

Que sais-je?

COLLECTION ENCYCLOPÉDIQUE

fondée par Paul Angoulvent

Derniers titres parus

- | | |
|--|---|
| 50 La Bretagne et les Bretons
(M. LE LANNOU) | 1773 Technique de la sculpture
(J. RUDEL) |
| 51 Espaces euclidiens et hermitiens
(G. CASANOVA) | 1774 La chirurgie des os
et des articulations (J. JUDET) |
| 52 Psychanalyse et littérature
(J. BELLEMIN-NOËL) | 1775 Les robots (P.-J. RICHARD) |
| 53 Racine et la tragédie classique
(A. NIDERST) | 1776 La programmation linéaire
appliquée (R. FAURE) |
| 54 Les institutions chinoises
(M. LESAGE) | 1777 Les indices de prix
(J.-L. BOURSIN) |
| 55 La linguistique appliquée
(C. BOUTON) | 1778 L'artisanat en France
(M. DURAND et J.-P. FRÉMONT) |
| 56 Chronologie internationale
(E. BERG) | 1779 La randonnée pédestre
(M. COTE-COLISSON) |
| 57 L'utopie (J. SERVIER) | 1780 La terminologie : noms et notions
(A. REY) |
| 58 La gynécologie
(M. JAMES et R. BORY) | 1781 La trésorerie de l'entreprise
(A. CHOINEL et G. ROUYER) |
| 59 Les maladies du nourrisson
(L. ROSSANT) | 1782 Le féminisme (A. MICHEL) |
| 60 Le western (C. GONZÁLEZ) | 1783 Les incendies (P. GRAPIN) |
| 61 La folie (R. JACCARD) | 1784 L'écologisme (D. SIMONNET) |
| 62 L'idée et la forme (J. CLARET) | 1785 L'expression orale
(L. BELLENGER) |
| 63 La Cinquième République
(P.-M. DE LA GORCE et
B. MOSCHETTO) | 1786 La population française aux
XVII ^e et XVIII ^e siècles
(J. DUPAQUIER) |
| 64 Géographie de l'Allemagne
fédérale (P. RIQUET) | 1787 Agences et associations de voyages
(R. LANQUAR) |
| 65 La philosophie du langage
(J.-P. RESWEBER) | 1788 Le droit des incapacités
(D. DENIS) |
| 66 L'anesthésie (J. BAUMANN et
J.-M. DESMONTS) | 1789 La médecine préventive
(J. ZOURBAS) |
| 67 Rabelais et la Renaissance
(M. LAZARD) | 1790 La région parisienne
(Ph. PINCHEMEL) |
| 68 Le terrorisme (J. SERVIER) | 1791 Les arts martiaux
(H. COURTINE) |
| 69 Les banques dans le monde
(J. RIVOIRE) | 1792 Le désarmement (Cl. DELMAS) |
| 70 Les projections économiques
d'ensemble (M. DIDIER) | 1793 Le Québec (P. GEORGE) |
| 71 L'Irak (Ph. RONDOT) | 1794 Le notariat français
(J. RIOUFOL et F. RICO) |
| 72 Les conditions du travail
(P. JARDILLIER) | 1795 La systémique (D. DURAND) |

*Que
sais-
je?*

LES LANGUES ROMANES

1562

*Que
sais-je?*

LES LANGUES ROMANES

CHARLES CAMPROUX



PRESSES UNIVERSITAIRES DE FRANCE

Les langues romanes

QUE SAIS-JE ?

Les langues romanes

CHARLES CAMPROUX

Professeur à l'Université de Montpellier III

Deuxième édition mise à jour

16^e mille

puf

DU MÊME AUTEUR

Naissance de la poésie française (en coll. de Pierre DAIX), Paris, Ed. Messidor, 1958.

Etude syntaxique des parlers gévaudanais, Paris, Presses Universitaires de France, 1958.

Essai de géographie linguistique du Gévaudan, Paris, Presses Universitaires de France, 1962.

Histoire de la littérature occitane, Paris, Payot, 1953.

Le Joy d'Amour des troubadours, Montpellier, Causse, 1965.

ISBN 2 13 035916 7

2^e édition : 3^e trimestre 1979

© Presses Universitaires de France, 1974
108, Bd Saint-Germain, 75006 Paris

INTRODUCTION

DÉFINITIONS

I. — Définition de l'expression langue romane

On entend par langues romanes les idiomes qui continuent d'une façon ininterrompue le latin des anciens Romains. On désigne le plus souvent du terme de Romania l'ensemble linguistique du domaine roman. On peut distinguer, grossièrement, dans cet ensemble, dix groupes linguistiques : le roumain, le sarde, l'italien, le rhéto-roman, le français, le franco-provençal, l'occitan, le catalan, l'espagnol, le portugais. Il faut y ajouter, du point de vue historique, le dalmate, aujourd'hui éteint.

L'adjectif « roman », qui se trouve attesté en 1765 dans l'*Encyclopédie*, ne remonte pas, tel qu'il est utilisé dans l'expression « langues romanes », au latin *romanus* qui est continué en français par « romain ». Il a été tiré du substantif « roman » ainsi glosé au XVI^e siècle par Pasquier : « On appelle roman nostre nouveau langage. » Il sert depuis le XIX^e siècle à désigner les langues issues du latin. Il avait d'abord été employé par F. Raynouard pour désigner l'ancien occitan. Raynouard avait, en effet, intitulé un de ses ouvrages : *Grammaire de la langue romane ou langue des troubadours*, signifiant par là qu'il considérait la langue des troubadours comme un avatar du latin de Rome. Par analogie avec

l'expression « langue romane », l'adjectif « roman » fut également utilisé pour désigner le style d'architecture qui précéda le gothique ; on signifiait par là que l'on considérait « l'art roman » comme une forme abâtardie de l'architecture romaine.

Si l'adjectif « roman » ne remonte pas directement au latin *romanus*, la valeur que nous lui connaissons s'explique par l'histoire des significations de ce terme et d'un autre adjectif dérivé de *Roma*, attesté sporadiquement depuis le latin archaïque : *romanicus*.

