of these verbs could replace some of the negated occurrences of *asking*, *listening*, and *making*, as these forms are found less frequently in negative contexts in real English than in “school” English. Also, I would suggest there should be fewer examples of progressives in interrogative contexts and more in if-clauses and relative clauses, of course paying attention to the lexical restrictions that apply here (cf. 4.5.7). Preferably those verb forms ought to be presented in questions that typically occur in questions in actual language data (e.g. *bringing*, *eating*, *feeling*, and *living*), rather than those forms that do less commonly occur in interrogative contexts (e.g. *doing*, *playing*, and *thinking*). With respect to the co-selection of progressives and adverbials, my suggestion would be to pay more attention to some verb-specific patterns in real data. For instance, progressive examples should preferably contain more collocations of *doing*, *making*, and *wearing* with time adverbials, less time adverbial collocations with *asking*, *going*, and *playing*, and less place adverbial collocation with *asking*, *doing*, *sitting*, and *trying*.

In general, my suggestion would be to adjust the proportions in which items co-occur in the coursebooks to those in real spoken English. I believe that learners may profit from that kind of changes since they would render textbook English more authentic and bring it closer to the type of English learners will encounter in reality.

7.5 Focussing on frequent functions of progressives

I will now turn again from contexts to functions and suggest a couple of changes which may improve pedagogical descriptions of what progressives express or refer to. My claim is that the focus in teaching progressives should be on a set of frequent functions, functions that are commonly expressed by progressive forms in real English discourse. Since “frequency of usage is clearly an important consideration because of communicative utility” (Barlow 1996:30),⁸⁹ I argue that more emphasis should be put

in teaching on the most common functions identified for progressives and less on the infrequent ones, at least in the earlier stages of a course. Elsewhere in a study on modal verbs, I suggested to “use similar proportions of the different senses of a polysemous verb in the English used in schools as found in the English used in real-life situations.” (Römer 2004a: 196) I would now like to make the same suggestion for the use of progressives and their functions in EFL textbooks.

With reference to aspectual functions, Niemeyer (1991:342) notes that their discussion in a range of German grammars of English “can hardly be called adequate.” This is basically also what I found in the comparison of BNC/BoE and GEFL TC data (cf. 6.2). The coursebooks provide a rather monolithic view of what progressives can express.⁹⁰ In this simplified picture, one function of progressive forms, the reference to continuous and non-repeated actions or events, is particularly emphasised while other uses are marginalised. However, in real English, some of these “other uses” are rather common. Progressives which express repeatedness were for example found to be frequently used in my spoken BrNSE data. I would argue that the identified central function “CF2” (reference to continuous and repeated actions or events) certainly deserves more attention in the teaching materials. If we do not use more (real) examples which refer to repeated situations, learners might get the wrong impression that progressives can only be used to talk about single and not about multiple actions or events. Especially with the verb forms *coming*, *doing*, *helping*, *looking*, *playing*, *wearing*, and *working*, this function (CF2) ought to be used more often in the coursebooks.

Another function of progressives which is frequently expressed in the datasets from real spoken English, though only rarely occurs in the coursebook data, is “general validity”. As this is a rather common function, it might be worth introducing it to learners and presenting them with typical examples which refer to generally valid actions. We may also want to put some more emphasis on the functions “politeness or softening”, “emphasis or attitude”, and “gradual change and development”. These functions are not particularly central in real spoken English, at least not for progressives in general, but they are repeatedly expressed by a few common verbs, that means there are some attested lexical-grammatical relations which are not at all captured in the coursebooks. For instance, the “politeness/softening” function is often found to occur with *asking*, *saying*, *telling*, *thinking*, and *trying* (cf. 4.5.11). *Costing*, *hoping*, and *telling* are frequently used in emphatic contexts. These relations maybe ought to be covered in the coursebooks as well. Less emphasis in the textbooks could perhaps be put on the function I labelled “framing” (cf. 4.4.5). Although framing-progressives are extremely rare in real spoken English, they occur repeatedly with a number of verbs in the GEFL TC data (e.g. with *looking*, *sitting*, *talking*, and *thinking*). I would furthermore suggest a couple of changes or adjustments related to progressives and time reference types. It might be better to, on the whole, focus less on future time orientation, in particular with verbs like *going* and *playing*. In the EG 2000 volumes, a number of progressives with present time reference could be replaced by examples which refer to actions or events in the past. Also, and this is valid for both examined coursebook series, more attention should perhaps be given to progressives which show an “inde-

terminate” time orientation and do refer to situations in the present and the future. The neglect of this time reference type in the coursebooks goes hand in hand with the under-representation of the “general validity” function.

On the whole I would argue for an expansion of the simplified functional spectrum we observe in textbook English and suggest to present learners with a more complex and more complete picture of progressive functions. Even though pedagogical descriptions should preferably be as simple as possible, there is a certain danger in over-simplification, especially when we move on to a more advanced stage in the course. At beginner level, it may be sufficient to focus on the two central progressive functions identified in actual spoken English, CF1 and CF2 (cf. 4.4.3), but at intermediate or advanced level I consider it important to introduce learners to the wider range of functions that progressive forms can express. As Barlow (1996: 11) notes with reference to teaching English reflexives, a knowledge of “less frequent patterns is important in moving the language learner from intermediate to more advanced levels of proficiency.” I believe that this also applies to the less frequent (additional) functions of progressive forms. These functions do not deserve as much emphasis as the more frequent central functions, but I consider it very risky to marginalise them completely or not to cover them at all.

7.6 Shifting emphasis to lexis

In Section 4.5 I asked the question “How lexical is grammar?” and found, after a detailed verb-specific analysis of a large number of individual progressive form types, their preferred contexts and functions, that grammar, at least when it comes to progressives, is in fact very lexical. That means that there are a number of strong relations between progressive forms of individual verbs and the structures, contexts, and functions in which they typically occur.

To give just one example, I have shown that the shares of negation are very unevenly distributed across progressive verb forms (cf. 4.5.6). While in real spoken English some progressives, e.g. with *bothering*, *expecting*, *letting*, *meaning*, *paying*, and *suggesting*, are very frequently negated (showing shares of up to 51.79%), others, e.g. with *bringing*, *checking*, *hoping*, *living*, *meeting*, and *wondering*, never or very rarely appear in negative contexts. I said that these and other findings presented in 4.5 cast strong doubt on the usefulness of average distributional values in empirical linguistic description, since such values may be misleading and obscure important co-selectional tendencies of lexical items.⁹¹ I hence stress the centrality of “lexical grammar” as an approach to language analysis and description that integrates lexis and grammar and argue, in Hunston and Francis’s (1998: 63) terms, that it would be much more appropriate to “speak of one system” instead of treating the two parts separately. If, within a grammatical structure such as the progressive, different lexical items have preferences for different realisations of this structure (which is what I have found), this implies that it perhaps does not make much sense to start from this grammatical structure and talk

about “the progressive” as such. Instead, it is probably safer to make statements about the behaviour of particular progressives (or groups of progressives), i.e. about concrete lexical realisations of the form and their typical use in real discourse. Put differently, what I am suggesting to do is move away from a *grammar of “empty” constructions* (e.g. “the progressive”) to a *grammar of lexical items*.

This proposed shift in emphasis from grammar or syntax on the one hand to lexis or vocabulary on the other is supposed to affect linguistic description in general and pedagogical language description in particular.⁹² It is unfortunately still a prevalent idea in language teaching that grammar and lexis are two independent parts of the language. In this context it is symptomatic that issues related to “teaching vocabulary” and “teaching grammar” are usually dealt with in separate chapters of methodological handbooks for teachers (cf. e.g. Bausch, Christ & Krumm ³1995, Gehring ²2004) and in the curricular guidelines (cf. Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen 1993:59, 60). In the light of current corpus-driven research findings like the ones presented in this study, this probably needs to be changed. As Francis and Sinclair (1994:200) state,

[t]here is little point in presenting learners with syntactic structures – how groups and clauses are built up – and then presenting lexis separately and haphazardly as a resource for slotting into these structures. In other words, we should not burden learners with vast amounts of syntactic information on the one hand and lexical (“vocabulary”) information on the other.

I fully support this view. The present study has demonstrated how misleading and dangerous it can be to treat “the progressive” as a construction with empty slots that can be arbitrarily filled with whatever lexical item may come to mind.

Corpus linguistic analyses show us that there are a number of lexical-grammatical restrictions and verb-specific distributional typicalities which do not go together with such a slot-and-filler model of language. I would therefore like to make a point for an altered presentation of progressives in teaching materials, which takes these restrictions into account, pays more attention to the preferences of individual verbs, and emphasises the most typical attested patterns. Referring back to the example of progressives and negation, one concrete suggestion could be to start with verb forms like *bothering*, *expecting*, *letting*, *meaning*, *paying*, or *suggesting*, i.e. with forms that are often used in negated progressives, and introduce learners to patterns such as *I’m not bothering*, *I wasn’t expecting*, *I’m not letting*, *I’m not paying*, or *I’m not suggesting*, instead of presenting them with a construction like *TO BE not V-ing*, thus giving them the idea that any verb could equally well be used in this construction with all possible forms of *TO BE* and with all types of subjects. So far, the concept of “lexical grammar” has not yet arrived in language descriptions and reference works (with the *COBUILD Grammar Patterns* forming a welcome exception; cf. Francis, Hunston, & Manning 1996a, 1996b), and I suspect that it is, unfortunately, still far from arriving in the EFL classroom. However, if in the future more people (in particular applied linguists, materials

writers, and teachers) have more faith in corpus-driven linguistics and the insights it may offer, the prospects do not have to be that grim.

In the next and final section of this chapter I will build on my previous considerations on lexical grammar, frequent functions, typical contexts, communicative competence, and authenticity (cf. 7.2 to 7.6) and try to connect them in a first sketch of a new approach to teaching progressives.

7.7 Towards a corpus-driven communicative didactic lexical grammar of progressives

In this first sketch of a new approach to teaching progressives I will integrate the points raised and the suggestions made in the previous sections of this chapter and examine what could be improved about the way in which progressive forms are presented in EFL coursebooks. What I would like to propose is the development of a corpus-driven communicative didactic lexical grammar of progressives.

This grammar is (i) *corpus-driven* as it incorporates research in the Birmingham corpus linguistic tradition, inspired by the work of John Sinclair, Gill Francis, Susan Hunston, and Elena Tognini-Bonelli. At every stage in my work, I followed the principles of corpus-driven linguistics, starting from large amounts of corpus data, accepting and reflecting the evidence, and trying not to let previous findings and existing theories influence my results. The grammar is (ii) *communicative* in that it puts an emphasis on spoken language and its functions in real-life contexts and in that it aims to enhance communicative competence in the learner (cf. Leech 1988). It is (iii) *didactic* in Dieter Mindt's sense, which means that it is descriptive and includes information on frequencies and typical co-occurrences of items. Last but not least, my grammar is (iv) *lexical*, meaning that it shifts emphasis from syntax to lexis and aims to integrate lexical items and the structures and functions they are typically used in. It pays attention to the co-selectional preferences of individual verbs. The grammar will try to respond to the problems of selection, progression, and presentation, since those are the three key problems in language teaching that are, according to Mindt's (1981a: 179) model, supposed to be solved by a didactic grammar.

Let us first address the problem of *selection*. The question here is: What facts about progressives and their use should learners be presented with? To answer that, we have to explain what we mean by the rather general term "facts". What a learner needs to know in order to be able to use a linguistic item or structure appropriately, are the functions it expresses and the contexts it is found in. Since I regard frequency of occurrence and typicality of use in real English as key criteria to be considered in the selectional process, the functions and contexts of progressives that I would like to include in my pedagogical account are chosen according to these criteria. As discussed earlier on (cf. 4.1.1), I base my description on spoken British native-speaker English and treat this type of English as an appropriate norm, though not as the only possible one.⁹³ I believe that learners can profit most from spoken language models when it

comes to improving their communicative competence, which is an important aim of the present approach.⁹⁴ With respect to the selection of progressive tense forms, I think that learners should get to know how to use the following forms (however at different stages in the course): present progressive, past progressive, present perfect progressive, and past perfect progressive.

Table 56 provides a list of functions and context phenomena that were found to be common with progressives in spoken BrNSE. The relative frequencies of co-occurrence in BNC and BoE data are given in parentheses. In addition to the listed functions, all possible time reference types, as there are past time reference, present time reference, future time reference, and present/future “indeterminate” time reference, should be included in the account.

The next issue we need to deal with is *progression*, which means that decisions have to be made on the sequencing in the course of forms, functions, and contexts. The main guiding principle is again frequency of occurrence in the corpus data. The order that I would like to suggest for an introduction of the different progressive forms and their functions and context phenomena (as listed in Table 56) is given in Table 57. I also attempt to assign the listed items roughly to the years of instruction in which they ought to be introduced, taking a six-year course with English as the first foreign language as the basis for my considerations.

Table 56. Selection of typical functions and context phenomena for inclusion in pedagogical descriptions of progressives

Functions	
–	CF1: continuous + non-repeated actions/events (54.21%) ⁹⁵
–	CF2: continuous + repeated actions/events (27.47%)
–	non-continuous + non-repeated actions/events (9.24%)
–	non-continuous + repeated actions/events (8.78%)
–	general validity (22.35%) ⁹⁶
–	politeness/softening (12.82%)
–	emphasis/attitude/shock (9.46%)
–	gradual change and development (5.17%)
Context phenomena	
–	all long and short forms of the PresProg, PastProg, PresPerfProg, and PastPerfProg, with particular emphasis on the patterns <i>'re V-ing</i> (23.68%), <i>was V-ing</i> (16.94%), <i>'m V-ing</i> (13.48%), and <i>'s V-ing</i> (13.20%)
–	personal pronoun subjects, with particular emphasis on <i>I</i> (24.28%), <i>you</i> (17.13%), and <i>we</i> (12.37%)
–	negation (8.39%)
–	interrogative contexts (10.94%)
–	if-clauses (4.47%) and relative clauses (6.17%)
–	time adverbial specification (25.04%)
–	place/“other” adv. specification (3.04%/7.68%)

Table 57. A possible progression in teaching progressive forms, functions, and contexts

Year	Form	Function and time reference	Context features
1	PresProg (<i>after</i> Simple Present)	CF1 (cont. + non-rep.) → CF2 (cont. + repeated); present time reference	short and long forms of the PresProg, focus on 're V-ing, 's V-ing, 'm V-ing, mainly in collocation with the subjects <i>I, you, and we</i> ; questions (with selected verbs, cf. 4.5.7 and 7.5); collocation with time adverbials
2	PastProg; (review PresProg)	past time reference → future time reference; (review CF1 and CF2)	<i>was/were V-ing</i> (focus on <i>was</i>), in collocation with all personal pronoun subjects; negated progressives (mainly of PresProg); collocation with "other" adverbials (with selected verbs)
3	PresPerf Prog → PastPerf Prog (review PresProg and PastProg)	"indeterminate" time reference; general validity; politeness/softening	short and long forms of the PresPerfProg and PastPerfProg, focus on 've <i>been V-ing</i> ; relative clauses and if-clauses (with selected verbs only); collocation with place adverbials (with selected verbs only)
4	(review all forms)	non-cont. + non-rep. actions/events non-cont. + rep. actions/events; emphasis/attitude/shock; gradual change and development	

The third and final problem to solve is the problem of *presentation*. How, i.e. in which lexical-grammatical contexts, are we supposed to present progressives at the different stages in the course? My answer to this question brings together two of the central pedagogical concerns that were discussed in Sections 7.2 and 7.6 above: an increase in authenticity in the classroom and a stronger emphasis on the preferred patterns of individual lexical items. In the presentation of progressives I hence suggest to pay highest possible attention to the most typical patterns found in real English, and to include these patterns in the teaching materials. Forms, functions, and typical context features (as listed in Tables 56 and 57) are supposed to be introduced by means of those verbs that are most common in the respective construction in real English (cf. the detailed descriptions in 4.5 and the suggestions made in 7.4 to 7.6).