Aux premiers temps de l'histoire romaine, *romanus* était l'ethnique du toponyme *Roma* qui passe pour être d'origine étrusque. Dès ces premiers temps, à la valeur ethnique de l'adjectif s'ajoutait une double valeur juridique et politique, telle qu'elle apparaît dans les expressions : *civis romanus*, *populus romanus*. Primitivement le *populus romanus* était formé par les *gentes* groupées dans les trente curies des trois tribus. Pour être *civis romanus*, citoyen romain, il fallait nécessairement faire partie d'une *gens*.

Peu à peu le droit de citoyenneté fut accordé à des clients et à des plébéiens d'origine diverse : ils devinrent citoyens romains. Ceux-ci, les *romani*, s'opposaient d'abord aux autres habitants du Latium, les *latini* qui, devenus *socii* (associés) des Romains, n'en entraient pas moins souvent en conflit avec eux. Par la suite, le droit de cité déborda de Rome et fut étendu non seulement aux habitants du Latium mais aussi à diverses cités italiques en deçà du Pô. En 49 av. J.-C. il fut accordé aux cités italiques d'au-delà du fleuve. Finalement, avec la constitution de Caracalla en 212 apr. J.-C. le droit de cité appartint à tous les habitants de l'*Imperium Romanum*.

Dès lors le terme de *romanus* était vide de toute valeur ethnique. Les *Romains* ne s'opposaient plus qu'aux *Barbari*, terme générique pour désigner les populations qui vivaient en dehors des limites de l'*Imperium Romanum* et qui menaçaient les confins d'une façon permanente. *Romanus* n'avait plus qu'une signification juridique et politique. La connotation politique était devenue prédominante dès les IV^e et V^e siècles si l'on en croit les écrivains de ce temps qui expriment souvent la fierté des habitants de l'*Imperium Romanum* d'être des *romani*, quelle qu'ait pu être la diversité de leurs origines ethniques.

Quelle que fût leur ethnie primitive, les habitants de l'*Imperium Romanum* avaient adopté la langue de Rome, tout au moins comme langue officielle et langue de culture. En conséquence, de bonne heure, *romanus*, outre sa signification politique, se chargea d'une signification linguistique. « Le *Romanus*, écrit Gaston Paris, est l'habitant, parlant latin, d'une partie quelconque de l'Empire. » Avec la chute de l'Empire romain, la signification politique s'effaça naturellement et désormais *romanus* ne désigna plus, à partir de la fin du V^e siècle, que celui qui parlait la langue de Rome.

Au dernier siècle de l'Empire apparaît le terme de *Romania*, attesté chez Orose au début du V^e siècle, dont le développement sémantique, tout au moins en Occident, fut parallèle à celui de *Romanus*. La valeur politique du mot est encore attestée au VI^e siècle, mais bientôt *Romania* ne désigna plus que l'ensemble de ceux qui parlaient la langue de Rome, les *romani*. La restauration de l'Empire romain par Charlemagne ne rendit au terme de *Romania* une valeur politique que pour peu de temps.

Les Romains de la Rome des premiers temps

n'appelaient pas leur langue *lingua romana*, mais *lingua latina*; l'adjectif *latinus* étant dérivé du toponyme *Latium*, le pays dont Rome était la capitale. L'expression *lingua latina* continua d'être utilisée tout au long de l'histoire de la Rome antique jusqu'à la chute de l'Empire romain. Elle le fut encore par la suite dans toute la Romania mais alors en tant qu'expression désignant le latin des intellectuels et des savants. D'assez bonne heure l'adverbe *latine* avait pris la signification de « en bon latin », par suite de l'opposition entre *latinus* et *barbarus*.

Au fur et à mesure que se développait la notion de *romanus* évoluant vers un signifié purement linguistique et que naissait la notion de *Romania* évoluant vers un signifié semblable, l'expression normale pour *parler la langue de Rome*, qui était, depuis les origines *latine loqui*, se trouva peu à peu concurrencée par l'expression *romane loqui*. Il y eut un moment où *latine loqui* dut signifier exclusivement *parler en bonne langue* tandis que *romane loqui* signifiait simplement *parler en langue vulgaire*.

Dès le latin archaïque, à côté de l'adjectif *romanus*, on trouve l'adjectif *romanicus* sporadiquement. Caton dans son *De rustica* emploie *romanicus* au sens de « originaire de Rome, » tandis que *romanus* signifiait simplement « romain ». On peut donc estimer que, tandis que *romanus* restait attaché à *Roma*, *romanicus* a appartenu plus spécialement au champ sémantique de *Romania*. L'expression pour « parler la langue des *romani* », seule attestée, est *romane loqui*. L'adverbe *romanice* avec valeur linguistique n'apparaît qu'au milieu du XI^e siècle. Comme d'autre part, dans les textes latins et moyens latins jusqu'au X^e siècle c'est toujours *romanus* et non *romanicus* que l'on rencontre avec la valeur

linguistique, et que dans la latinité tardive beaucoup d'adjectifs ont des dérivés en *-anicus* qui ne diffèrent guère ou pas du tout des adjectifs simples (comme par exemple, dans le cas des adjectifs en *-anus*, *-anicus* qui ont servi en Languedoc à former des noms de lieu sur des noms de personnes), on peut penser que *romanice* a été facilement formé à partir de *romane*.

En tout cas, en ancien français, c'est à partir de l'adverbe *romanice* au sens linguistique de : « à la façon des Romains » par opposition aux Francs comme on peut le voir dans la Loi salique par exemple, que nous avons le substantif et l'adjectif *romanz* employés pour désigner la langue vulgaire aussi bien celle du Nord (pays d'oïl) que celle du Sud (pays d'oc). *Romanz* s'employait dès lors par opposition au latin, devenu la langue savante.

C'est de même à partir de *romanice* que dans les pays rhéto-romans fut utilisé le terme de *rumontsh*, *rumantsh*.

En moyen français, lors de la disparition de la déclinaison à deux cas, qui opposait *romanz* cas sujet à *romant* cas régime, triompha la forme *roman* telle qu'elle est employée par Pasquier au XVI^e siècle. En tant qu'adjectif sur la forme *roman* a été tirée la forme *romane* tandis que dans la Suisse romande on eut, au XVI^e siècle également, *romand*, *romande*.