For example, when we introduce the PresProg with future time reference, we ought to make clear which verbs are usually found to express futurity in which progressive patterns (e.g. *coming* in *I'm (not) coming* + preposition, *leaving* in *I'm leaving* or *(are) you leaving?*, and *stopping* in *I'm (not) stopping* or *we're (not) stopping*). All exam-

ples should be taken from authentic sources; that means they should present “used” language instead of “invented” language. As an alternative to simply presenting single example sentences, (filtered) concordances of selected -ing forms could be used. With a basic training in how such displays of multiple language examples of an item in context are supposed to be read,⁹⁷ intermediate or advanced learners could be given the task to work out the most typical usage patterns of progressives themselves. They could be enabled to make their own discoveries about language regularities and hence learn inductively.⁹⁸ A large number of authentic examples from spoken English corpora which could be used to illustrate the selected forms, functions, and contexts can be found in Chapter 4 in the sections which deal with the use of progressives in BNC_spoken and BoE_brspek (cf. 4.3. and 4.4). The most typical lexical-grammatical associations that we may want to emphasise in language teaching are exemplified in the Sections of 4.5.

To sum up my considerations related to the three problems a didactic grammar is supposed to solve, it can be said that the key criteria which have led to the decisions I made about the selection, progression, and presentation of progressives in pedagogical descriptions were *typicality*, *authenticity*, and *communicative utility*. I think that these three criteria must not be underrated as they may have a considerable positive impact on language teaching. My draft of a corpus-driven communicative didactic lexical grammar of progressives has tried to highlight these key issues.

Of course, the pedagogical concept that I have just outlined does not represent a comprehensive model and is probably not immediately transferable to the classroom. What this draft of a grammar is supposed to do, though, is provide some useful information for materials writers and teachers. Given the stresses of their job, teachers cannot be expected to keep track with all sorts of recent developments in linguistic research, or even to carry out linguistic research themselves in order to improve their teaching. It should be part of the responsibility of (applied) linguists to help out here and assist in bridging the gap between corpus research and practice, hence making life easier for language practitioners while at the same time improving the situation for language learners. There is a need to make teachers aware of the fact that, unless more attention is being paid in materials design to the actual facts of language, they run the risk of neglecting frequent and communicatively important patterns and features (easy to uncover by means of corpus analysis), while at the same time focussing on less central, less important ones.

It goes without saying that we as corpus linguists cannot and should not impose too much on teachers and coursebook writers or even force them to change their ways of teaching and presenting grammar to learners. We can and should, however, offer them alternatives of how they could present learners with the language they have chosen to learn in more appropriate ways. What they need, and what they certainly deserve, are better, more reliable descriptions of language – descriptions of language as it really is, and this is where corpus-driven linguistics can help. Picking up Owen’s (1993: 185) master and servant metaphor,⁹⁹ I think that the corpus is in fact in a rather powerful position. However, instead of thinking about the relationship between grammarian, teacher, and learner on the one side versus corpus and concordancer on the

other in terms of “master” and “servant” (both terms which share rather unfavourable semantic prosodies), it may be more appropriate to think about a relation between “giver” and “receiver” of invaluable information about the language.

CHAPTER 8

Conclusions

Corpus, practice, theory

As the title indicates, it was the aim of *Progressives, Patterns, Pedagogy* to provide new insights into the use of progressive forms, the structures they normally occur in, and their treatment in language teaching. The study has dealt with some modest and basic questions, as there were: How are different progressive forms distributed? In which lexical or syntactical contexts do they usually occur? What do different progressive forms typically express? and Is it possible to identify a “generally valid” behaviour of progressives, or do different verbs show largely dissimilar context and function patterns? These questions have been addressed to two types of language data: (i) data collected from the largest available corpora of spoken British English, the spoken part of the British National Corpus (BNC_spoken) and the spoken British subcomponent of The Bank of English (BoE_brspok), and (ii) data retrieved from GEFL TC, a small and specialised corpus of German EFL textbook texts. Starting from these resources, I have explored how progressives are used by native speakers of English in actual communicative situations, how they are presented in English language teaching materials, and thirdly in how far coursebook language differs from real language use.

The topic has been approached from a corpus-driven linguistic angle, following the pioneering work of John Sinclair, Gill Francis, Susan Hunston, Elena Tognini-Bonelli, and others. Working corpus-driven meant that I put the corpus first in my analysis, approached it without any fixed theoretical concepts or preconceived ideas in mind, derived classificatory labels and categories from the data, and accepted and reflected the evidence at every stage in my discussion. Like Sinclair (2002) I consider it “intellectually unhealthy” to use corpus data only to illustrate existing theories and not in order to question and adjust them.¹⁰⁰ In my approach, corpus linguistics, or to be more precise, corpus-driven linguistics (CDL) is regarded as more than a mere methodology. CDL is seen as an altogether new way of doing linguistics, since it has the potential to affect both the practical and the theoretical side of the study of language.

Detailed objective descriptions based on large amounts of data as provided by corpora can be of major interest and use for language pedagogy and also for the construction of new linguistic frameworks which aim to come closer to capturing the structure of language. A central purpose of the present corpus-driven study was hence a major stock-taking of progressives in spoken English. I have tried to observe and describe as closely as possible what was actually there in the language data and then draw conclusions about progressives on the basis of these observations and de-

scriptions.¹⁰¹ This detailed analysis of progressives in spoken English constituted the “general linguistic interest” part of the study. The primary motivation of my work, however, was pedagogical. Starting from the problems which the use of progressives causes for learners, even on an advanced level, I wanted to find out whether these problems are “progressive-inherent” (i.e. whether progressives are just more difficult to use than other items), or whether they possibly arise from inadequate descriptions in the teaching materials that learners are exposed to. I hence took stock of the treatment of progressives in EFL coursebooks. This in depth analysis of progressives in “school” English and the subsequent comparison of “school” English and spoken English findings constituted the “pedagogical interest” part of the study. Under the headings “From corpus to practice” and “From corpus to theory” I will now again summarise some central findings of both parts, discuss limitations, and make suggestions for future research.

8.1 From corpus to practice

In general, the exploration of EFL coursebook language which was at the heart of the pedagogical part of the study has contributed to a better understanding of the way in which progressive forms, their functions, and contexts are presented in the language teaching classroom. More specifically, my analysis of the concordances retrieved from the two GEFL TC subcorpora, GLN and EG 2000, has highlighted which progressive form types are used in what proportions; what kinds of subjects, objects, prepositions, and adverbials they mainly co-occur with; how frequently they are found in negated or interrogative contexts, or in if-clauses or relative clauses; whether non-contracted or contracted forms prevail; how often they are used to refer to actions or events in the past, present, or future; and what types of situations they mainly express or refer to. These findings about the use of progressives in “school” English have been presented and discussed in some detail in Chapter 5 of the book.

For the pedagogical purpose of this study, however, it was important to go one step further from the mere accumulation of facts and compare the coursebook findings with the results based on real spoken English corpus data which were obtained in the general linguistic part of the analysis. This comparison of empirical findings from the “school” English and spoken English analyses brought to light a number of significant differences between progressives in real language use and their treatment in the coursebooks (cf. Ch. 6) and enabled me to derive some pedagogical implications from the observed differences (cf. Ch. 7). An integration of these implications and the larger number of concrete practical suggestions made for an adjustment of the way in which progressives are presented in the textbooks has then led to a first draft of a new concept of teaching progressives (outlined in 7.7) which tries to incorporate all major findings of the study.

As stated in the introduction and in Chapter 7, a central purpose of my work is to help bridging the gap between research and practice, to assist teachers and make life easier for learners. In this context, an aim that I share with Mindt (1996: 247) is to

“bring textbooks for teaching English as a foreign language into closer correspondence with actual English”, thus helping to improve language teaching. Right now, the input data that goes into the teaching materials is rather questionable. Instead of natural “used” language, most textbooks contain invented language that has been constructed for the purpose of teaching. I have argued in some detail that something needs to be done to change that. At the beginning of Chapter 7 I said that, if we let it, corpus-driven linguistics can indeed make a significant contribution to language pedagogy and positively affect teaching materials. It has hopefully become clear from my discussions (particularly in Chs. 6 and 7) what this contribution may look like.

Taking the presentation of progressives as an example, this study has provided a systematic empirical account of the relationship between English in use and English in the classroom and thus presents a typical case of the indirect use of corpora in language pedagogy. Even though I have tried to be as comprehensive as possible in my coverage of topic and data, I was of course not able to investigate all the things that would have been worth investigation; you can never do everything. Leaving aside the time criterion, the limitations I see are mainly related to the restricted availability of certain types of corpus resources. Like Sinclair (cf. e.g. 1991: 18, 100) I believe that it is always better to have more data and larger corpora since a broader coverage will imply a higher degree of representativeness and will ensure a greater confidence about the validity and significance of the findings. As far as the accessibility of spoken British English corpora was concerned, I was in a comparatively lucky situation. Taking into account how difficult, time-consuming, and thus expensive it is to compile a speech corpus of only several hundred thousands of words, the 10 and 20 million word corpora my research is based on represent two amazing achievements. And, in fact, these resources proved to be big enough to provide me with reliable results. Supportive evidence for the representativeness of my findings can be found repeatedly throughout Chapter 4.

The situation was much more difficult when it came to finding appropriate resources of “school” English. As pointed out in 5.3, when I started the project, no electronic collections of EFL teaching materials were readily available. It was also pointed out that the corpus I have compiled to at least partially close this availability gap is very small judged by today’s standards. It would hence be desirable to have more and larger corpora of the GEFL TC kind, preferably from a wider range of EFL teaching backgrounds so that the language of coursebooks used in Germany could be easily compared with the language of Austrian, Chinese, Italian, Japanese, or Swedish EFL textbooks, to name just a few countries. A new welcome development in this particular area is a research project entitled “Foreign language learning: Phraseology and discourse” which has recently been launched at the Catholic University of Louvain, Belgium; a project which involves the compilation of the TeMa corpus, consisting of ELT textbook material used in Belgian secondary schools (cf. Gouverneur & Meunier 2004; Meunier 2004).

Another data resource which may prove invaluable in this pedagogical context and which would allow for a more comprehensive analysis of “school” English, is a

spoken corpus of classroom discourse that would fully capture the language used in the EFL classroom, ideally not only in an orthographically transcribed format but with sound or video files linked to the transcripts. It would be worth going back to Sinclair and Coulthard's early model study based on teacher-pupil discourse (cf. Sinclair & Coulthard 1975).

A third component that I also would have liked to include in my comparative analysis, in addition to what I have termed "real" and "ideal" language learner *input* (cf. also Römer 2004b), is language learner *output*. Learner corpora, i.e. systematic collections of the language produced by learners, provide an invaluable resource for the investigation of this type of language. Although a number of corpora exist which capture written learner data,¹⁰² there is still a need to foster the development of learner corpora, in particular on the spoken side. To ensure comparability with the spoken English and spoken-type "school" English data, I would have required access to a spoken learner corpus, preferably of German L1 learner production. Such a corpus is only now being compiled by a research team around Joybrato Mukherjee at the University of Gießen, Germany, as part of the LINDSEI project.¹⁰³

Yet another dimension which the study may have profited from is the contrastive dimension. Learning problems, like those many learners have with progressives and their use, can only be fully understood on the background of the learners' L1, in our case German (cf. 5.1). To find out more about language contrasts and the ways in which progressive forms and functions are realised in the German language, a parallel corpus would be needed, ideally of spoken or spoken-type language. Such a parallel corpus would include transcripts of spoken (British) English texts that are aligned to their German translations. Obviously, the compilation of parallel speech corpora poses some serious practical problems since we usually do not have any translations of conversations. However, what we often find are simultaneous oral translations in interviews or discussions, so it should be possible to record what a German interpreter makes out of the spoken English input he or she gets and create a parallel corpus from these two versions of a dialogue. Besides, scripted dialogues as they occur in films are one possible resource which is commonly available in original and translated versions. The extraction of subtitles in different languages from movie DVDs could be a first step in the direction of spoken(-type) parallel corpora which could provide a very useful tool in translation studies and contrastive analysis, also for pedagogical purposes.

In future studies that follow my comparative corpus vs. coursebook approach, it would also be important to take into account the regional variation dimension and cover varieties of English other than BrNSE. I have discussed some possible objections to using a native-speaker language norm earlier on (cf. 4.1.1), but criticisms from people who call for coverage of non-British varieties, especially American English, are as justified as criticisms from people who make a point for international or lingua franca English. If we want to include this regional variation component, we are again faced with problems related to the availability, or unavailability as the case may be, of appropriate spoken corpora. At present, there are no corpora of different English language varieties around which beat the size of BNC_spoken and BoE_brspok.¹⁰⁴ Lastly, there

is probably also a need for further studies which examine the teaching of progressives and other lexical-grammatical trouble-makers from an SLA point of view and investigate what current theories of Second Language Acquisition have to offer to solving existing learning problems (on this topic, see Meunier 2002).

The list of desiderata, mainly on the resources side, indicates that, although quite a lot has already been achieved in the area of corpus linguistics and language teaching, we are still in the early (or earlier) days of this movement. A large number of challenges still lie ahead waiting for us. With regard to the main research purpose in the present study at the interface of corpus linguistics and pedagogy, the type of challenge is pretty clear. Analyses of real English in comparison with “school” English have to be expanded so that they cover more language phenomena and a wider range of types of data. To reiterate what Johansson and Stavestrand said back in 1987, “[t]here is a need for more investigations of the relationship between textbook language and natural discourse.” (Johansson & Stavestrand 1987: 147)

8.2 From corpus to theory

At several stages in my work, I stressed the importance of putting the data first in linguistic analysis and description, and of doing a careful stocktaking of the phenomena under scrutiny, in my case of progressive constructions. A detailed account of my large-scale “progressive” stock-taking is given in Chapter 4 of this book. I have also tried to show that it is essential in any kind of linguistic work to go beyond the stage of mere description, and I have argued that corpus-driven linguistics, taken seriously, does exactly that.

Not only have my analyses provided detailed information on the actual use of progressives in spoken BrNSE, the contexts and functions they typically occur in, and the relations that exist between individual lexical items and progressive patterns, they have also shown that some existing accounts of the progressive are inappropriate in several respects or that they emphasise less central aspects while neglecting other important ones. A case in point is the treatment of progressives as a uniform group, going hand in hand with the lack of attention being paid to lexical-grammatical relations. The analysis reported on in Section 4.5, which took the progressive forms of 100 individual high-frequency verbs as the starting point, clearly demonstrated that average values and statements about “the progressive” as such, as can be found in most existing accounts of this construction, do not tell us a lot about the actual use and behaviour of a large number of progressive forms. What we obviously need is an account of progressives, and perhaps of language in general, that pays more attention to lexical preferences and co-selection patterns; in other words, we need an account which individualises more and generalises less. Of course, it is unjust to blame existing descriptions for not capturing new evidence, but it ought to be taken seriously what Sinclair (2004c: 271) discusses in a recent paper, namely that new evidence calls for “new priorities, [and] new attitudes”. There certainly is a necessity of adjusting lin-

guistic descriptions and approaches to language to the new findings and insights that corpora provide us with.

It may be that certain existing linguistic theories lend themselves to the corpus-driven approach. I am, for instance, thinking of usage-based theories in the framework of Cognitive Grammar. I would, however, not attempt to try and bridge the gap between corpus linguistics and the generative models developed by Noam Chomsky, for whom corpus linguistics “doesn’t exist” (B. Aarts 2000: 5). Being usage-based, the cognitive models I have in mind account for language in performance and give prominence to frequency of occurrence, collocation, and language patterning (cf. the contributions in Barlow & Kemmer 2000). Also, we have no strict separation of lexis and grammar in models such as Construction Grammar (cf. Goldberg 1995; Croft & Cruse 2004: 256), or in Langacker’s usage-based model (cf. Langacker 2000a: 3, 2000b: 32). Thus, we might envisage a marriage of cognitive and corpus approaches, possibly with a corpus-driven usage-based theory being the offspring. However, it might also be that an entirely new theory, built more or less from scratch, will eventually emerge from corpus-driven work; a theory which requires an altogether different set of descriptive labels and categories. At present, we cannot tell. We just have to concordance and see.

Notes

Chapter 1

1. Restrictions as to what types of verbs are usually not found in the progressive will be dealt with in Chapter 3 of this study.
2. Cf. Stubbs's remarks on observation in science: "the development of science would have been impossible without observational technologies." (2000:21)
3. In this context it is worth quoting Susan Hunston, who in the introduction to her book *Corpora in Applied Linguistics* stresses the surplus value of a corpus in language studies as compared to other types of data: "Corpora allow researchers not only to count categories in traditional approaches to language but also to observe categories and phenomena that have not been noticed before." (2002a:1)
4. The term "lexical grammar" is here preferred over the more widely used "lexicogrammar" (also "lexico-grammar") as it serves better to express a field which truly integrates lexis and grammar. According to Sinclair, lexicogrammar "is fundamentally grammar with a certain amount of attention to lexical patterns within the grammatical frameworks; it is not in any sense an attempt to build together a grammar and lexis on an equal basis." (2000b:191) As Sinclair (2004c:277) states in a more recent article, "[l]exicogrammar is still firmly a kind of grammar, laced, or perhaps spiked with some lexis."

Chapter 2

5. Later in his article Leech (1992:106) argues that CCL (computer corpus linguistics) may be a "new philosophical approach to the subject", not just a methodological one. He does not go as far as to call CCL a theory.
6. COBUILD is an acronym for COLLins BirmIngham University International Language Database. For information on this project see Sinclair 1987a.
7. This chapter can only give a brief account of the major differences between CDL and CBL (corpus-based linguistics). The reader is referred to Tognini-Bonelli 1996a, 1996b, and 2001 for more detailed information.
8. This approach was first used by the COBUILD team at Birmingham University in the 1980s and has since then been further developed mainly by Elena Tognini-Bonelli, John Sinclair, Gill Francis, and Susan Hunston.

9. An example of a corpus-based study on the progressive is Legenhausen's 1985 analysis of progressive and non-progressive forms in (British) English radio programmes (including football and tennis reports, and the wedding ceremony of Prince Charles and Lady Diana), using Reichenbach's (1947) time interval system. Legenhausen chooses a particular theoretical framework on the basis of which he analyses the data.
10. We will come back to this and other Sinclairian precepts in Chapter 7.
11. For information on the Tuscan Word Centre, check <http://www.twc.it> (consulted: 22.08.04).
12. The 3C2S-method is comparable to Francis's 1993 "item-environment method".
13. For further exemplifications of semantic prosodies, see Bublitz 1996, Esser 1999, Louw 1993, and Partington 2004.
14. See Bausch 1979, Börner & Vogel 1976, Dirven 1990, Hüllen 1973 and 1976, Jung 1975, Klein 1987, Mindt 1971, 1981b, 1982, and 1986, Ungerer 1974, and Zimmermann 1984.
15. I will respond to these three central problems in Section 7.7 in the context of developing a new didactic grammar concept.
16. For information on communicative grammar see also Candlin (1973:64), who suggests "to reorganise LT materials on a communicative basis" and Leech & Svartvik ²1994.

Chapter 3

17. The Brown corpus (size 1 million words), compiled by researchers at Brown University, Providence, Rhode Island, is an electronic collection of samples of written American English taken from a variety of text types. The Lancaster-Oslo/Bergen-Corpus (LOB) can be regarded as the British answer to Brown. It consists of 1 million words of written British English and was created in a British-Norwegian collaborative project.
18. On the topic of recent grammatical change and the usability of Frown and FLOB, see also Mair 1997.
19. In his monograph *A Grammar of Speech*, David Brazil notes that his concern is "used language: that is to say, language which has occurred under circumstances in which the speaker was known to be doing something more than demonstrate the way the system works." (1995:24) This type of real language contrasts with invented language the prime purpose of which is *not* communication.
20. See the description of the grammar in the Cambridge University Press *Language and Linguistics Catalogue* 2002, p. 1. The label "grammar for the 21st century" has also been used by the promoters of Biber et al. 1999 (cf. the Longman website at http://www.longman.com/dictionaries/which_dict/lgswe.html; consulted: 22.08.04).
21. In his review of LGSWE, Schulze calls the grammar an "Eldorado" for everyone interested in (mainly functional) syntactic, semantic, and pragmatic variation in English (2003:134).
22. This lack of transparency concerning the empirical basis of EGEVS has also been criticised by Görlach (2001 and 2003), who notes that "there is no point in merging different corpora with the sole aim of enlarging the data base, since individual collections are necessarily compiled by different scholars who have different aims in mind." (2003:21)

23. See also the critical review of CamGEL by Mukherjee and the following discussion on Linguist List in July 2002, Linguist List issues 13.1853, 13.1932, 13.1952, and 13.2005 (check the online archives at <http://www.ling.ed.ac.uk/linguist/issues/>; consulted: 22.08.04).
24. A number of checks of randomly chosen parts from CamGEL examples in Brown, ACE, and LOB did not render any positive results. It is one major advantage of corpus-based and corpus-driven studies from publically available corpora that they permit *replication* and thus help rendering linguistics as scientific as physics or chemistry. Of course, scientific results can only be tested when we know what exactly they are based on and where the provided examples are taken from.
25. Kennedy (1998: 122) gives its size as “about 150,000 words”. There is no word count provided in Ota’s book.
26. A similar frequency is given by Quirk et al. (1985: 198) who write that “less than 5 per cent of verb phrases are progressive”. It is not fully clear, though, whether the authors have carried out their own quantitative analyses or just refer back to what can be found in the literature on the topic.
27. In the present account, I do not differentiate between the labels “function” and “meaning”. I use the more general term “function(s)” throughout with reference to what other linguists may term “meaning(s)”, namely with reference to what a certain form expresses or refers to. What is expressed in my system of function classification may be a semantic or a pragmatic meaning.
28. The only explicitly numerical information found in Joos’s book, for instance, is related to the form *be going to*: “In *Trial* *be going to* is used 28 times with future meaning” (1964: 134).
29. BFBS is the British Forces Broadcasting Service. For more information check the BFBS website at <http://www.bfbs.com> (consulted: 22.08.04).
30. As meanings can combine with other meanings, percentages add up to more than 100% (cf. Mindt 2000: 256).
31. A few pages later Scheffer writes the following: “Of all the verbs in the progressive about 25% occur with a temporal adverbial in either direct or removed context.” It remains unclear whether this higher percentage results from an extension of the collocational span.

Chapter 4

32. Comparable guide-lines also exist for teachers in all other German federal states and are published under titles such as “Rahmenpläne” (e.g. Berlin, Mecklenburg-Western Pomerania), “Rahmenrichtlinien” (e.g. Saxony), “Lehrpläne” (e.g. Bavaria, Rhineland-Palatinate), and “Bildungspläne” (e.g. Baden-Württemberg).
33. This is not valid for the learners’s *receptive* competence where a training based on spoken and written materials from different language varieties is essential for an improvement of language awareness.
34. On the importance of accounting for language variation in ELT, see Conrad 2004.
35. The Santa Barbara Corpus of Spoken American English (SBCSAE), for example, has a target size of only 200,000 words. According to a wordcount carried out by Norbert Schlüter, the first released part of the SBCSAE contains less than 60,000 words (cf. Schlüter 2002: 65). Important specialised spoken American corpora are the 2-million word Corpus of Spoken Pro-

fessional American English (CSPAЕ) and the Michigan Corpus of Academic Spoken English (MICASE), currently at a size of about 1.8 million words (August 2004). For further information check <http://www.athel.com/cspatg.html> and <http://www.hti.umich.edu/m/micase/> (consulted: 22.08.04).

36. At a colloquium on “Research into English as a lingua franca: the state of the art” at the BAAL 36th Annual Meeting (September 4th–6th 2003, University of Leeds) it was for instance reported that in conversations among non-native speakers the use of non-standard forms, such as “brought up” instead of “brought up”, usually does not cause problems or breaks in the communication.

37. VOICE is sometimes also referred to as “Vienna-Oxford ELF Corpus” (cf. e.g. Seidlhofer 2001a:76), apparently a preliminary name of the corpus.

38. The Birmingham Collection of English Text (also simply called “Birmingham Corpus”) was later included in The Bank of English. For more information see Section 4.1.4 and Sinclair (ed.) 1987a.

39. The Survey of English Usage (SEU), founded by Sir Randolph Quirk in 1959, is a pre-electronic one-million word corpus of equal shares of spoken and written British English from the 1960s.

40. The Cambridge International Corpus, for example, can only be accessed by authors who write for Cambridge University Press. CANCODE is part of this corpus. For more information readers are referred to the CUP ELT website at <http://uk.cambridge.org/elt/corpus>. More information on the Longman Corpus Network can be found at <http://www.longman.com/dictionaries/corpus/lccont.html> (consulted: 22.08.04).

41. The Reuters Corpus Volume 1 (RCV1) consists of more than 800,000 news stories from August 1996 to August 1997, the typical news output of one year. RCV1 is available free of charge from the global news group Reuters. More information about the corpus can be accessed via the company website at <http://about.reuters.com/researchandstandards/corpus> (consulted: 22.08.04).

42. The issues of corpus size and representativeness are touched upon in almost every introductory corpus linguistic textbook (cf. e.g. Biber, Conrad & Reppen 1998, Hunston 2002a, Kennedy 1998, McEnery & Wilson ²2001, Meyer 2002). See also Bowker & Pearson 2002, Gavioli 2002, and Ghadessy, Henry & Roseberry 2001 for interesting discussions of representativeness and small corpora. Relevant CORPORA list issues are stored online in the mailing list archives at <http://www.hit.uib.no/corpora> (consulted: 22.08.04).

43. In the compilation of a spoken corpus it is important to pay attention to a number of sociolinguistic variables, such as speaker gender, age, ethnicity, or social status, and to sample speech from different geographical regions. See Crowdy (1993:259) who states that “[r]epresentativeness is achieved by sampling a spread of language producers in terms of age, gender, social group, and region, and recording their language output over a set period of time.”

44. Detailed information on the BNC design and composition can be found in Aston & Burnard 1998 and Burnard 1995. See also the project website at <http://www.natcorp.ox.ac.uk> (consulted: 22.08.04).

45. For detailed information on the World Edition, see Burnard 2000 and 2002.

46. For more information on the construction of BNC_spoken (sampling strategies, transcription conventions, etc.), see Crowdy 1993 and 1995.

47. As mentioned above, COBUILD stands for Collins Birmingham University International Language Database. The collaborative COBUILD project between HarperCollins publishers and researchers at the University of Birmingham was initiated in 1980 by John Sinclair (cf. Krishnamurthy 2002a and Sinclair 1993).

48. For up-to-date information on COBUILD and the current size of The Bank of English, check the website of Collins publishing at <http://www.collins.co.uk/> and the project description at <http://www.collins.co.uk/books.aspx?group=153> (consulted: 13.01.05).

49. For more information on typical features of monitor corpora, see Clear 1988 and Sinclair 1991.

50. I would like to thank Ramesh Krishnamurthy for his kind assistance during my research stay at the Centre for Advanced Research in English at the University of Birmingham in Spring 2001 and for providing me with detailed information on the composition of The Bank of English, especially about the brspok subsection.

51. For further information on BNCweb, see <http://homepage.mac.com/bncweb/home.html> (consulted: 22.08.04). I wish to thank Sebastian Hoffmann for providing me with some verb frequency lists at a time when BNCweb was only available for researchers at the University of Zürich.

52. Since its first release with BNC version 1.0 SARA has been further developed. The SARA version that is now distributed with the BNC World Edition is a lot more user-friendly than the earlier version.

53. See Reppen 2001 for a comparative review of MonoConc Pro and WordSmith Tools.

54. My message is stored in the archives of the CORPORA list at <http://www.hit.uib.no/corpora/2000-1/0196.html> (consulted: 22.08.04).

55. I would like to thank again all the helpful CORPORA list subscribers who responded to my query (in alphabetical order): Guy Aston, Michael Barlow, Ylva Berglund, Oliver Christ, Alejandro Curado Fuentes, Arne Fitschen, Stefan Th. Gries, Naomi Hallan, Sebastian Hoffmann, Tony Jappy, David Lee, Chris Tribble, and Alexander S. Yeh. Particular thanks go to Sebastian Hoffmann, David Lee, and Chris Tribble.

56. For detailed information on the use of *Lookup* and its query options, the reader is referred to Potter 1999, available from COBUILD*Direct*, Westmere, 50 Edgbaston Park Road, Birmingham, B15 2RX, UK.

57. In the analysis of functions and contexts, the defined span of 200 characters (roughly 3 lines in an average Word document printout) was found to be appropriate. There was only a small number of examples that would have required more context for the function determination. Of course, “200” is a fairly arbitrary cut-off, but with a sum of 27,252 collected concordance lines, I had to find a way to keep the datasets to a manageable size.

58. There were no progressives of the forms *knowing* and *supposing* in BNC_spoken and no progressives of *matter* in BNC_spoken and BoE_brspok.

59. Readers might argue that “preposition” is not the correct category for *at*, *for*, *up* etc., at least not the only one. In a multi-word verbal construction like *look at*, *at* is often labelled “particle” (cf. e.g. Quirk et al. 1985: 1150). As in the present analysis it was not taken into account whether the *looking* + *at/for/up* etc. combination behaved as a single unit or not and in order to avoid further complexity in description, I have decided to stick to the single label “preposition”.