En ancien français ainsi qu'en ancien occitan *romanz* eut également le sens de « ouvrage écrit en langue vulgaire », après avoir eu, probablement d'abord, celui de « texte vulgaire — français ou occitan — résultat d'une traduction ou d'un remaniement d'un texte latin ». Les verbes *romancier* (français), *romansar* (occitan) ont signifié « traduire en langue vulgaire » et « parler, chanter, écrire,

célébrer en langue vulgaire ». A partir du ^{xiv}^e siècle, le terme de *roman* s'appliquera plus spécialement aux romans d'aventure alors écrits en vers, puis au ^{xv}^e siècle aux récits de chevalerie en prose. Depuis le ^{xviii}^e siècle le mot *roman* et les dérivés de sa famille ont connu la fortune que l'on sait.

Si, en français, *romanus* a abouti à nous donner les deux champs sémantiques que nous connaissons au mot *roman*, celui de sa valeur linguistique et celui de sa valeur littéraire, dans la partie orientale de la *Romania* le terme de *romanus* a connu une destinée différente. Sous la forme *rumân*, il a eu d'abord le sens d'*esclave* en ancien roumain, la domination slave ayant enlevé leurs droits civils aux descendants des *romani*. La forme *Român* (Romîn) est une réfection opérée sous l'influence du latin, et désigne désormais le citoyen de la Roumanie, la *România* forme rendue officielle en 1859 pour désigner la nouvelle formation politique lors de l'union des deux principautés de Valachie et de Moldavie. En dehors du cas de la Roumanie, on retrouve le nom de *Romania* dans la région de l'Italie dénommée la *Romagna*, forme qui continue directement en formation populaire le terme ancien (1).

(1) Dans l'Antiquité et jusqu'à la chute de l'Empire d'Orient le terme de *Romania* désignait en Orient tout l'Empire. Lorsque Byzance était devenue la capitale de l'Empire romain d'Orient le nom de *Rômaïoi* que les Grecs donnaient aux Romains fut appliqué aux Grecs eux-mêmes. Leur langue s'appela la langue *rômaïkê*, d'où, de nos jours, le nom de *româique* pour désigner le grec moderne populaire.

Les termes *latinus*, *latine* sont représentés dans toute la *Romania* comme termes savants. Pour désigner la langue latine l'Italien emploie *il latino*, le français *le latin*, l'occitan *lo latin*, le catalan *el llatí*, formes savantes qui représentent l'adjectif latin neutre substantivé *latinum*, attesté avec cette valeur pour la première fois chez Quintilien (1^{er} siècle de notre ère). L'espagnol *el latín* et le portugais *o latim* tirent leur origine de l'adverbe *latine* (dans l'expression consacrée *latine loqui*). Le roumain a utilisé le type classique *lingua latina* : *limba latină*. Le caractère savant de l'origine de ces

II. — Définition de la place occupée par les langues romanes

Définir les langues romanes par la place qu'elles occupent parmi les langues du monde est une tâche difficile, sinon impossible, si l'on veut les définir sous le triple aspect, habituellement évoqué, géographique, génétique et typologique, ne serait-ce que parce que, sauf pour le premier aspect, la recherche scientifique est bien loin d'être arrivée à son terme.

Définir géographiquement nos langues romanes n'a qu'une simple valeur documentaire, fort intéressante, d'ailleurs, à certains points de vue. Mais on n'aura pas appris grand-chose, en dehors de l'énoncé, quand on aura dit qu'elles s'étendent en Europe, en Amérique du Sud, en Amérique du Nord (Mexique, Canada), et en quelques autres régions de moindre importance dans le monde, sur le territoire des peuples dits latins.

Définir typologiquement les langues romanes suppose que la même définition est déjà réalisée pour toutes les autres langues du monde. On est loin de cette réalité, malgré les résultats récents obtenus

termes ne les a pas empêchés de développer des nuances sémantiques remarquables. D'une part *latin* a pu devenir synonyme de « langage en général » et de « langue spéciale » (d'où langage incompréhensible, jargon), cf. Guillaume de Poitiers : « *e li aucel -chanton, chascun en lor latin* ». D'autre part il s'est employé au sens de « érudition, savoir », cf. : *y perdre son latin*.

Comme terme populaire le continuateur de *latinus* a servi à désigner des populations romanes en contact avec des populations d'origine et de langue différentes, d'une part sur la côte dalmate (où le parler roman est aujourd'hui éteint), d'autre part, dans les Alpes (Dolomites, Engadine, Grisons) où les habitants se dénomment *Ladins* et appellent leur langue *lingua ladina*. Au Moyen Âge, en Espagne l'usager de l'espagnol, par opposition à l'arabophone, fut appelé *ladino*. Ce terme a été conservé par les Séfardites (juifs expulsés d'Espagne) pour se désigner eux-mêmes et leur parler. Pour signifier « parlant espagnol », *ladino* survit encore dans les pays de langue espagnole en Amérique du Sud.

grâce à la faveur dont jouit le structuralisme contemporain. Il est évident qu'une typologie achevée des langues du monde en synchronie permettrait une exacte définition des diverses langues par leur classement typologique même.

Ce classement pourrait, par ailleurs, modifier le classement génétique de la linguistique historique qui s'est efforcée depuis un siècle et demi de définir les langues en les groupant en familles, par le système de la filiation. Cependant le classement des langues par leur parenté, qui exprime uniquement le fait historique de la continuité, quoi qu'il en puisse être de leurs ressemblances ou divergences de structure sur le plan synchronique, reste encore pour le moment, surtout en ce qui concerne les langues indo-européennes, le moyen le plus simple et le plus clair de définir les langues romanes par la place qu'elles occupent parmi les autres langues.