60. For more information on statistics in (corpus) linguistics, see Albert & Koster 2002, Butler 1985, Church et al. 1991, Dunning 1993, Kilgariff 2001, and Oakes 1998.
61. The Web Chi Square Calculator is accessible via a website maintained by Catherine N. Ball and Jeffrey Connor-Linton (see http://www.georgetown.edu/faculty/cball/webtools/web_chi.html; consulted: 22.08.04). For further information on the test, requirements, and interpretations, the reader is referred to the Georgetown University online tutorial at http://www.georgetown.edu/faculty/cball/webtools/web_chi_tut.html (consulted: 22.08.04).
62. In the cases of the PresPerfProg and PastPerfProg the *to be* form *been* is combined with forms of the auxiliary *have*.
63. For the PresProg, the PresPerfProg, and the PastPerfProg the chi-square test did not show any significance at the .05 level. According to chi-square, only the PastProg distribution in the two datasets is just about significant.
64. In the cases of PresPerfProg and PastPerfProg forms of *TO HAVE* (plus *been*) are used in auxiliary function.
65. Repeatedly used adverbials in the “others” group (cf. Figure 30) include *always*, *for* + noun group (e.g. *for a moment*), *this* + noun group (e.g. *this morning*), *again*, *all the time*, *today*, *tonight*, *tomorrow*, *yesterday*, *last* + noun (e.g. *last night*), *before*, and *on* + PP (e.g. *on Thursday*).
66. In Mindt’s system, prediction is used as a “[t]erm for something which is confidently expected to happen” (2000:249). His meaning “volition/intention” is defined as “[t]erm for willingness or readiness to do something or for a plan or an arrangement to do something” (2000:250).
67. There were no occurrences of *asking*, *mattering*, and *suggesting* that expressed this function.
68. The ten verb forms that did not occur in this function are: *asking*, *believing*, *calling*, *costing*, *knowing*, *living*, *mattering*, *sending*, *suggesting*, and *supposing*.
69. Mindt (2000:250) states that “[n]on-progressive meaning is characterized by the absence of the meanings incompleteness, temporariness, iteration/habit and comprises the presence of at least one of the following meanings: highlighting/prominence, emotion, politeness/downtoning, prediction, volition/intention, matter-of-course.”
70. This emphatic use of progressives is comparable to the “foregrounded progressive” discussed by Couper-Kuhlen 1995.
71. There were no progressive forms in the concordances of *mattering*.
72. As *knowing* only occurs once in a progressive construction, the PastProg share of 100% cannot be considered representative.
73. CEC is the abbreviation of Corpus of English Conversation which is part of the 500,000 word London-Lund Corpus of spoken British English. The Lancaster-Oslo/Bergen Corpus (LOB) and the Brown Corpus both consist of exclusively written material (one million words each; LOB: British English, Brown: American English).
74. Support for this decision can be found in Williams’s non-CDL study in which the author states that “[t]he ‘be going to’ form is to be considered as a form of the progressive” (Williams 2002:197).

Chapter 5

75. I would like to thank Ulla Dvořák and Sven Naujokat for providing me with these examples from their pupils's essays and classroom discussions.
76. I am most grateful to Karsten Lütteken for his invaluable technical help in the compilation of GEFL TC.
77. To my knowledge, all previous linguistic analyses of German EFL teaching materials have been carried out manually. Norbert Schlüter (personal communication) confirms this with reference to his own and his colleague Dieter Mindt's work.
78. This in part echoes Hunston and Francis's (2000: 15) definition of corpus linguistics as "a way of investigating language by observing large amounts of naturally-occurring, electronically-stored discourse, using software which selects, sorts, matches, counts and calculates."
79. Note that Willis (2003: 223) uses the term "pedagogic corpus" in a somewhat different sense with reference to a corpus that "is made up of those texts which learners have read or listened to in the course of their studies." The texts Willis has in mind for inclusion in this pedagogic corpus are "natural texts rather than texts specifically designed for language teaching" (2003: 224).
80. According to the chi-square test, the chance probability is less than 0.01.
81. The shares of fragmentary constructions with a missing form of to be have been omitted here. Of the 22 selected verb forms, they only affect *doing* (cf. the shorter bar charts in the table).
82. See the highlighted parts in Table 39. All above-average shares have been shaded grey.

Chapter 7

83. For two recent collections which are dedicated to the latest developments in the field of corpus linguistics and language teaching, see Aston, Bernardini & Stewart 2004 and Sinclair 2004a.
84. Being shown the textbook passage given in (267), a native speaker of American English said that he thought the use of "Is she packing?" in a textbook for ten-year-olds was indeed rather odd. He said that, at least in colloquial American English, "Is she/he packing?" would often be used in a "gangster" context, meaning "Is she/he carrying a gun (down her/his trousers)?" In fact, plenty of evidence for this usage can be found on the web, often in drama or film scripts or in chat-room conversations. Here are a couple of illustrative examples: *People stare at the girl in baggy pants, black nails, black shirt and chains. It makes me wonder what they're thinking. "Is she packing?"; Am I seeing things or is he packing a thingy down there?; I saw Ed in some tights one night, and I figured you know, since you did the damn thing and all... Well I know you know... Is he PACKING? huh?; How many weapons is she packing anyway?*
85. It can be argued in this context that coursebook dialogues put too much emphasis on semantic meaning at the expense of pragmatic meaning, and that the concocted examples tend to be pragmatically deprived, since their only purpose is to introduce new vocabulary items and a certain grammatical structure in a meaningful exchange (cf. also Lausevic 2002).
86. Michael McCarthy made this statement during a discussion at the first Inter-Varietal Applied Corpus Studies Group (IVACS) Conference in Limerick, Ireland, June 15th 2002.

87. In this context, de Beaugrande (2001b:25) convincingly points out “that working with non-authentic data is certainly no safeguard against political incorrectness either” and discusses the invented textbook example “the boy kicked the girl”.

88. See for instance the so-called *Richtlinien* published by the Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (1993:37), which state: “Ziel des Englischunterrichts ist es, die Fähigkeit und Bereitschaft der Schülerinnen und Schüler zur Kommunikation in der englischen Sprache zu entwickeln. Sie sollen befähigt werden, ihren Absichten entsprechend in für sie bedeutsamen Realsituationen sach- und situationsgerecht sprachlich handeln zu können.” [It is the aim of ELT to develop the ability and preparedness of pupils to communicate in the English language. They are supposed to be enabled to act in purposeful and contextually relevant ways according to their (communicative) purposes in actual speech situations that are important to them.]

89. On the importance of and problems with the issue of frequencies and language teaching, see also Bald et al. 1972, Barlow 2003, Leech 1997, Mindt 1997b, and van Els et al. 1984.

90. This is in line with what Willis (2003:95–98) has found. He also gives an example of an “unsystematic and incomplete” textbook description of progressives and their functions.

91. The average share of negation we determined for BNC_spoken and BoE_brspok progressives taken together was 8.39%.

92. For a pioneering approach to language teaching which takes lexis as a starting point, the reader is referred to Willis’s (1990) monograph *The Lexical Syllabus*. A recent review of the lexical approach is provided in an article by Harwood (2002).

93. A similar pedagogic account could be imagined to be given, for instance, on General American English or on International or lingua franca English.

94. See also Carter & McCarthy (1995:141), who make a case for basing language teaching on the grammar of spoken English since it helps a lot “to foster speaking skills and natural spoken interaction”.

95. The percentages in Table 56 relate to both spoken English corpora taken together. They have been collapsed from the BNC_spoken and the BoE_brspok Figures (cf. Sections 4.3 and 4.4).

96. The functions “general validity”, “politeness/softening”, “emphasis/attitude/shock”, and “gradual change and development” were found to occur in addition to the four combinations of central function features (“continuousness” and “repeatedness”). The sum of the given percentages thus amounts to more than 100%.

97. Sinclair 2003 offers a very good training on how to read concordances.

98. This concept of using corpus data and concordances directly in the classroom is usually referred to as “data-driven learning” (DDL). Detailed information on DDL, including a number of practical suggestions of implementing it, can be found in Bernardini 2000, 2002, and 2004, Bodin 2002, Fox 1998, Granger & Tribble 1998, Hadley 2000, Johns 1994 and 2002, Johns & King 1991, Kettemann 1995, Stevens 1995, and Tribble & Jones 1990.

99. In his critical discussion of the COBUILD approach, Owen (1993:185) claims that “[t]he grammarian and the language teacher need the corpus as servant, not as master.” See also the response to Owen’s paper by Francis & Sinclair 1994.

Chapter 8

100. On a similar note, Michael Hoey, in a discussion at the 24th ICAME conference in April 2003 in Guernsey, UK, called it “bad science” when researchers only look at the data in order to illustrate what they already know.

101. The importance of adequate observations that have to precede any kind of explaining or theorising steps was stressed by Tognini-Bonelli in a panel discussion on “Grammar and corpus linguistics” at the 24th ICAME conference in April 2003 in Guernsey, UK.

102. For some recent state-of-the-art accounts on learner corpus resources and their research potential, the reader is referred to Barlow 2005, Granger 2004, Nesselhauf 2004, and Pravec 2002.

103. LINDSEI, the Louvain International Database of Spoken English Interlanguage, is a spoken learner corpus project coordinated by Sylviane Granger at the Catholic University of Louvain, Belgium. See <http://www.fltr.ucl.ac.be/fltr/germ/etan/cecl/Cecl-Projects/Lindsei/> (consulted: 22.08.04).

104. Note however, that the American response to the BNC, the American National Corpus (ANC), is currently being compiled. The first release of ANC contains 3.2 million words of spoken American English. For more information, see <http://americannationalcorpus.org/> (consulted: 22.08.04).

References

- Aarts, B. (2000). Corpus linguistics, Chomsky, and Fuzzy Tree Fragments. In C. Mair & M. Hundt (Eds.), *Corpus Linguistics and Linguistic Theory: Papers from the Twentieth International Conference on English Language Research on Computerized Corpora (ICAME 20), Freiburg im Breisgau 1999* (pp. 5–13). Amsterdam: Rodopi.
- Aarts, J. (2000). Towards a new generation of corpus-based English grammars. In B. Lewandowska-Tomaszczyk & P. J. Melia (Eds.), *PALC '99: Practical Applications in Language Corpora* (pp. 17–36). Frankfurt: Lang.
- Aarts, J. (2002). Does corpus linguistics exist? Some old and new issues. In L. E. Breivik & A. Hasselgren (Eds.), *From the Colt's Mouth ... and Others. Language Corpora Studies. In Honour of Anna-Brita Stenström* (pp. 1–17). Amsterdam: Rodopi.
- Abney, S. (1996). Statistical methods and linguistics. In J. Klavans & P. Resnik (Eds.), *The Balancing Act: Combining Symbolic and Statistical Approaches to Language* (pp. 1–26). Cambridge, MA: MIT Press.
- Albert, R., & C. J. Koster (2002). *Empirie in Linguistik und Sprachlehrforschung*. Tübingen: Narr.
- Allen, R. L. (1966). *The Verb System of Present-day American English*. The Hague: Mouton.
- Allen, J. P. B., & H. G. Widdowson (1974). Teaching the communicative use of English. *IRAL* XII(1), 1–20. Reprinted in J. Swales (Ed.). (1985). *Episodes in ESP* (pp. 73–87). Oxford: Pergamon.
- Arndt, H. (1969). Wissenschaftliche Grammatik und pädagogische Grammatik. *Neusprachliche Mitteilungen aus Wissenschaft und Praxis*, 22, 65–76.
- Ashford, S. et al. (Eds.). (1996). *Learning English – Green Line New 2*. Stuttgart: Klett.
- Ashford, S. et al. (Eds.). (1997). *Learning English – Green Line New 3*. Stuttgart: Klett.
- Ashford, S. et al. (Eds.). (1998). *Learning English – Green Line New 4*. Stuttgart: Klett.
- Ashford, S. et al. (Eds.). (1999). *Learning English – Green Line New 5*. Stuttgart: Klett.
- Ashford, S. et al. (Eds.). (2000). *Learning English – Green Line New 6*. Stuttgart: Klett.
- Aston, G. (1997). Enriching the learning environment: Corpora in ELT. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 51–64). London: Longman.
- Aston, G. (2000). Corpora and language teaching. In L. Burnard & T. McEnery (Eds.), *Rethinking Language Pedagogy from a Corpus Perspective* (pp. 7–17). Frankfurt: Lang.
- Aston, G., & L. Burnard (1998). *The BNC Handbook. Exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.
- Aston, G., S. Bernardini, & D. Stewart (Eds.). (2004). *Corpora and Language Learners*. Amsterdam: John Benjamins.
- Aston, P. et al. (Eds.). (1995). *Learning English – Green Line New 1*. Stuttgart: Klett.
- Bache, C., & N. Davidsen-Nielsen (1997). *Mastering English. An Advanced Grammar for Non-Native and Native Speakers*. Berlin: Mouton de Gruyter.

- Bachmann, L. F. (1990). *Fundamental Considerations in Language Testing*. Reading, MA: Addison-Wesley Publishing Company.
- Bald, W.-D. (1977). *Testmethoden und Linguistische Theorie. Beispiele und Übungen zur Linguistischen Methodologie*. Tübingen: Niemeyer.
- Bald, W.-D. (1995). Gegenstands- und Zielbestimmungen der anglistischen Sprachwissenschaft im Lichte der fachinternen Entwicklungen. In B. Korte & K. P. Müller (Eds.), *Anglistische Lehre Aktuell: Probleme, Perspektiven, Praxis* (pp. 93–107). Trier: Wissenschaftlicher Verlag Trier.
- Bald, W.-D., B. Carstensen, & M. Hellinger (Eds.). (1972). *Die Behandlung grammatischer Probleme in Lehrwerken für den Englischunterricht*. Frankfurt: Diesterweg.
- Barlow, M. (1996). Corpora for theory and practice. *IJCL*, 1 (1), 1–37.
- Barlow, M. (2003). Corpora and language teaching. *KATE Forum*, 27 (1), 6–9.
- Barlow, M. (2005). Computer-based analyses of learner language. In G. Barkhuizen & R. Ellis (Eds.), *Analysing Learner Language* (pp. 335–358). Oxford: Oxford University Press.
- Barlow, M., & S. Kemmer (Eds.). (2000). *Usage-Based Models of Language*. Stanford: CSLI Publications.
- Barnbrook, G. (1996). *Language and Computers: A Practical Introduction to the Computer Analysis of Language*. Edinburgh: Edinburgh University Press.
- Bausch, K.-R. (Ed.). (1979). *Beiträge zur Didaktischen Grammatik. Probleme, Konzepte, Beispiele*. Königstein: Scriptor.
- Bausch, K.-R., H. Christ, & H.-J. Krumm (Eds.). (1995). *Handbuch Fremdsprachenunterricht*. Tübingen: Francke.
- Bazell, C. E., J. C. Catford, M. A. K. Halliday, & R. H. Robins (Eds.). (1966). *In Memory of J. R. Firth*. London: Longmans, Green and Co Ltd.
- Beaugrande, R. de (2001a). 'If I were you...': Language standards and corpus data in EFL. *Revista Brasileira de Linguística Aplicada*, 1 (1), 117–154. Available online: <http://beaugrande.bizland.com/Ifiwereyou.htm> (consulted: 22.08.04).
- Beaugrande, R. de (2001b). Twenty challenges to corpus research. And how to answer them. Available online: <http://beaugrande.bizland.com/Twenty%20questions%20about%20corpus%20research.htm> (consulted: 22.08.04).
- Berber Sardinha, T. (1996). WordSmith Tools. *Computers & Texts*, 12, 19–21.
- Berglund, Y. (2000). "You're gonna, you're not going to": A corpus-based study of colligation and collocation patterns of the (BE) *going to* construction in present-day spoken British English. In B. Lewandowska-Tomaszczyk & P. J. Melia (Eds.), *PALC '99: Practical Applications in Language Corpora* (pp. 161–192). Frankfurt: Lang.
- Bernardini, S. (2000). *Competence, Capacity, Corpora. A Study in Corpus-aided Language Learning*. Bologna: CLUEB.
- Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann & G. Marko (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 165–182). Amsterdam: Rodopi.
- Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. McH. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 15–36). Amsterdam: John Benjamins.
- Berry, R. (2001). Understanding tense and aspect in English: The roles of creativity and convention. Paper presented at IATEFL 2001, Brighton, April 2001.
- Biber, D., S. Conrad, & R. Reppen (1994). Corpus-based approaches to issues in applied linguistics. *Applied Linguistics*, 15 (2), 169–189.

- Biber, D., S. Conrad, & R. Reppen (1998). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Biber, D., S. Johansson, G. Leech, S. Conrad, & E. Finegan (1999). *Longman Grammar of Spoken and Written English*. London: Longman.
- Bland, S. K. (1988). The present progressive in discourse: Grammar versus language usage revisited. *TESOL Quarterly*, 22, 53–68.
- Bloomfield, L. (1933). *Language*. London: Allen & Unwin.
- Bodelsen, C. A. (1936). The expanded tenses in modern English. An attempt at an explanation. *Englische Studien*, 71, 220–238.
- Bodin, E. (2002). Classroom concordancing options. In M. Tan (Ed.), *Corpus Studies in Language Education* (pp. 179–185). Bangkok: IELE Press.
- Börner, W., & K. Vogel (1976). Zum Verhältnis von wissenschaftlicher und pädagogischer Grammatik im Fremdsprachenunterricht. In W. Börner, B. Kielhöfer, & K. Vogel. *Französisch lehren und lernen* (pp. 7–39). Kronberg: Scriptor.
- Bowker, L., & J. Pearson (2002). *Working with Specialized Language. A Practical Guide to Using Corpora*. London: Routledge.
- Brazil, D. (1995). *A Grammar of Speech*. Oxford: Oxford University Press.
- Breen, M. P. (1985). Authenticity in the language classroom. *Applied Linguistics*, 6 (1), 60–69.
- Breen, M. P., & C. N. Candlin (1980). The essentials of a communicative curriculum in language teaching. *Applied Linguistics*, 1 (2), 89–112.
- Bublitz, W. (1995). Tense, aspect, and modality: Introduction. In W. Riehle & H. Keiper (Eds.), *Anglistentag 1994 Graz: Proceedings* (pp. 135–137). Tübingen: Niemeyer.
- Bublitz, W. (1996). Semantic prosody and cohesive company: 'Somewhat predictable'. *Leuvense Bijdragen (Leuven Contributions in Linguistics and Philology)*, 85 (1–2), 1–32.
- Burnard, L. (Ed.). (1995). *Users' Reference Guide for the British National Corpus Version 1.0*. Oxford: Oxford University Computing Services.
- Burnard, L. (Ed.). (2000). *Reference Guide for the British National Corpus (World Edition)*. Oxford: Oxford University Computing Services.
- Burnard, L. (2002). Where did we go wrong? A retrospective look at the British National Corpus. In B. Kettemann & G. Marko (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 51–70). Amsterdam: Rodopi.
- Burnard, L. (Ed.). (2003). *Reference Guide for BNC-baby*. Oxford: Oxford University Computing Services.
- Butler, C. (1985). *Statistics in Linguistics*. New York: Blackwell.
- Bybee, J., R. Perkins, & W. Pagliuca (1994). *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: Chicago University Press.
- Canale, M. (1983). From communicative competence to communicative language pedagogy. In J. C. Richards & R. W. Schmidt (Eds.), *Language and Communication* (pp. 2–27). London: Longman.
- Canale, M., & M. Swain (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1 (1), 1–47.
- Canale, M., & M. Swain (1988). Some theories of communicative competence. In W. Rutherford & M. Sharwood-Smith (Eds.), *Grammar and Second Language Teaching* (pp. 61–84). New York: Newbury House.
- Candlin, C. N. (1973). The status of pedagogical grammar. In S. P. Corder & E. Roulet (Eds.), *Theoretical Linguistic Models and Applied Linguistics* (pp. 55–64). Paris: Didier.

- Candlin, C. N. (1976). Communicative language teaching and the debt to pragmatics. In C. Ramesh (Ed.), *Georgetown University Round Table on Language and Linguistics 1976* (pp. 237–256). Washington, DC: Georgetown University Press.
- Carter, R. (1998). Orders of reality: CANCODE, communication, and culture. *ELT Journal*, 52 (1), 43–56.
- Carter, R., & M. McCarthy (1995). Grammar and the spoken language. *Applied Linguistics*, 16 (2), 141–158.
- Chalker, S. (1994). Pedagogical grammar: Principles and problems. In M. Bygate, A. Tonkyn, & E. Williams (Eds.), *Grammar and the Language Teacher* (pp. 31–44). Hemel Hempstead: Prentice Hall.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1966). *Topics in the Theory of Generative Grammar*. The Hague: Mouton.
- Church, K., W. Gale, P. Hanks, & D. Hindle (1991). Using statistics in lexical analysis. In U. Zernik (Ed.), *Lexical Acquisition: Exploiting On-line Resources to Build a Lexicon* (pp. 115–164). Hillsdale, NJ: Lawrence Erlbaum Associates. Available online: <http://www.patrickhanks.com/papers/usingStats.pdf> (consulted: 22.08.04).
- Clear, J. (1988). Trawling the language: Monitor corpora. In M. Snell-Hornby (Ed.), *Zürilex '86 Proceedings* (pp. 383–389). Tübingen: Francke.
- Clear, J. (1993). From Firth principles. Computational tools for the study of collocation. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 271–292). Amsterdam: John Benjamins.
- Cobuild (1990). *Collins COBUILD English Grammar* ed. J. McH. Sinclair et al. London: HarperCollins.
- Cobuild (1992). *Collins COBUILD English Usage* ed. J. McH. Sinclair et al. London: HarperCollins.
- Cobuild (32001). *Collins COBUILD English Dictionary for Advanced Learners*, ed. J. McH. Sinclair et al. London: HarperCollins.
- Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.
- Comrie, B. (1985). *Tense*. Cambridge: Cambridge University Press.
- Comrie, B. (1998). Tense. In J. L. Mey (Ed.), *Concise Encyclopedia of Pragmatics* (pp. 994–999). Oxford: Elsevier.
- Conan Doyle, A. (1892). *The Adventures of Sherlock Holmes*. London: Penguin.
- Conrad, S. (2000). Will corpus linguistics revolutionize grammar teaching in the 21st Century? *TESOL Quarterly*, 34, 548–560.
- Conrad, S. (2004). Corpus linguistics, language variation, and language teaching. In J. McH. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 67–85). Amsterdam: John Benjamins.
- Cook, G. (1998). The uses of reality: A reply to Ronald Carter. *ELT Journal*, 52 (1), 57–62.
- Cook, G. (2001). 'The philosopher pulled the lower jaw of the hen.' Ludicrous invented sentences in language teaching. *Applied Linguistics*, 22 (3), 366–387.
- Cook, G. (2003). The functions of example sentences: A reply to Vivian Cook. *Applied Linguistics*, 24 (2), 249–255.
- Cook, V. (2002). The functions of invented sentences: A reply to Guy Cook. *Applied Linguistics*, 23 (2), 262–269.
- Corder, S. P. (1973). *Introducing Applied Linguistics*. London: Penguin.
- Corder, S. P. (1988). Pedagogic grammars. In W. Rutherford & M. Sharwood-Smith (Eds.), *Grammar and Second Language Teaching* (pp. 123–145). New York: Newsbury House.

- Couper-Kuhlen, E. (1995). On the foregrounded progressive in American conversational narrative: A new development. In W. Riehle & H. Keiper (Eds.), *Anglistentag 1994 Graz: Proceedings* (pp. 229–246). Tübingen: Niemeyer.
- Croft, W., & A. Cruse (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Crowdy, S. (1993). Spoken corpus design. *Literary and Linguistic Computing*, 8 (4), 259–265.
- Crowdy, S. (1995). The BNC spoken corpus. In G. Leech, G. Myers, & J. Thomas (Eds.), *Spoken English on Computer: Transcription, Mark-up and Application* (pp. 224–234). London: Longman.
- Crystal, D. (1966). Specification of English tenses. *Journal of Linguistics*, 2, 1–34.
- Dahl, Ö. (1998). Aspect. In J. L. Mey (Ed.), *Concise Encyclopedia of Pragmatics* (pp. 64–71). Oxford: Elsevier.
- Declerck, R. (1991). *Tense in English. Its Structure and Use in Discourse*. London: Routledge.
- Dietrich, G. (1955). *Erweiterte Form, Präteritum und Perfektum im Englischen*. Munich: Hueber.
- Dirven, R. (1990). Pedagogical grammar. *Language Teaching*, 23 (1), 1–18.
- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19 (1), 61–74. Available online: www.comp.lancs.ac.uk/ucrel/papers/tedstats.pdf (consulted: 22.08.04).
- Eastwood, J. (1994). *Oxford Guide to English Grammar*. Oxford: Oxford University Press.
- Ek, J. A. van (1969). The ‘progressive’ reconsidered. *English Studies*, 50, 579–585.
- Els, T. van, T. Bongaerts, G. Extra, C. van Os, & A.-M. Janssen-van Dieten (1984). *Applied Linguistics and the Learning and Teaching of Foreign Languages*. London: Edward Arnold.
- Elsness, J. (1994). On the progression of the progressive in Early Modern English. *ICAME Journal*, 18, 5–25.
- Esser, J. (1999). Collocation, colligation, semantic preference and semantic prosody: New developments in the study of syntagmatic word relations. In W. Falkner & H.-J. Schmid (Eds.), *Words, Lexemes, Concepts – Approaches to the Lexicon. Studies in Honour of Leonhard Lipka* (pp. 155–165). Tübingen: Narr.
- Firth, J. R. (1957). *Papers in Linguistics 1934–1951*. London: Oxford University Press.
- Fleischhack, E., & H. Schwarz (2001). *Cornelsen English Grammar*. Berlin: Cornelsen.
- Fox, G. (1998). Using corpus data in the classroom. In B. Tomlinson (Ed.), *Materials Development in Language Teaching* (pp. 25–43). Cambridge: Cambridge University Press.
- Francis, G. (1993). A corpus-driven approach to grammar. Principles, methods and examples. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 138–156). Amsterdam: John Benjamins.
- Francis, G., & J. McH. Sinclair (1994). ‘I bet he drinks Carling Black Label’: A riposte to Owen on corpus grammar. *Applied Linguistics*, 15 (2), 190–202.
- Francis, G., S. Hunston, & E. Manning (Eds.). (1996a). *Grammar Patterns 1: Verbs*. London: HarperCollins.
- Francis, G., S. Hunston, & E. Manning (Eds.). (1996b). *Grammar Patterns 2: Nouns and Adjectives*. London: HarperCollins.
- Freckmann, N. (1995). The progressive and adverbial collocations: Corpus evidence. In W. Riehle & H. Keiper (Eds.), *Anglistentag 1994 Graz. Proceedings* (pp. 255–276). Tübingen: Niemeyer.
- Gavioli, L. (2002). Some thoughts on the problem of representing ESP through small corpora. In B. Kettemann & G. Marko (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 293–303). Amsterdam: Rodopi.
- Gehring, W. (2004). *Englische Fachdidaktik*. Berlin: Schmidt.

- Ghadessy, M., A. Henry, & R. L. Roseberry (Eds.). (2001). *Small Corpus Studies and ELT. Theory and Practice*. Amsterdam: John Benjamins.
- Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal*, 58 (4), 363–374.
- Glisan, E., & V. Drescher (1993). Textbook grammar: Does it reflect native speaker speech? *The Modern Language Journal*, 77 (1), 23–33.
- Gnutzmann, C. (Ed.). (1999). *Teaching and Learning English as a Global Language. Native and Non-native Perspectives*. Tübingen: Stauffenburg.
- Gnutzmann, C., & F. G. Königs (Eds.). (1995). *Perspektiven des Grammatikunterrichts*. Tübingen: Narr.
- Görlach, M. (1991). *Introduction to Early Modern English*. Cambridge: Cambridge University Press.
- Görlach, M. (1999). Varieties of English and language teaching. In C. Gnutzmann (Ed.), *Teaching and Learning English as a Global Language. Native and Non-native Perspectives* (pp. 3–21). Tübingen: Stauffenburg.
- Görlach, M. (2001). Review: Dieter Mindt, *An Empirical Grammar of the English Verb System. Arbeiten aus Anglistik und Amerikanistik*, 26 (2), 221–224.
- Görlach, M. (2003). Some reflexions on corpus linguistics. In M. Görlach. *Topics in English Historical Linguistics* (pp. 18–27). Heidelberg: Winter.
- Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: The University of Chicago Press.
- Gouverneur, C., & F. Meunier (2004). The treatment of phraseology in ELT textbooks. Paper presented at the 6th Teaching and Language Corpora conference (TaLC 6), Granada, 6–9 July 2004.
- Grabowski, E., & D. Mindt (1994). Die unregelmäßigen Verben des Englischen: eine Lernliste auf empirischer Grundlage. *Die Neueren Sprachen*, 93 (4), 334–353.
- Grabowski, E., & D. Mindt (1995). A corpus-based learning list of irregular verbs in English. *ICAME Journal*, 19, 5–22.
- Granger, S. (2004). Computer learner corpus research: Current status and future prospects. In U. Connor & Upton T. (Eds.), *Applied Corpus Linguistics: A Multidimensional Perspective* (pp. 123–145). Amsterdam: Rodopi.
- Granger, S., & C. Tribble (1998). Learner corpus data in the foreign language classroom: Form-focused instruction and data-driven learning. In S. Granger (Ed.), *Learner English on Computer* (pp. 199–209). London: Longman
- Graustein, G. et al. (1987). *English Grammar. A University Handbook*. Leipzig: VEB Enzyklopädie.
- Greenbaum, S. (1988). *Good English and the Grammarian*. London: Longman.
- Greenbaum, S. (1996). *The Oxford English Grammar*. Oxford: Oxford University Press.
- Greenbaum, S. (2000). *The Oxford Reference Grammar* (Ed. by E. Weiner). Oxford: Oxford University Press.
- Haase, I. (1995). *Konditionalsätze im authentischen englischen Sprachgebrauch und im Englischunterricht. Eine empirische Untersuchung*. Dissertation im Fachbereich Neuere Fremdsprachliche Philologien, Englische Philologie. Berlin: Privatdruck.
- Hadley, G. (2000). Sensing the winds of change: An introduction to data-driven learning. Available online: <http://web.bham.ac.uk/johnstf/winds.htm> (consulted: 22.08.04).
- Haegeman, L., & J. Guéron (1999). *English Grammar. A Generative Perspective*. Oxford: Blackwell.

- Hahn, A., S. Reich, & J. Schmied (2000). Aspect in the Chemnitz internet grammar. In C. Mair & M. Hundt (Eds.), *Corpus Linguistics and Linguistic Theory: Papers from the Twentieth International Conference on English Language Research on Computerized Corpora (ICAME 20)*, Freiburg im Breisgau 1999 (pp. 131–139). Amsterdam: Rodopi.
- Harris, Z. S. (1951). *Methods in Structural Linguistics*. Chicago: University of Chicago Press.
- Harwood, N. (2002). Taking a lexical approach to teaching: Principles and problems. *IJAL*, 12 (2), 139–155.
- Hatav, G. (1993). The aspect system in English: An attempt at a unified analysis. *Linguistics*, 31, 209–237.
- Hatcher, A. G. (1951). The use of the progressive form in English: A new approach. *Language*, 27, 254–280.
- Hedge, T. (2000). *Teaching and Learning in the Language Classroom*. Oxford: Oxford University Press.
- Heine, B., U. Claudi, & F. Hünemeyer (1991). *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.
- Hirtle, W. H. (1967). *The Simple and Progressive Forms. An Analytical Approach*. Québec: Les Presses de l'Université Laval.
- Hirtle, W. H., & V. N. Curat (1986). *The simple and the progressive*. Oxford: Blackwell.
- Honeybone, P. (Forthcoming). J. R. Firth. In S. Chapman & C. Routledge (Eds.), *Key Thinkers in Linguistics and the Philosophy of Language*. Edinburgh: Edinburgh University Press.
- Hopper, P. J., & E. C. Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.
- House, J. (1999). Misunderstanding in intercultural communication: Interactions in English as *lingua franca* and the myth of mutual intelligibility. In C. Gnutzmann (Ed.), *Teaching and Learning English as a Global Language. Native and Non-native Perspectives* (pp. 73–89). Tübingen: Stauffenburg.
- Hüllen, W. (1973). Linguistische Modelle und Didaktische Grammatik. *Neusprachliche Mitteilungen aus Wissenschaft und Praxis*, 26, 8–16.
- Hüllen, W. (1976). *Linguistik und Englischunterricht. Didaktische Analysen*. Heidelberg: Quelle & Meyer.
- Huddleston, R., & G. K. Pullum (Eds.). (2002). *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hudson, R. (1992). *Teaching Grammar. A Guide for the National Curriculum*. Oxford: Blackwell.
- Hunston, S. (2002a). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Hunston, S. (2002b). Pattern grammar, language teaching, and linguistic variation: Applications of a corpus-driven grammar. In R. Reppen, S. M. Fitzmaurice, & D. Biber (Eds.), *Using Corpora to Explore Linguistic Variation* (pp. 167–183). Amsterdam: John Benjamins.
- Hunston, S., & G. Francis (1998). Verbs observed: A corpus-driven pedagogic grammar. *Applied Linguistics*, 19 (1), 45–72.
- Hunston, S., & G. Francis (2000). *Pattern Grammar. A Corpus-driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins.
- Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269–293). Harmondsworth: Penguin.
- Hymes, D. (1992). The concept of communicative competence revisited. In M. Pütz (Ed.), *Thirty Years of Linguistic Evolution. Studies in Honour of René Dirven on the Occasion of his Sixtieth Birthday* (pp. 31–57). Amsterdam: John Benjamins.
- Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford: Oxford University Press.