Sans accorder à ce classement une valeur réelle du double point de vue génétique et typologique, on divise généralement les langues du monde en : *a*) langues américaines des indigènes du continent du même nom ; *b*) langues de l'Océanie, notamment le groupe des langues malayo-polynésiennes ; *c*) langues de l'Afrique centrale et australe (groupe nilo-saharien, groupe Niger-cordofanien, groupe Khoïn) ; *d*) langues de l'Asie du Sud-Est (chinois, tibétain, birman, etc.) ; *e*) langues de l'Eurasie et de l'Asie septentrionale (vaste groupe qui va de la Finlande et de la Hongrie jusqu'au Japon) ; *f*) langues caucasiennes (certains veulent y rattacher le basque) ; *g*) langues afro-asiatiques appelées encore « chamito-sémitiques » (certains rattachent l'ibère au groupe libyco-berbère qui fait partie de ces langues) ; *h*) langues asianiques et méditerranéennes (ne comportent que des langues mortes, assez peu connues,

qui furent en usage en Asie Mineure et dans le Bassin méditerranéen, parmi lesquelles l'étrusque et le ligurien) et enfin les langues indo-européennes auxquelles appartiennent les langues romanes.

Les langues indo-européennes sont celles qui nous sont le mieux connues grâce à la méthode de classement génétique. On peut affirmer qu'on est arrivé à classer dans leur groupe des langues apparemment fort différenciées, à en fixer sûrement le degré de parenté et à établir convenablement les correspondances entre les divers sous-groupes qu'elles constituent.

On divise le plus souvent la grande famille indo-européenne en neuf groupes : *a*) le tokharien, disparu aujourd'hui, connu par des textes antérieurs au *x*^e siècle de notre ère, découverts dans le Turkestan chinois ; *b*) l'indo-iranien (sanskrit, vieux perse, avestique et, de nos jours, bengali et hindoustani) ; *c*) hittite seulement connu par des textes du *II*^e millénaire av. J.-C., découverts en Asie Mineure ; *d*) balto-slave (lituanien et letton d'une part, d'autre part vieux slave et, de nos jours, bulgare, serbe, tchèque, slovaque, polonais, russe) (grand-russien, ukrainien), etc. ; *e*) illyrien et albanais ; *f*) germanique (gothique aujourd'hui éteint, nordique représenté actuellement par les langues scandinaves, anglo-saxon d'où vient l'anglais, vieux haut allemand d'où provient l'allemand moderne, vieux bas allemand représenté aujourd'hui par le néerlandais) ; *g*) l'hellénique (grec ancien et ses dialectes, moyen grec langue du monde byzantin jusqu'au *xvi*^e siècle, grec moderne) ; *h*) italo-celtique ou, suivant certains, italique et celtique.

Le celtique nous intéresse particulièrement puisque, outre le brittonique, langue de la Grande-Bretagne dans l'Antiquité, qui subsiste dans le

gallois et sans doute dans le breton, et le gaélique (irlandais, écossais des Hautes-Terres, parler de l'île de Man), il comprenait le gaulois qui s'était répandu de la Gaule à l'Asie Mineure et dont il ne reste plus que des traces éparses, plus nombreuses peut-être qu'on ne le dit si l'on en croit le savant celtisant et toponymiste F. Falc'hun (1).

Quant à l'italique, il comprenait les parlers de l'Italie antique : les dialectes osco-ombriens d'une part et d'autre part le latin, parler du Latium, dialecte usité dans une région qui s'étendait seulement entre la rive gauche du Tibre inférieur, la mer Tyrrhénienne et le pied des Apennins.

Issues du latin parlé à Rome, les langues romanes se définissent donc comme appartenant originellement au sous-groupe italique du groupe italo-celtique au sein de la famille des langues indo-européennes.

(1) Cf. François FALC'HUN, *Les noms de lieux celtiques*, Rennes, Editions armoricaines, 1966.

CHAPITRE PREMIER

DÉVELOPPEMENT HISTORIQUE DES ÉTUDES ROMANES

I. — La préhistoire de la linguistique romane

Au Moyen Age, la seule langue qui était jugée digne d'être étudiée et enseignée était le latin. A telle enseigne que le terme de *grammaire* (la *grammatica* dont l'étude occupait la première place dans l'enseignement du *trivium* qui comprenait grammaire, rhétorique et dialectique) désignait exclusivement la langue latine que l'on enseignait dans les écoles. Uc Faidit commence ainsi son *Donatz Proensals* : « Las oit partz que om troba en *grammatica* troba om en vulgar provençal. » (Les huit parties du discours que l'on trouve dans la *grammaire* on les trouve dans le *vulgaire provençal*.) On le voit : « grammaire » s'opposait à « idiome vulgaire ».

Un des premiers ouvrages à traiter de l'une des langues romanes nées du latin fut précisément l'ouvrage d'Uc Faidit dont on place la rédaction vers 1240. Il avait probablement été précédé des *Razos de trobar* du troubadour catalan Raimon Vidal de Besalu, troubadour de la première moitié du XIII^e siècle, ouvrage qui fut mis en vers entre 1270 et 1280 par Terramagnino de Pisa et inspira les

Regles de trobar ouvrage écrit en Sicile vers 1290 par Jaufre de Foixà sur l'ordre du roi d'Aragon Jacme II.

Ces premières grammaires d'une langue romane furent composées dans une optique essentiellement philologique et rhétorique et non point exactement linguistique. Elles visaient à aider ceux qui désiraient composer dans la langue des troubadours d'une façon correcte et s'adressent particulièrement à ceux qui sont les plus éloignés du centre de la *parladura de Lemozi* (c'est ainsi que Raimon Vidal de Besalu dénommait la langue des troubadours) : Catalans, Italiens, Siciliens. De même, dans l'autre partie du domaine linguistique français, le domaine d'oïl, ce fut, comme dans le domaine d'oc, en vue de servir aux étrangers que fut composée la première grammaire du français : l'*Aprise de la langue française*, œuvre de Gautier ou Walter de Bibbesworth, écrite vers 1290 pour une noble dame de l'Essex. L'ouvrage est, d'ailleurs, rédigé en dialecte anglo-normand.