- Jespersen, O. (1931). The expanded tenses. *SPE Tract*, 36, 524–530.
- Johansson, S., & H. Stavestrand (1987). Problems in learning – and teaching – the progressive form. In I. Lindblad & M. Ljung (Eds.), *Proceedings from the Third Nordic Conference for English Studies, Hässelby, Sept 25–27, 1986, vol. 1* (pp. 139–148). Stockholm: Almqvist & Wiksell.
- Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (Ed.), *Perspectives on Pedagogical Grammar* (pp. 27–45). Cambridge: Cambridge University Press.
- Johns, T. (2002). Data-driven learning: The perpetual challenge. In B. Kettemann & G. Marko (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 107–117). Amsterdam: Rodopi.
- Johns, T., & P. King (Eds.). (1991). *Classroom Concordancing [ELR Journal 4]*.
- Joos, M. (1964). *The English Verb. Form and Meanings*. Madison: The University of Wisconsin Press.
- Jørgensen, E. (1991). The progressive tenses and the so-called ‘non-conclusive’ verbs. *English Studies*, 72 (2), 173–182.
- Jung, L. (1975). *Linguistische Grammatik und Didaktische Grammatik*. Frankfurt a.M.: Diesterweg.
- Kennedy, G. (1998). *An Introduction to Corpus Linguistics*. London: Longman.
- Kettemann, B. (1995). On the use of concordancing in ELT. *Arbeiten aus Anglistik und Amerikanistik*, 20, 29–41.
- Kettemann, B., & G. Marko (Eds.). (2002). *Teaching and Learning by Doing Corpus Analysis*. Amsterdam: Rodopi.
- Kilgarriff, A. (2001). Comparing corpora. *IJCL*, 6 (1), 97–133.
- Klein, E. (1987). Ein Netzwerk-Modell für die Erstellung didaktischer und pädagogischer Grammatiken. *Die Neueren Sprachen*, 86 (6), 517–530.
- Klein, H. G. (1974). *Tempus, Aspekt, Aktionsart*. Tübingen: Niemeyer.
- Klein, W. (1995). A simplest analysis of the English tense-aspect system. In W. Riehle & H. Keiper (Eds.), *Anglistentag 1994 Graz. Proceedings* (pp. 139–151). Tübingen: Niemeyer.
- Knapp, K., & C. Meierkord (Eds.). (2002). *Lingua Franca Communication*. Frankfurt: Lang.
- Knowles, G., A. Wichmann, & P. Alderson (Eds.). (1996). *Working with Speech: Perspectives on Research into the Lancaster/IBM Spoken English Corpus*. London: Longman.
- König, E. (1995). On analyzing the tense-aspect system of English: A state-of-the-art report. In W. Riehle & H. Keiper (Eds.), *Anglistentag 1994 Graz. Proceedings* (pp. 153–169). Tübingen: Niemeyer.
- Königs, K. (2000). *Übersetzen Englisch – Deutsch. Systemischer Ansatz*. Munich: Oldenbourg.
- Krause, O. (2002). *Progressiv im Deutschen. Eine empirische Untersuchung im Kontrast mit Niederländisch und Englisch*. Tübingen: Niemeyer.
- Krishnamurthy, R. (2002a). The Bank of English past, present, and future: corpus size, composition, annotation, and software. Paper presented at the 2nd ILASH half-day workshop on “Computational Language Resources”, University of Sheffield, Feb 8th 2002. Available online: <http://www.dcs.shef.ac.uk/research/ilash/Seminars/rameshLR2.doc> (consulted: 22.08.04).
- Krishnamurthy, R. (2002b). Classroom cornucopia: The new COBUILD dictionary and the Bank of English corpus. In M. Tan (Ed.), *Corpus Studies in Language Education* (pp. 121–128). Bangkok: IELE Press.
- Lampater, P., & P. Pasch (1997). *Learning English – Green Line New 1 und 2: Grammatisches Beiheft*. Stuttgart: Klett.

- Lampater, P., & P. Pasch (1998). *Learning English – Green Line New 3 und 4: Grammatisches Beiheft*. Stuttgart: Klett.
- Landman, F. (1992). The progressive. *Natural Language Semantics*, 1, 1–32.
- Langacker, R. W. (2000a). Cognitive linguistics, language pedagogy, and the English present tense. *LAUD Paper* No. 298, Series B. Essen: LAUD.
- Langacker, R. W. (2000b). A dynamic usage-based model. In M. Barlow & S. Kemmer (Eds.), *Usage-Based Models of Language* (pp. 1–63). Stanford: CSLI Publications.
- Lausevic, C. (2002). Spoken discourse: From linguistics to pedagogy. In C. Finkbeiner (Ed.), *Wholeheartedly English: A Life of Learning. Festschrift for Johannes-Peter Timm* (pp. 196–209). Berlin: Cornelsen.
- Leech, G. (1971). *Meaning and the English Verb*. London: Longman.
- Leech, G. (1988). Varieties of English grammar: The state of the art from the grammarian's point of view. In W.-D. Bald (Ed.), *Kernprobleme der Englischen Grammatik* (pp. 19–37). Munich: Langenscheidt-Longman.
- Leech, G. (1992). Corpora and theories of linguistic performance. In J. Svartvik (Ed.), *Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4–8 August 1991* (pp. 105–122). Berlin: Mouton de Gruyter.
- Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McNery, & G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 1–23). London: Longman.
- Leech, G., & J. Svartvik (1975). *A Communicative Grammar of English*. London: Longman.
- Leech, G., & J. Svartvik (²1994). *A Communicative Grammar of English*. London: Longman.
- Legenhausen, L. (1985). Die SF/EF-Distribution in englischen Rundfunkreportagen. *Anglistik und Englischunterricht*, 27, 117–135.
- Leisi, E. (1960). Die Progressive Form im Englischen. *Die Neueren Sprachen N. F.*, 9, 217–226.
- Lenko-Szymanska, A. (2004). Past Progressive or Simple Past? The acquisition of the progressive aspect by Polish advanced learners of English. Paper presented at the 6th Teaching and Language Corpora conference (TaLC 6), Granada, 6–9 July 2004.
- Levickij, V. V., & T. A. Romanova (1997). Use of tenses of verbs and adverbs in the English language: A statistical study. *Journal of Quantitative Linguistics*, 4 (1–3), 135–138.
- Lewis, M. (1993). *The Lexical Approach. The State of ELT and a Way Forward*. Hove: Language Teaching Publications.
- Lewis, M. (Ed.). (2000). *Teaching Collocation. Further Developments in the Lexical Approach*. Hove: Language Teaching Publications.
- Ljung, M. (1980). *Reflections on the English Progressive*. Gothenburg: Gotab.
- Ljung, M. (1990). *A Study of TEFL Vocabulary*. Stockholm Studies in English 78. Stockholm: Almqvist & Wiksell International.
- Lock, G. (1996). *Functional English Grammar. An Introduction for Second Language Teachers*. Cambridge: Cambridge University Press.
- Lorenz, G. (2002). Language corpora rock the base: On standard English grammar, perfective aspect and seemingly adverse corpus evidence. In B. Kettemann & G. Marko (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 131–145). Amsterdam: Rodopi.
- Louw, W. E. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 152–176). Amsterdam: John Benjamins.
- Lucko, P. (1995). Between aspect, actionality and modality: The function of the expanded form. In W. Riehle & H. Keiper (Eds.), *Anglistentag 1994 Graz. Proceedings* (pp. 171–182). Tübingen: Niemeyer.

- Mair, C. (1997). On the advantages of not being a native speaker... : The continental tradition in English linguistics. In U. Fries, V. Müller, & P. Schneider (Eds.), *From Ælfric to the New York Times. Studies in English Corpus Linguistics* (pp. 249–256). Amsterdam: Rodopi.
- Mair, C., & M. Hundt (1995). Why is the progressive becoming more frequent in English? *Zeitschrift für Anglistik und Amerikanistik*, 43 (2), 111–122.
- Markus, M. (1977). *Tempus und Aspekt. Zur Funktion von Präsens, Präteritum und Perfekt im Englischen und Deutschen*. Munich: Fink.
- Mauranen, A. (2003). The corpus of English as lingua franca in academic settings. *TESOL Quarterly*, 37, 513–527.
- Mauranen, A. (2004). Spoken corpus for an ordinary learner. In J. McH. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 89–105). Amsterdam: John Benjamins.
- McCarthy, M. (1988). Questions of naturalness in language. *ELR Journal (New Series)*, 2, 47–60.
- McEnery, T., & A. Wilson (²2001). *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Meunier, F. (2002). The pedagogical value of native and learner corpora in EFL grammar teaching. In S. Granger, J. Hung, & S. Petch-Tyson (Eds.), *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching* (pp. 119–141). Amsterdam: John Benjamins.
- Meunier, F. (2004). Native corpora, learner corpora and ELT: The winning team? Paper presented at the 25th conference of the International Computer Archive of Modern and Medieval English (ICAME 25), Verona, 19–23 May 2004.
- Meyer, C. F. (2002). *English Corpus Linguistics*. Cambridge: Cambridge University Press.
- Mindt, D. (1971). *Strukturelle Grammatik, Generative Transformationsgrammatik und Englische Schulgrammatik*. Frankfurt: Diesterweg.
- Mindt, D. (1981a). Angewandte Linguistik und Grammatik für den Englischunterricht. In P. Kunsmann & O. Kuhn (Eds.), *Weltsprache Englisch in Forschung und Lehre: Festschrift für Kurt Wächter* (pp. 175–186). Berlin: Schmidt.
- Mindt, D. (1981b). Linguistische Grammatik, didaktische Grammatik und pädagogische Grammatik. *Neusprachliche Mitteilungen*, 34, 28–35.
- Mindt, D. (1982). Linguistik und Didaktik im Fremdsprachenunterricht: Historische Entwicklung und Forschungsperspektiven. In C. Gnutzmann & D. Stark (Eds.), *Grammatikunterricht: Beiträge zur Linguistik und Didaktik des Fremdsprachenunterrichts* (pp. 29–45). Tübingen: Narr.
- Mindt, D. (1985). Das Tempus in einer didaktischen Grammatik am Beispiel des futurischen Zeitbezugs im Englischen. *Anglistik und Englischunterricht*, 27, 171–183.
- Mindt, D. (1987). *Sprache – Grammatik – Unterrichtsgrammatik. Futurischer Zeitbezug im Englischen I*. Frankfurt: Diesterweg.
- Mindt, D. (1991). Syntactic evidence for semantic distinctions in English. In K. Aijmer & B. Altenberg (Eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik* (pp. 182–196). London: Longman.
- Mindt, D. (1992). *Zeitbezug im Englischen. Eine didaktische Grammatik des englischen Futurs*. Tübingen: Narr.
- Mindt, D. (1995). *An Empirical Grammar of the English Verb. Modal Verbs*. Berlin: Cornelsen.
- Mindt, D. (1996). English corpus linguistics and the foreign language teaching syllabus. In J. Thomas & M. Short (Eds.), *Using Corpora for Language Research* (pp. 232–247). London: Longman.

- Mindt, D. (1997a). Complementary distribution, gradience and overlap in corpora and in ELT: Analysing and teaching the progressive. In U. Fries, V. Müller, & P. Schneider (Eds.), *From Ælfric to the New York Times. Studies in English Corpus Linguistics* (pp. 227–237). Amsterdam: Rodopi.
- Mindt, D. (1997b). Corpora and the teaching of English in Germany. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 40–50). London: Longman.
- Mindt, D. (2000). *An Empirical Grammar of the English Verb System*. Berlin: Cornelsen.
- Mindt, D. (2002). A corpus-based grammar for ELT. In B. Kettemann & G. Marko (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 91–104). Amsterdam: Rodopi.
- Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (Ed.). (1993). *Richtlinien und Lehrpläne für das Gymnasium – Sekundarstufe I – in Nordrhein-Westfalen*. Frechen: Ritterbach.
- Mohan, T. (1976). Present Simple and Present Continuous. *Zielsprache Englisch*, 3, 41–42.
- Mossé, F. (1938). *Histoire de la forme périphrastique être + participe présent en Germanique*. Paris: Klincksieck.
- Mukherjee, J. (2004). Corpus linguistics and the native-speaker concept. Paper presented at the 25th conference of the International Computer Archive of Modern and Medieval English (ICAME 25), Verona, 10–23 May 2004.
- Nehls, D. (1974). *Synchron-diachrone Untersuchungen zur Expanded Form im Englischen. Eine struktural-funktionale Analyse*. Munich: Hueber.
- Nehls, D. (1988). Der englische Verbalaspekt. In W.-D. Bald (Ed.), *Kernprobleme der englischen Grammatik. Sprachliche Fakten und ihre Vermittlung* (pp. 192–208). Berlin: Langenscheidt-Longman.
- Nelson, M. (2000). A corpus-based study of business English and business English teaching materials. Unpublished PhD thesis. Available online: <http://www.kielikanava.com/thesis.html> (consulted: 22.08.04)
- Nelson, G., S. Wallis, & B. Aarts (2002). *Exploring Natural Language. Working with the British Component of the International Corpus of English*. Amsterdam: John Benjamins.
- Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In J. McH. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 125–152). Amsterdam: John Benjamins.
- Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam: John Benjamins.
- Niemeyer, J. (1991). Tense and aspect in German grammars of English in the past fifty years. In G. Leitner (Ed.), *English Traditional Grammars. An International Perspective* (pp. 329–347). Amsterdam: John Benjamins.
- Oakes, M. P. (1998). *Statistics for Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Ota, A. (1963). *Tense and Aspect in Present-day American English*. Tokyo: Kenkyusha.
- Owen, C. (1993). Corpus-based grammar and the Heineken effect: Lexico-grammatical description for language learners. *Applied Linguistics*, 14 (2), 167–187.
- Owen, C. (1996). Do concordances require to be consulted? *ELT Journal*, 50 (3), 219–224.
- Palmer, F. R. (Ed.). (1968). *Selected Papers of J. R. Firth 1952–59*. London: Longman.
- Palmer, F. R. (1988). *The English Verb*. London: Longman.
- Palmer, F. R. (1998). Firth and the London School. In J. L. Mey (Ed.), *Concise Encyclopedia of Pragmatics* (pp. 1061–1065). Oxford: Elsevier.
- Parrot, M. (2000). *Grammar for English Language Teachers*. Cambridge: Cambridge University Press.