Les ouvrages ci-dessus visaient donc un but essentiellement pratique et étaient rédigés en vulgaire d'oc et d'oïl. C'est en latin que Dante rédigea son plaidoyer en faveur de l'élévation de son idiome vulgaire à la dignité d'illustre dans son *De vulgari eloquentia* (1315). Il s'agit donc toujours d'une optique non linguistique. Le *De vulgari eloquentia* est, en vérité, une poétique du « vulgaire ». Mais Dante a élargi sa réflexion grammaticale par une méditation critique approfondie sur les langues et principalement sur les idiomes vulgaires romans. À côté des conceptions traditionnelles sur le langage, les langues, la grammaire, telles qu'on les trouve au Moyen Âge chez les philosophes, chez Roger Bacon, par exemple, qui déclare : « Grammatica

una et eadem est secundum substantiam in omnibus linguis, licet accidentaliter varietur », on trouve dans l'ouvrage de Dante certaines affirmations originales remarquables qui ont toujours leur place dans la linguistique moderne. Dante se fait déjà une conception droitement historique des langues qui évoluent, à la fois, dans le temps et dans l'espace. La langue primitivement une des peuples de l'Europe s'est divisée en trois catégories d'idiomes. Dante fait preuve d'une connaissance remarquable, pour ces temps-là, de ce qu'était alors l'Europe linguistique. Le grec à l'Orient, les parlers germaniques à l'Est et au Nord, « quant au reste de l'Europe, écrit-il, une troisième langue le tient tout entier, bien qu'à présent elle semble elle-même « triparlière » : car certains pour affirmer disent *oc*, d'autres *oïl*, d'autres *si*, comme font, par exemple les *Yspani*, les *Franci* et les *Latini* ». Ce disant, Dante définit, pour la première fois, le « domaine roman », du moins le domaine roman occidental. On peut s'étonner de voir Dante identifier ceux qui parlent l'*oc* avec les *Yspani* sans citer, par ailleurs, une quatrième langue qui serait l'ancêtre de l'espagnol. C'est que Dante se fait une idée précise des vulgaires illustres. De ces vulgaires il n'y en a, pour lui, que trois, l'*oc*, l'*oïl* et le *si*. Il n'ignore pas qu'il existe des parlers locaux : il en existe autant qu'il y a de cités relativement importantes, mais ils s'effacent devant le vulgaire illustre. Pour le *si*, par exemple, ce parler vulgaire illustre n'est « aucun des dialectes italiques » ; « nous appelons vulgaire illustre en Italie, cardinal, et royal, et courtois, celui qui appartient à toute ville italienne et n'apparaît le bien propre d'aucune, celui d'après quoi tous les vulgaires municipaux d'Italie se mesurent, pèsent et comparent ». Ainsi le vulgaire illustre est une

koiné. Tel était le caractère de l'*oc* créé par les troubadours et qui, au temps de Dante, était la langue de culture des Yspani peut-être plus que des gens du Midi actuel de la France qui de ce point de vue étaient entraînés dans l'orbite de l'oïl, vulgaire illustre des rois de France, depuis que le Languedoc était devenu domaine royal. On se rend facilement compte de ce fait quand on observe les limites que Dante reconnaît à la langue d'oïl :

« Quant à ceux qui parlent en langue d'oïl, ils sont en quelque manière au septentrion, eu égard à ceux-là (ceux qui parlent en langue d'oc et ceux qui parlent en langue de si). Car ils sont au levant des Alamans ; au septentrion et au couchant ils ont pour retranchement la mer d'Angleterre ou de Galles, et pour bornage les *montagnes d'Aragon*. »

Par ailleurs Dante estime qu'à côté des vulgaires naturels existent toujours les langues grammaticales, et notamment, la *grammatica* dont parlait Uc Faïdit. Cette *grammatica* d'après Dante, comme pour l'ensemble des grammairiens de son temps, dépend essentiellement d'une convention maintenue par l'école. C'est une langue fixée et fixe. De cette langue ne pouvaient donc « dériver » les vulgaires romans. Dante ne dit rien de la langue « naturelle » qui est la source des langues romanes. On peut déduire toutefois de son argumentation que le latin littéraire devait être la forme grammaticale de cette langue mère, cette *troisième langue* européenne qui à présent semble elle-même triparlière.

Pendant plusieurs siècles les préoccupations linguistiques resteront surtout orientées soit vers l'aspect philosophico-logique, soit vers l'aspect grammatico-littéraire du langage et des langues. On étudiera les problèmes que posent la grammaire normative, le bel usage littéraire de la langue, le développement et l'enrichissement des langues na-

tionales, les fondements philosophiques du langage et des langues, les rapports de celles-ci avec la logique et les structures de l'esprit humain. Le problème de l'origine du langage et des langues par lesquelles il se manifeste ne sera la plupart du temps évoqué qu'en fonction de principes qui se rattachent à ces orientations. Toutes positions qui ne pouvaient se rapporter que très indirectement à la connaissance et à l'étude des langues romanes.

Toutefois l'argumentation rhétorique, par exemple, pouvait trouver appui dans la constatation de certaines relations génétiques. Ainsi dès le *xv^e* siècle, suivant plus ou moins la direction suggérée par Dante, un certain nombre de philologues italiens en étaient arrivés à penser que les vulgaires romans, tout spécialement l'italien, étaient d'origine latine. A partir du *xvi^e* siècle surtout, la volonté d'ennobler le vulgaire national ne pouvait que déboucher sur l'affirmation d'une origine noble desdits vulgaires. Les travaux étymologiques qui se multipliaient en France depuis le deuxième tiers du siècle devaient soutenir cette affirmation. Ne disposant pas des matériaux ni des instruments de recherche dont nous disposons, les méthodes de recherche elles-mêmes étant encore à inventer, les étymologistes échafaudèrent des hypothèses fort fantaisistes : l'origine hébraïque du français et de l'italien, l'origine grecque (ou gauloise) du français, l'origine étrusque du toscan, etc. Peu à peu la vérité se faisait jour à travers ces tentatives. Hotman qui mettait cependant en lumière la part de l'allemand dans les origines françaises était déjà tout latin. Mais c'est à H. Estienne, en France, et à C. Cittadini en Italie que l'on doit l'affirmation nette, vers la fin du siècle, que les langues romanes ont leur source dans le latin vulgaire. Désormais, cette

conception devait prévaloir chez les érudits.

Au XVII^e siècle, malgré la prépondérance des préoccupations rhétoriques qui se manifestent, par exemple, en Italie avec le *Vocabolario degli Accademici della Crusca* publié en 1612 à Venise, en France avec la fondation de l'Académie en 1635, la publication des *Remarques sur la langue française* de Vaugelas, et surtout, à partir de 1660 avec la publication de la *Grammaire générale et raisonnée* de Port-Royal, les recherches étymologiques continuées apportèrent aux études romanes des résultats non négligeables. Le représentant le plus connu et le plus habile dans ce domaine fut Gilles Ménage qui publia en 1650 ses *Origines de la langue française* et en 1669 les *Origini della lingua italiana*. Un nombre non négligeable des étymologies de Ménage ont été confirmées par les romanistes modernes. Ménage comprit l'importance de l'étude des textes de la basse latinité, des textes de l'ancien français. Mieux encore il prit conscience de la nécessité d'étudier les dialectes et les parlers des paysans.

Désormais prévaut l'idée que la latinité vulgaire est à la base des langues romanes. Dès lors est reconnu l'intérêt historique des comparaisons linguistiques et l'on découvre certaines lois constantes dans les évolutions et les transformations. C'est au XVII^e siècle que l'on doit l'œuvre, de la plus haute importance, de Charles du Cange le *Glossarium mediae et infimae latinitatis* qui vit le jour à Paris en 1678. L'ancien français a déjà sa place dans ce trésor lexical de la latinité médiévale, résultat du dépouillement d'une très riche récolte de toutes sortes de documents : grammaires, glossaires, chartes, chroniques, livres liturgiques, œuvres hagiographiques et œuvres littéraires, etc.

Au XVIII^e siècle l'attention considérant les faits

linguistiques se porte essentiellement vers la spéculation philosophique. On s'intéresse aux problèmes posés par l'origine du langage dans ses rapports avec la pensée. On se passionne pour une grammaire générale conçue à la fois comme une création et une manifestation de la logique. Bien connus, à ces points de vue, sont les noms de Leibniz, Berkeley, Condillac, Dumarsais, etc. A ces spéculations se rapporte assez directement le problème de la valeur relative des langues. On connaît le *Discours sur l'universalité de la langue française* de Rivarol publié en 1784, tandis que Fichte affirme la supériorité de l'allemand sur le français pendant que Janisch juge le français raffiné et l'allemand philosophique.

Cependant, le préromantisme pousse les érudits dans une direction qui deviendra dominante quelques années plus tard. On s'applique à exhumer le patrimoine littéraire du Moyen Âge. On étudie avec amour les débuts des littératures nationales. La tentative de Jehan de Nostredame au XVI^e siècle publiant ses *Vies des plus célèbres et anciens poètes provençaux* (1575) était restée à peu près sans écho, du moins en France. Désormais le mouvement ne s'arrêtera plus. Si la publication à Madrid du *Diccionario de la Lengua Castellana* (1726-1739) se rattache encore essentiellement à une préoccupation rhétorique (ce dictionnaire est qualifié de *diccionario de autoridades*), les recherches philologico-linguistiques sont orientées vers l'histoire. C'est alors que Lacurne de Sainte-Palaye (qui meurt en 1781) compose son monumental *Dictionnaire historique de l'ancien français* en dix volumes (publié seulement de 1875 à 1882) et recueille plus de quatre mille pièces des troubadours. On peut dire que le XVIII^e siècle finissant ouvre déjà la première

période vraiment scientifique des études romanes : celle de la linguistique romane comme science historique qui caractérisera le XIX^e siècle.

II. — La linguistique romane au dix-neuvième siècle

1. La linguistique romane, science historique : la méthode historico-comparative. — On l'a dit bien des fois : la linguistique romane en tant que science historique, telle qu'elle se présente dans la première moitié du XIX^e siècle, est fille du romantisme. Elle dépend essentiellement d'une conception historico-évolutionniste de la langue. Le sens romantique de l'histoire fait rejeter la conception immobiliste du langage soumis aux seules lois d'une logique générale. Le goût pour les réalités historiques du passé conduit, en même temps, à accorder une importance majeure aux faits de langue en face de la pure spéculation. Ces deux attitudes expliquent la naissance de la méthode historico-comparative.

Le souci d'une connaissance objective des réalités concrètes des langues prévaut déjà chez le philologue allemand Friedrich August Wolf (1759-1824). En 1808 Friedrich Schlegel montrait les relations qui unissent les langues indo-européennes. La preuve définitive de la parenté des langues indo-européennes est établie par Franz Bopp dans sa *Vergleichende Grammatik des Sanskrit, Send, Armenischen, Griechischen, Lateinischen, Litauischen, Altslavischen, Gotischen und Deutschen* (1833-1852). Mais il avait exposé ses vues dès 1816 à peu près en même temps que le Danois Rasmus Rask dont l'œuvre écrite en danois resta peu connue des milieux scientifiques. On peut considérer que c'est la *Deutsche Grammatik* de Jacob Grimm publiée entre 1819

et 1847 qui fonde pratiquement la grammaire historique. Grimm a énoncé le principe des « lois phonétiques », formulé les règles de la mutation consonantique du germanique, auxquelles reste attaché son nom : « loi de Grimm », défini la « métaphonie » ou *Umlaut*, le concept de « fort » et de « faible » dans la flexion verbale, etc. A la même époque (1833-1836) un autre Allemand, A. Friedrich Pott, créait l'étymologie scientifique dans le domaine indo-européen avec ses *Etymologische Forschungen*.

C'est dans ce mouvement que s'insère la science linguistique romane. Les romanistes y prendront une large place parce qu'ils étaient bien placés pour faire progresser la linguistique scientifique par la méthode historico-comparative. En effet, ils n'ignoraient rien de la langue latine qui se trouve à l'origine des langues romanes et ils disposaient de nombreux documents s'étendant à travers le temps qui leur permettaient de suivre l'évolution des langues romanes dans le détail. Par le fait, l'étude des langues romanes pouvait jouer le rôle de « préceptrice » des études linguistiques.

Deux savants méritent d'être regardés comme les fondateurs de la philologie et de la linguistique romanes, François Raynouard (1761-1836) en France et Friedrich Diez (1794-1876) en Allemagne.