- Partington, A. (1998). *Patterns and Meanings. Using Corpora for English Language Research and Teaching*. Amsterdam: John Benjamins.
- Partington, A. (2004). "Utterly content in each other's company": Semantic prosody and semantic preference. *IJCL*, 9 (1), 131–156.
- Payne, J. (1995). The COBUILD spoken corpus: Transcription conventions. In G. Leech, G. Myers, & J. Thomas (Eds.), *Spoken English on Computer: Transcription, Mark-up and Application* (pp. 203–207). London: Longman.
- Piepho, H.-E. (1974). *Kommunikative Kompetenz als übergeordnetes Lernziel des Englischunterrichts*. Dornberg-Frickenhofen: Frankonius.
- Piepho, H.-E. (1977). 'Progressive' und 'simple form' im englischen Anfangsunterricht. *Westermanns Pädagogische Beiträge*, 29, 71–74.
- Potter, L. (1999). *CobuildDirect Tutorial for University of Birmingham Distance MA Students*. Unpublished manuscript.
- Potter, S. (²1975). *Changing English*. London: Deutsch.
- Pravec, N. A. (2002). Survey of learner corpora. *ICAME Journal*, 26, 81–114.
- Pürschel, H. (1981). Der Gebrauch der progressiven Form im gesprochenen Englisch. In J. Esser & A. Hübler (Eds.), *Form and Functions. Papers in General English, and Applied Linguistics Presented to Vilém Fried on the occasion of His Sixty-Fifth Birthday* (pp. 83–91). Tübingen: Narr.
- Quirk, R., & S. Greenbaum (1973). *A University Grammar of English*. London: Longman.
- Quirk, R., & S. Greenbaum (²1980). *A University Grammar of English*. London: Longman.
- Quirk, R., & S. Greenbaum (1990). *A Student's Grammar of the English Language*. Harlow: Longman.
- Quirk, R., S. Greenbaum, G. Leech, & J. Svartvik (1972). *A Grammar of Contemporary English*. London: Longman.
- Quirk, R., S. Greenbaum, G. Leech, & J. Svartvik (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Recktenwald, R. P. (1975). *The English Progressive: Semantics and History*. Providence, RI: Brown University.
- Reichenbach, H. (1947). *Elements of Symbolic Logic*. London: MacMillan.
- Renouf, A. (2002). The time dimension in modern English corpus linguistics. In B. Kettemann & G. Marko (Eds.), *Teaching and Learning by Doing Corpus Analysis* (pp. 27–41). Amsterdam: Rodopi.
- Reppen, R. (2001). Review of Monoconc Pro and Wordsmith Tools. *Language Learning and Technology*, 5 (3), 32–36. Available online: <http://llt.msu.edu/vol5num3/review4/default.html> (consulted: 22.08.04).
- Rödel, M. (2003). Die Entwicklung der Verlaufsform im Deutschen. *Muttersprache*, 2, 3–15.
- Römer, U. (2004a). A corpus-driven approach to modal auxiliaries and their didactics. In J. McH. Sinclair (Ed.), *How to Use Corpora in Language Teaching* (pp. 185–199). Amsterdam: John Benjamins.
- Römer, U. (2004b). Comparing real and ideal language learner input. The use of an EFL textbook corpus in corpus linguistics and language teaching. In G. Aston, S. Bernardini, & D. Stewart (Eds.), *Corpora and Language Learners* (pp. 151–168). Amsterdam: John Benjamins.
- Römer, U. (Forthcoming). Looking at *looking*: Functions and contexts of progressives in spoken English and 'school' English. In A. Renouf & A. Kehoe (Eds.), *The Changing Face of Corpus Linguistics*. Amsterdam: Rodopi.

- Rutherford, W. E. (1987). *Second Language Grammar: Learning and Teaching*. London: Longman.
- Rydén, M. (1997). On the panchronic core meaning of the English progressive. In T. Nevalainen & L. Kahlas-Tarkka (Eds.), *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen* (pp. 419–429). Helsinki: Société Néophilologique.
- Sammon, G. (2002). *Exploring English Grammar*. Berlin: Cornelsen.
- Sasse, H.-J. (2001). *Recent Activity in the Theory of Aspect: Accomplishments, Achievements or just Non-progressive State?* Arbeitspapier Nr. 40 NF. Cologne: Institut für Sprachwissenschaft.
- Savignon, S. J. (1983). *Communicative Competence: Theory and Classroom Practice*. Reading, MA: Addison-Wesley Publishing Company.
- Savignon, S. J. (1990). Communicative language teaching: Definitions and directions. In J. E. Alatis (Ed.), *Georgetown University Round Table on Language and Linguistics 1990* (pp. 207–217). Washington, D. C.: Georgetown University Press.
- Scheffer, J. (1975). *The Progressive in English*. Amsterdam: North-Holland Publishing Company.
- Schlüter, N. (2000). The present perfect in British and American English: Selected results of an empirical analysis. In C. Mair & M. Hundt (Eds.), *Corpus Linguistics and Linguistic Theory. Papers from the Twentieth International Conference on English Language Research on Computerized Corpora (ICAME 20), Freiburg im Breisgau 1999* (pp. 313–320). Amsterdam: Rodopi.
- Schlüter, N. (2002a). *Present Perfect. Eine korpuslinguistische Analyse des englischen Perfekts mit Vermittlungsvorschlägen für den Sprachunterricht*. Tübingen: Narr.
- Schlüter, N. (2002b). Temporal specification of the present perfect: A corpus-based study. In P. Peters, P. Collins & A. Smith (Eds.), *New Frontiers in Corpus Research* (pp. 307–315). Amsterdam: Rodopi.
- Schneider, E. (1980). Tempus und Aspekt in österreichischen Lehrbüchern für den Englischunterricht. *Arbeiten aus Anglistik und Amerikanistik*, 5 (2), 201–233.
- Schulze, R. (1985). A selective bibliography of recent writings on the English tense and aspect system. *Anglistik und Englischunterricht*, 27, 11–38.
- Schulze, R. (2003). Review of Longman Grammar of Spoken and Written English. *Zeitschrift für Sprachwissenschaft*, 22 (1), 133–135.
- Schwarz, H. (Ed.). (1997). *English G 2000 A1*. Berlin: Cornelsen.
- Schwarz, H. (Ed.). (1998a). *English G 2000 A2*. Berlin: Cornelsen.
- Schwarz, H. (Ed.). (1998b). *English G 2000. Grammatikheft 1/2*. Berlin: Cornelsen.
- Schwarz, H. (Ed.). (1999a). *English G 2000 A3*. Berlin: Cornelsen.
- Schwarz, H. (Ed.). (1999b). *English G 2000 A4*. Berlin: Cornelsen.
- Schwarz, H. (Ed.). (2000). *English G 2000. Grammatikheft 3/4*. Berlin: Cornelsen.
- Schwarz, H. (Ed.). (2001). *English G 2000 A5*. Berlin: Cornelsen.
- Schwarz, H. (Ed.). (2002). *English G 2000 A6*. Berlin: Cornelsen.
- Scott, M. (1996). *WordSmith Tools*. Oxford: Oxford University Press.
- Scott, M. (1997). *WordSmith Tools Manual*. Oxford: Oxford University Press.
- Seidlhofer, B. (2000). Mind the gap: English as a mother tongue vs. English as a lingua franca. *AILA Review*, 14, 357–385.
- Seidlhofer, B. (2001a). The case for a corpus of English as a lingua franca. In G. Aston & L. Burnard (Eds.), *Corpora in the Description and Teaching of English* (pp. 70–85). Bologna: CLUEB.
- Seidlhofer, B. (2001b). Closing a conceptual gap: The case for a description of English as a lingua franca. *International Journal of Applied Linguistics*, 11 (2), 133–158.

- Seidlhofer, B. (2002). *Habeas corpus and divide et impera*: 'Global English' and applied linguistics. In K. Spelman Miller & P. Thompson (Eds.), *Unity and Diversity in Language Use* (pp. 198–217). London: Continuum.
- Seuren, P. A. M. (1998). *Western Linguistics. An Historical Introduction*. Oxford: Blackwell.
- Sinclair, J. McH. (1966). Beginning the study of lexis. In C. E. Bazell, J. C. Catford, M. A. K. Halliday, & R. H. Robins (Eds.), *In Memory of J. R. Firth* (pp. 410–430). London: Longmans, Green and Co Ltd.
- Sinclair, J. McH. (1984). Naturalness in language. In J. Aarts & W. Meijs (Eds.), *Corpus Linguistics. Recent Developments in the Use of Computer Corpora in English Language Research* (pp. 203–210). Amsterdam: Rodopi.
- Sinclair, J. McH. (1985). Selected issues. In R. Quirk & H. G. Widdowson (Eds.), *English in the World* (pp. 248–254). Cambridge: Cambridge University Press.
- Sinclair, J. McH. (Ed.). (1987a). *Looking Up: An Account of the COBUILD Project in Lexical Computing*. London: Collins ELT.
- Sinclair, J. McH. (1987b). Collocation: A progress report. In R. Steele & T. Threadgold (Eds.), *Language Topics. Essays in Honour of Michael Halliday. Volume II* (pp. 319–331). Amsterdam: John Benjamins.
- Sinclair, J. McH. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. McH. (1993). The Bank of English: A British and international corpus of English. *Zeitschrift für Anglistik und Amerikanistik*, 41 (2), 166–167.
- Sinclair, J. McH. (1994). Trust the text. In M. Coulthard (Ed.), *Advances in Written Text Analysis* (pp. 13–25). London: Routledge. Reprinted in J. McH. Sinclair, 2004. *Trust the Text. Language, Corpus and Discourse* (pp. 9–23), ed. with R. Carter. London: Routledge.
- Sinclair, J. McH. (1996). The search for the units of meaning. *Textus IX*: 75–106. Reprinted in J. McH. Sinclair, 2004. *Trust the Text. Language, Corpus and Discourse* (pp. 24–48), ed. with R. Carter. London: Routledge.
- Sinclair, J. McH. (1997). Corpus evidence in language description. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 27–39). London: Longman.
- Sinclair, J. McH. (1999). The computer, the corpus and the theory of language. *LMS Lingua*, 1, 21–32.
- Sinclair, J. McH. (2000a). Current issues in corpus linguistics. In R. Favretti (Ed.), *Linguistica e informatica. Corpora, multimedialità e percorsi di apprendimento* (pp. 29–38). Rome: Bulzoni Editore. Reprinted in J. McH. Sinclair, 2004. *Trust the Text. Language, Corpus and Discourse* (pp. 185–193), ed. with R. Carter. London: Routledge.
- Sinclair, J. McH. (2000b). Lexical grammar. *Darbai ir dienos*, 24, 191–204. Available online at <http://donelaitis.vdu.lt/publikacijos/sinclair.pdf> (consulted: 22.08.04). Reprinted in J. McH. Sinclair, 2004. *Trust the Text. Language, Corpus and Discourse* (pp. 164–176), ed. with R. Carter. London: Routledge.
- Sinclair, J. McH. (2001). Preface. In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.), *Small Corpus Studies and ELT. Theory and Practice* (pp. vii–xv). Amsterdam: John Benjamins.
- Sinclair, J. McH. (2002). Corpus-driven linguistics. In K. Aijmer (Ed.), *ICAME 2002. The Theory and Use of Corpora. Abstracts of the Conference*. Gothenburg: Gothenburg University. Available online: http://www.eng.gu.se/sv/news/icame2002/ICAME_2002abstract_v2.pdf (consulted: 22.08.04).
- Sinclair, J. McH. (2003). *Reading Concordances*. London: Pearson.
- Sinclair, J. McH. (Ed.). (2004a). *How to Use Corpora in Language Teaching*. Amsterdam: John Benjamins.

- Sinclair, J. McH. (2004b). Introduction. In J. McH. Sinclair (Ed.) *How to Use Corpora in Language Teaching* (pp. 1–10). Amsterdam: John Benjamins.
- Sinclair, J. McH. (2004c). New evidence, new priorities, new attitudes. In J. McH. Sinclair (Ed.) *How to Use Corpora in Language Teaching* (pp. 271–299). Amsterdam: John Benjamins.
- Sinclair, J. McH., & R. M. Coulthard (1975). *Towards an Analysis of Discourse. The English Used by Teachers and Pupils*. London: Oxford University Press.
- Smith, N. (2002). Ever moving on? The Progressive in recent British English. In P. Peters, P. Collins, & A. Smith (Eds.), *New Frontiers in Corpus Research* (pp. 317–330). Amsterdam: Rodopi.
- Smitterberg, E. (2005). *The progressive in 19th century English. A process of integration*. Amsterdam: Rodopi.
- Smitterberg, E., S. Reich, & A. Hahn (2000). The present progressive in political and academic language in the 19th and 20th centuries: A corpus-based investigation. *ICAME Journal*, 24, 99–118.
- Stevens, V. (1995). Concordancing with language learners: Why? When? What? *CAELL Journal*, 6 (2), 2–10. Available online: <http://web.bham.ac.uk/johnstf/stevens.htm> (consulted: 22.08.04).
- Strang, B. M. H. (²1968). *Modern English Structure*. London: Edward Arnold.
- Stubbs, M. (1993). British traditions in text analysis. From Firth to Sinclair. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 1–33). Amsterdam: John Benjamins.
- Stubbs, M. (1996). *Text and Corpus Analysis: Computer-assisted Studies of Language and Culture*. Oxford: Blackwell.
- Stubbs, M. (2000). Society, education and language: The last 2,000 (and the next 20?) years of language teaching. In H. Trappes-Lomax (Ed.), *Change and Continuity in Applied Linguistics* (pp. 15–34). Clevedon: BAAL and Multilingual Matters.
- Stubbs, M. (2001a). Texts, corpora, and problems of interpretation: A response to Widdowson. *Applied Linguistics*, 22 (2), 149–172.
- Stubbs, M. (2001b). *Words and Phrases. Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- Svartvik, J. (1990). *The London-Lund Corpus of Spoken English*. Lund: Lund University Press.
- Svartvik, J., & R. Quirk (Eds.). (1980). *A Corpus of English Conversation*. Lund: CWK Gleerup.
- Sweet, H. (1899). *The Practical Study of Languages: A Guide for Teachers and Learners*. London: Dent & Sons.
- Szałek, M. (1982). Die linguistische und die didaktische Grammatik. *Glottodidactica: An International Journal of Applied Linguistics*, 15, 43–51.
- Tesch, F. (1988). Die Anwendung des chi-quadrat-Tests in der Linguistik. In D. Mindt (Ed.), *EDV in der Angewandten Linguistik. Ziele – Methoden – Ergebnisse* (pp. 54–58). Frankfurt: Diesterweg.
- Tesch, F. (1990). *Die Indefinitpronomina some und any im authentischen englischen Sprachgebrauch und in Lehrwerken: Eine empirische Untersuchung*. Tübingen: Narr.
- Tesch, F. (1999). Verb profiles and the simple present. In F.-W. Neumann & S. Schülting (Eds.), *Anglistentag 1998 Erfurt. Proceedings* (pp. 501–507). Trier: Wissenschaftlicher Verlag Trier.
- Teubert, W. (1999). Corpus linguistics – a partisan view. Corpus linguistics as a theoretical approach. *TELRI Newsletter*, 8, 4–19. Available online: http://solaris3.ids-mannheim.de/ijcl/teubert_cl.html (consulted: 22.08.04).
- Tobin, Y. (1993). *Aspect in the English Verb: Process and Result in Language*. London: Longman.
- Tognini-Bonelli, E. (1996a). *The Role of Corpus Evidence in Linguistic Theory and Description*. PhD Thesis, University of Birmingham.