Le *Choix de poésies originales des troubadours*, en six volumes, publié à Paris de 1816 à 1821, et *Le lexique roman ou Dictionnaire de la langue des troubadours* également en six volumes, publié à Paris de 1836 à 1844, sont les deux œuvres magistrales de F. Raynouard. A côté de l'abondante récolte des textes de troubadours, le *Choix de poésies* contient un ensemble de *Recherches sur l'origine et la formation de la langue romane*, une *Grammaire de la langue romane ou langue des troubadours* et une

Grammaire comparée des langues de l'Europe latine, dans leurs rapports avec la langue des troubadours qui fait incontestablement de F. Raynouard le fondateur de la grammaire comparée des langues romanes. Sans doute F. Raynouard s'est trompé gravement en considérant l'occitan des troubadours comme la première langue romane issue du latin, comme, pour ainsi dire, une seconde langue mère reliant le latin aux autres langues et idiomes romans. Mais il a eu le mérite, en étudiant un domaine jusqu'à lui peu exploré, celui de l'ancien occitan, langue et littérature, d'avoir donné la première description de la phase ancienne d'une langue romane, d'avoir porté l'attention sur un certain nombre de traits communs aux langues romanes, soulignant ainsi leur parenté, et aussi, il convient de ne pas l'oublier, d'avoir porté son examen sur les idiomes populaires et locaux, ouvrant ainsi un champ d'études qui prendra une importance très grande dans les études romanes aussi bien au ^{xx}e siècle qu'au ^{xix}e.

Un autre mérite de F. Raynouard est d'avoir suscité la vocation de romaniste de Friedrich Diez par l'entremise de Goethe qui fit se connaître les deux chercheurs en 1818 à Jena, même si l'influence de Grimm fut prépondérante sur l'esprit de Diez lorsqu'il se proposa d'appliquer la méthode établie par Grimm, en étudiant les langues germaniques, aux langues romanes. Les œuvres essentielles de Diez, dans le domaine qui nous intéresse, sont sa *Grammatik der romanischen Sprachen*, en trois volumes, publiée à Bonn de 1836 à 1843 et son *Etymologisches Wörterbuch der romanischen Sprachen* également publié à Bonn en 1854. Dans sa *Grammatik* Diez répartit les six langues romanes qu'il examine en un groupe oriental (italien et valaque, c'est-à-dire roumain), un groupe sud-occidental

(espagnol et portugais) et un groupe nord-occidental (occitan et français). Il ne commet pas l'erreur de perspective de Raynouard et fait remonter ces six langues au latin vulgaire. Bien que tenant compte seulement des langues littéraires, il ne commet pas l'erreur de Grimm qui, au début, confondait lettres et sons ; il distingue nettement les unes des autres et fonde son examen sur les sons. Il souligne l'importance du rôle tenu par les changements phonétiques tout en mettant en évidence celle de l'analogie dans la morphologie. Bien que la phonétique et la morphologie de Diez n'aient plus actuellement qu'un intérêt historique, on peut dire que ces deux disciplines en linguistique romane ont été fondées par lui. En ce qui concerne la science de l'étymologie, si la recherche de Diez ne se souciait pas de faire l'histoire des mots, ce qui est une préoccupation de la science étymologique moderne, et s'il se contentait de vouloir remonter à l'origine du mot, Diez n'en a pas moins établi solidement une méthode de travail rigoureuse, appuyée sur la connaissance exacte des changements phonétiques, ce qui lui permit, entre autres choses, de distinguer entre termes savants (*Kunstproducten*) et mots populaires (*Naturproducten*). Par la comparaison des vocabulaires il distinguait les mots communs à l'ensemble des langues examinées et les mots propres à chaque groupe, jetant ainsi les bases d'une lexicologie dialectale de la Romania qui devait se révéler plus tard pleine d'intérêt.

2. La linguistique romane, science positive : les néo-grammairiens. — Pour Diez comme pour les autres linguistes influencés par le romantisme, la langue, bien qu'elle se manifeste au moyen de la matière que sont les sons, est avant tout un produit

de l'esprit humain, même si on peut la considérer comme un organisme naturel. Pour les linguistes de l'époque romantique cet organisme était une création continue de l'esprit humain. Cette conception n'excluait nullement chez eux une attitude positive en face de la réalité du langage. Gaston Paris, disciple de Diez, caractérisait ainsi sa méthode :

« Les faits sont mon seul sujet ; je les rassemble et je les juge aussi bien qu'il m'est possible, voilà tout. »

Vers le milieu du XIX^e siècle, l'attitude positiviste devait devenir prédominante en linguistique, sous l'influence générale des théories positivistes exposées par Auguste Comte, dont on pensait qu'elles étaient confirmées par la théorie biologique de l'évolutionnisme soutenue par Darwin. Sous l'influence de Darwin, A. Schleicher, linguiste allemand qui était également botaniste, inventa la théorie dite de « l'arbre généalogique » (*Stammbaumtheorie*). De même que, suivant l'évolutionnisme darwinien, toutes les variétés d'êtres vivants résultent de la diversification progressive de quelques très rares espèces originelles, de même la diversité linguistique actuelle remonte à un petit nombre de langues mères dont sont génétiquement issues les langues diverses. L'étude du degré de parenté des langues permettait donc d'établir rigoureusement l'arbre généalogique linguistique. La linguistique devenait ainsi une science naturelle. Par là, A. Schleicher en venait à nier absolument que la linguistique pût être une discipline historique ayant pour objet la libre activité de l'esprit humain. Organisme naturel se développant, comme tout autre organisme appartenant aux sciences naturelles, d'une façon mécanique et déterminée, il devait être possible, grâce

à l'observation des faits et à leur systématisation rigoureuse, d'établir l'évolution des langues et d'en déterminer les caractères par des lois fixes, constantes et sans exception. La tâche essentielle de la linguistique c'était donc de promulguer ces lois. L'élément linguistique qui se prêtait le mieux à l'établissement de telles lois était naturellement l'élément phonique, les sons du langage étant définis uniquement par leurs propriétés physiologiques et physiques. C'est pourquoi les linguistes contemporains de A. Schleicher devaient porter leur attention presque uniquement sur la phonétique et voir dans les changements phonétiques l'essentiel de l'évolution linguistique. L'objectif final de l'étude des langues devenait un travail de reconstruction permettant de remonter à l'unité première de la langue mère par la connaissance des lois régissant l'évolution phonétique.