- Tognini-Bonelli, E. (1996b). *Corpus Theory and Practice*. Birmingham: TWC Monographs.
- Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam: John Benjamins.
- Tribble, C., & G. Jones (1990). *Concordances in the Classroom*. Houston, TX: Athelstan.
- Turner, J. F. (1981). Grading the present progressive. *Die Neueren Sprachen*, 80, 194–202.
- Ungerer, F. (1974). Überlegungen zu einer didaktischen Grammatik für den Englischunterricht. *Der Fremdsprachliche Unterricht*, 8, 21–35.
- Ungerer, F., P. Pasch, P. Lampater, & R. Hellyer-Jones (2001). *Learning English Grundgrammatik. Ausgabe für Gymnasien*. Stuttgart: Klett.
- Virtanen, T. (1997). The progressive in NS and NNS student compositions: Evidence from the international corpus of learner English. In M. Ljung (Ed.), *Corpus-based Studies in English: Papers from the Seventeenth International Conference on English Language Research on Computerized Corpora (ICAME 17)* (pp. 299–309). Amsterdam: Rodopi.
- Weinrich, H. (1971). *Tempus. Besprochene und erzählte Welt*. Stuttgart: Kohlhammer.
- Westergren Axelsson, M., & A. Hahn (2001). The use of the progressive in Swedish and German advanced learner English: A corpus-based study. *ICAME Journal*, 25, 5–30.
- Widdowson, H. G. (1996). Comment: Authenticity and autonomy in ELT. *ELT Journal*, 50 (1), 67–68.
- Widdowson, H. G. (1998). Context, community, and authentic language. *TESOL Quarterly*, 32 (4), 705–716.
- Widdowson, H. G. (2000). On the limitations of linguistics applied. *Applied Linguistics*, 21 (1), 3–25.
- Widdowson, H. G. (2001). Interpretations and correlations: A reply to Stubbs. *Applied Linguistics*, 22 (4), 531–538.
- Williams, C. (2002). *Non-progressive and Progressive Aspect in English*. Fasano: Schena Editore.
- Williams, C. (2004). Pragmatic and cross-cultural considerations in translating verbal constructions in prescriptive legal texts in English and Italian. *TEXTUS*, 17 (1), 217–246.
- Willis, D. (1990). *The Lexical Syllabus. A New Approach to Language Teaching*. London: HarperCollins.
- Willis, D. (2003). *Rules, Patterns, and Words. Grammar and Lexis in English Language Teaching*. Cambridge: Cambridge University Press.
- Wittgenstein, L. (1958). *Philosophical Investigations*. Translated by G. E. M. Anscombe. Oxford: Blackwell.
- Wong, J. (2002). “Applying” conversation analysis in applied linguistics: Evaluating dialogue in English as a second language textbooks. *IRAL*, 40, 37–60.
- Young, R. B. (1971). Evolutionary biology and ideology. *Science Studies*, 1, 177–206.
- Zimmermann, G. (1984). *Erkundungen zur Praxis des Grammatikunterrichts*. Frankfurt: Diesterweg.
- Zydatiś, W. (1976a). *Tempus und Aspekt im Englischunterricht*. Kronberg: Scriptor.
- Zydatiś, W. (1976b). Learning problem expanded form: A performance analysis. *IRAL*, 14, 351–371.
- Zydatiś, W. (1977). Die Bestimmung der Interimsgrammatik von Fremdsprachenlernern mit Hilfe von Fehleranalyse und elicitation procedures – dargestellt am Beispiel der Progressive Form und des Present Perfect. *Die Neueren Sprachen*, 76 (1), 14–37.

Index

A

Aarts 9, 11, 31, 170, 277
Allen 31, 32, 34, 38, 63
annotated 9, 10
annotation 9, 10
annotational 9
aspect 1, 5, 19–28, 30–32, 89, 173, 261
attitude 58, 88, 95, 96, 98–101, 107,
108, 110, 111, 161, 165, 166, 203,
229, 234, 240, 241, 256, 267, 268,
273, 284, 288, 289
authentic 3, 11, 23, 24, 102, 277–280,
290
authenticity 3, 276–278, 280, 287, 289,
290

B

Bald 8, 60, 172, 306
Bank of English (BoE) 5, 41, 45, 46,
53, 302, 303
Barlow 2, 10, 11, 49, 268, 282, 283,
285, 298, 303, 306, 307
Biber 2, 25, 27, 28, 30–32, 34, 38, 65,
68, 69, 73–75, 113, 116, 117, 120,
232, 300, 302
Birmingham Collection of English Text
40, 302
BNCweb 48, 303
Brazil 24, 170, 300
British National Corpus (BNC) 5, 41,
43–45, 49–52

C

chi-square test 60, 63, 183, 203, 304,
305
Chomsky 15, 37, 281, 298
Cobuild 8, 45, 57, 299, 303, 306
colligation 13, 14, 60, 123, 133, 181
colligational 14, 100, 151, 191, 206

collocates 87, 98, 102, 107, 135, 136,
195, 232
collocation 12–14, 34, 60, 70, 75, 79,
85, 86, 102, 103, 127, 136, 146, 149,
157, 164, 181, 204, 233–235,
237–242, 272, 282, 298
collocational 12, 14, 105, 151, 206,
213, 282
collocations 5, 12, 34–36, 51, 70,
77–80, 136, 146, 169, 212, 249, 272
communicative competence 276, 280,
281, 287, 288
communicative grammar 17, 300
communicative utility 283, 290
Comrie 20, 22, 23, 30, 89, 101, 104,
116
Concord 50–52, 179
context of situation 12
contextualist 12, 14
contextualists 11
continuous 1, 58, 86–93, 157, 161,
199–204, 205, 225, 230, 231,
260–262, 266, 273, 284, 288
continuousness 20, 24, 58, 86–88,
93–95, 105, 111, 116, 157–159,
199–205, 225, 260–265
core 14
corpus-based 5, 7–11, 27, 28, 32, 37,
299–301
corpus-driven 3–14, 16, 17, 19, 20, 28,
29, 32, 42, 57, 60, 80, 104, 112, 170,
181, 275–277, 286, 287, 293, 297,
298, 301

D

data-driven learning 306
didactic grammar 14–17, 287, 290,
300

disbelief 58, 95, 96, 99–101, 107, 108,
110, 111, 161, 165, 166, 203, 256
downtoning 29, 34, 88, 99, 107, 165,
304
duration 20, 26, 27, 30, 34, 87, 90, 93,
101, 106

E

ELFA 40

emphasis 58, 88, 95, 96, 99–101, 107,
108, 110, 111, 116, 156, 161, 165,
166, 203, 229, 241, 256, 267, 268,
273, 284, 288, 289, 306

F

Firth 11–13, 38, 60, 277

FLOB 21, 32, 34, 36, 73, 75, 83, 300

framing 23, 30, 58, 95, 96, 104–106,
108–111, 161, 168, 203, 204, 229,
232, 235, 238, 241, 256, 267, 268,
273, 284

Francis 8, 10, 11, 45, 51, 112, 170, 279,
282, 285–287, 293, 299, 300, 305,
306

future time reference 24, 30, 82, 83,
85, 86, 107, 151, 154–157, 198, 199,
224, 241, 258, 272, 288, 289

G

general validity 58, 81, 95–98, 106,
107, 109, 111, 157, 161, 164,
203–205, 229, 241, 256, 266–268,
273, 285, 288, 289, 306

going to future 154, 155, 224, 232,
233, 235, 238, 240

gradual change 58, 95, 96, 101, 102,
108–111, 115, 161, 166, 167, 169,
203, 204, 227, 241, 256, 267, 268,
273, 284, 288, 306

grammatical progression 5, 171, 231,
234, 237, 242

grammaticalization 154, 155, 235, 238

Greenbaum 15, 21, 25, 87, 94, 99, 105

H

habit 26, 29, 34, 90, 93, 95, 104,
108–110, 168, 304

habits 26, 95, 96, 103, 104, 108–111,
161, 167, 203, 204, 227, 256, 268

habitual 23, 89, 93, 103, 104, 227

habituality 26, 58, 167

Hunston 2, 8, 10, 11, 45, 51, 112, 170,
175, 282, 285–287, 293, 299, 302,
305

Hymes 281

I

imperfective 23, 30

imperfectivity 20, 87

incomplete 24

incompleteness 20, 173

incompletion 28, 29, 34, 87, 93, 304

instrumentalists 9

iteration 29, 34, 90, 93, 304

J

Joos 21, 31–34, 63, 64, 116, 301

K

KWIC 48–50, 53

L

learner input 39, 296

Leech 4, 7, 9, 14, 15, 17, 25, 27, 104,
106, 287, 299, 300, 306

Legenhausen 31, 33, 35, 300

lexical grammar 5, 17, 71, 111, 113,
170, 285–287, 290, 299

lexicogrammar 299

limited duration 26, 27, 30, 87, 106

lingua franca 39, 40, 296, 302, 306

LOB 21, 26, 30, 32, 34, 36, 73, 75, 83,
122, 300, 301, 304

London-Lund Corpus 41, 304

Lookup 53, 303

M

Mauranen 39, 40, 270, 279, 280

Mindt 7, 15–17, 28–34, 36, 68, 70,
72–74, 77, 78, 83, 90, 92, 93, 97, 99,
101, 270, 277, 287, 294, 300,
304–306

N

naturalness 279, 280

O

Ota 20, 31–34, 63, 64, 75, 78, 301

P

past time reference 84–86, 107, 108,
110, 151, 196–199, 205, 258, 272,
288, 289

pedagogic corpus 175, 305

pedagogic grammar 14–16

place adverbials 75, 78, 79, 193, 194,
222, 253, 289

politeness 26, 29, 34, 58, 85, 95–99,
107, 110, 122, 126, 161, 164–166,
198, 203–205, 241, 267, 268, 273,
284, 288, 289, 306

present time reference 81–83, 85, 86,
96, 109–111, 151, 154, 196, 199,
205, 222, 257, 258

present/future time reference 82, 86,
151, 156, 157, 199, 224, 258

presentation 5, 16–17, 237, 238,
241–243, 286, 287, 289, 290

Pürschel 31, 33, 35, 101, 231

Q

Quirk 20, 25–31, 38, 90, 93, 98, 116,
154, 301–303

R

repeated 58, 87, 89–93, 157–159,
199–202, 225–227, 230, 241,
260–262, 265, 266, 273, 284, 288,
289

repeatedness 58, 86, 87, 89, 90, 92–95,
111, 157–159, 199–203, 225, 226,
241, 260–262, 273, 284, 306

Richtlinien 38, 306

S

Scheffer 19–21, 31, 33–35, 75, 78, 83,
90, 94, 98, 101, 104, 106, 301

selection 16, 17, 287, 288, 290

semantic preference 13, 14

semantic prosodies 291, 300

semantic prosody 13, 14, 102

shock 58, 95, 96, 99–101, 107, 108,
111, 161, 165, 166, 203, 204, 229,
256, 288, 289, 306

Sinclair 4, 10–14, 37, 38, 40–42, 45,
46, 57, 60, 112, 170, 173, 175, 176,
181, 275, 277, 279, 280, 282, 286,
287, 293, 295–297, 299, 302, 303,
305, 306

slot-and-filler model 13, 81, 286

Smith 21, 22, 31–34, 36, 38, 70, 73, 75,
81, 83, 94, 250

softening 58, 85, 88, 95–99, 101, 107,
110, 126, 161, 164–166, 203, 204,
227, 241, 256, 267, 268, 273, 284,
288, 289, 306

Stubbs 8, 13, 43, 48, 60, 275, 277, 281,
299

susceptibility to change 24, 25, 102

T

temporariness 21, 26, 28, 29, 34, 93,
104, 236, 304

time adverbials 35, 75–77, 80, 97, 99,
102, 105, 146, 192–194, 221,
232–234, 238, 239, 242, 253, 283,
289

Tognini-Bonelli 4, 8–11, 13, 16, 17, 28,
112, 170, 270, 287, 299, 307

typicality 281, 287, 290

U

used language 24, 28, 170, 280, 295,
300

V

Virtanen 31, 33, 35, 36, 78, 90, 232

VOICE 40, 302

W

Widdowson 277, 279–281

Williams 1, 2, 19, 20, 22, 24, 25, 38,
99, 102, 116, 154, 172, 304

WordSmith Tools 49, 179, 303

Z

Zydatiř 116, 172, 173

In the series *Studies in Corpus Linguistics (SCL)* the following titles have been published thus far or are scheduled for publication:

- 19 **TOGNINI-BONELLI, Elena and Gabriella DEL LUNGO CAMICIOTTI (eds.):** Strategies in Academic Discourse. *Expected December 2005*
- 18 **RÖMER, Ute:** Progressives, Patterns, Pedagogy. A corpus-driven approach to English progressive forms, functions, contexts and didactics. 2005. xiv + 327 pp.
- 17 **ASTON, Guy, Silvia BERNARDINI and Dominic STEWART (eds.):** Corpora and Language Learners. 2004. vi, 312 pp.
- 16 **CONNOR, Ulla and Thomas A. UPTON (eds.):** Discourse in the Professions. Perspectives from corpus linguistics. 2004. vi, 334 pp.
- 15 **CRESTI, Emanuela and Massimo MONEGLIA (eds.):** C-ORAL-ROM. Integrated Reference Corpora for Spoken Romance Languages. 2005. xviii, 304 pp. (incl. DVD).
- 14 **NESSELHAUF, Nadja:** Collocations in a Learner Corpus. 2005. xii, 332 pp.
- 13 **LINDQUIST, Hans and Christian MAIR (eds.):** Corpus Approaches to Grammaticalization in English. 2004. xiv, 265 pp.
- 12 **SINCLAIR, John McH. (ed.):** How to Use Corpora in Language Teaching. 2004. viii, 308 pp.
- 11 **BARNBROOK, Geoff:** Defining Language. A local grammar of definition sentences. 2002. xvi, 281 pp.
- 10 **AIJMER, Karin:** English Discourse Particles. Evidence from a corpus. 2002. xvi, 299 pp.
- 9 **REPPEN, Randi, Susan M. FITZMAURICE and Douglas BIBER (eds.):** Using Corpora to Explore Linguistic Variation. 2002. xii, 275 pp.
- 8 **STENSTRÖM, Anna-Brita, Gisle ANDERSEN and Ingrid Kristine HASUND:** Trends in Teenage Talk. Corpus compilation, analysis and findings. 2002. xii, 229 pp.
- 7 **ALTENBERG, Bengt and Sylviane GRANGER (eds.):** Lexis in Contrast. Corpus-based approaches. 2002. x, 339 pp.
- 6 **TOGNINI-BONELLI, Elena:** Corpus Linguistics at Work. 2001. xii, 224 pp.
- 5 **GHADESSY, Mohsen, Alex HENRY and Robert L. ROSEBERRY (eds.):** Small Corpus Studies and ELT. Theory and practice. 2001. xxiv, 420 pp.
- 4 **HUNSTON, Susan and Gill FRANCIS:** Pattern Grammar. A corpus-driven approach to the lexical grammar of English. 2000. xiv, 288 pp.
- 3 **BOTLEY, Simon and Tony McENERY (eds.):** Corpus-based and Computational Approaches to Discourse Anaphora. 2000. vi, 258 pp.
- 2 **PARTINGTON, Alan:** Patterns and Meanings. Using corpora for English language research and teaching. 1998. x, 158 pp.
- 1 **PEARSON, Jennifer:** Terms in Context. 1998. xii, 246 pp.