Cette attitude simpliste ne pouvait manquer de soulever des objections. A la *Stammbaumtheorie* Johan Schmidt (1843-1901) opposa la *Wellentheorie*, théorie des ondes linguistiques, innovations linguistiques qui, à l'instar des ondes sur une surface liquide, se répandent sur des aires diverses. L'opposition fut surtout efficace de la part des romanistes disciples de Diez. Gaston Paris écrivait, avec bon sens, en 1868 :

« Le développement du langage n'a pas sa cause en lui-même, mais bien dans l'homme, dans les lois physiologiques et psychologiques de la nature humaine : par là il diffère essentiellement du développement des espèces, qui est le résultat exclusif de la rencontre des conditions essentielles de l'espèce avec les conditions extérieures du milieu. »

Les adversaires de la linguistique positivo-naturaliste ne manquaient pas de souligner les « exceptions » aux lois phonétiques de l'évolution en don-

nant une place importante au rôle de l'« analogie », par laquelle le travail de l'esprit humain est réintroduit dans le mécanisme automatique des « lois phonétiques ».

Dans le dernier quart du XIX^e siècle, un essai de rapprocher ces deux conceptions opposées, qui considéraient respectivement la langue comme un produit purement physique et comme une production de l'activité psychologique, fut tenté par ceux qu'on a appelés les néo-grammairiens. Il s'agit d'un groupe de linguistes allemands, notamment de Leipzig, qui furent dénommés *Junggrammatiker* par analogie avec le mouvement littéraire *das junge Deutschland* de la première moitié du siècle. Les néo-grammairiens s'appuyaient, en théorie, sur les deux principes méthodologiques suivants : 1) le *principe des lois phonétiques* : tout changement phonétique, en tant que tel, obéit mécaniquement à des lois sans exception et donc tous les mots qui se trouvent dans des conditions identiques obéissent aux mêmes règles ; 2) le *principe de l'analogie* : l'analogie joue un grand rôle dans l'évolution des langues par la réfection des formes linguistiques.

En théorie donc, les néo-grammairiens corrigeaient le caractère aveugle des lois phonétiques par les modifications apportées dans le fonctionnement de ces lois grâce à l'influence des formes déjà existantes dans la langue : ils réintroduisaient donc dans le déterminisme matérialiste de la phonétique une certaine liberté psychologique et humaine. Dans la pratique les néo-grammairiens demeurèrent fortement imprégnés d'esprit positiviste : ils ont toujours considéré les « lois phonétiques » comme l'élément prédominant dans la langue et tenu l'analogie, l'aspect psychique, comme un fait occasionnel, accessoire, exceptionnel. L'analogie ne fut pour

eux qu'une sorte de refuge, un dernier recours, une échappatoire lorsque le fonctionnement des lois phonétiques aboutissait à une impasse. L'analogie, suivant l'expression imagée de I. Jordan, ne recueillait que les miettes qui tombaient de la table des lois phonétiques.

La linguistique romane fut naturellement fortement influencée par l'école néo-grammairienne. À cette école, se rattachent, plus ou moins directement, un grand nombre d'ouvrages fondamentaux de la dernière partie du XIX^e siècle et de la première du XX^e : le *Grundriss der romanischen Philologie* de G. Gröber, les six volumes de la *Grammaire historique de la langue française* de K. Nyrop, l'*Etymologisches Wörterbuch der französischen Sprache* d'E. Gamillscheg, les ouvrages d'E. Bourciez : *Éléments de linguistique romane* et *Précis historique de phonétique française*, qui sont toujours des ouvrages de base dans l'enseignement universitaire en France. D'orientatoin néo-grammairienne fut également celui que l'on a appelé « le prince des romanistes », Wilhelm-Meyer-Lübke (1861-1936) à qui l'on doit les quatre volumes de la *Grammatik der romanischen Sprachen* publiés de 1890 à 1902, l'*Einführung in das Studium der romanischen Sprachwissenschaft* (1901), l'*Historische Grammatik der französischen Sprache* (1^{er} volume en 1909, second volume en 1921) et le *Romanisches Etymologisches Wörterbuch* (1^{re} éd. de 1911 à 1920). Meyer-Lübke a intégré à la linguistique romane ce qu'il y avait de meilleur dans la théorie des néo-grammairiens, ce qui ne l'a point empêché de faire preuve d'un certain éclectisme grâce à un esprit toujours ouvert aux idées neuves. Mais, esprit rigoureux, il demeura également toujours attaché aux faits concrets, se méfiant, trop au gré de certains, mais juste-

ment peut-on penser, des discussions purement théoriques.

Dès 1885 une forte opposition s'éleva contre les néo-grammairiens en tant que représentant pratiquement la rigidité mécaniste en linguistique. Georg Curtius s'élève contre l'idée que l'analogie ne serait qu'une tendance pathologique et contre l'attention exclusive apportée au matériel phonique des langues en négligeant le matériel lexical et morphologique. H. Schuchardt nie même l'existence de lois opérant aveuglément dans le langage. En 1897 dans son *Essai de sémantique* Michel Bréal résumera le fond du débat en écrivant :

« Si on avait dit : *développement régulier, marche constante*, personne ne s'en serait soucié. Mais *lois aveugles*, précision astronomique, l'attente générale était mise en éveil. »

Ce disant, il ramenait le rôle de la phonétique du domaine des lois à celui de la tendance.

C'est à la même relativisation des lois phonétiques que devaient aboutir les recherches de la dialectologie qui naît à la fin du XIX^e siècle et qui devait s'épanouir au XX^e avec la géographie linguistique. Celui qu'on a pu appeler le grand initiateur de la dialectologie romane, G. I. Ascoli qui publia en 1873 ses *Saggi ladini*, croyait fermement aux lois phonétiques et ne voyait dans la linguistique autre chose qu'une « histoire naturelle démontrée avec une évidence mathématique ». Malgré cette attitude naturaliste et mécaniste, les travaux de dialectologie d'Ascoli devaient mettre en lumière l'aspect individuel du langage et réintroduire, par exemple, par le concept historique de substrat qu'on lui doit, le facteur humain dans une évolution qu'on avait voulu considérer comme purement organique, donc étrangère à l'histoire.

Ascoli considérait les dialectes comme des unités

délimitables. A cette conception s'opposèrent Paul Meyer (1840-1917) et Gaston Paris en 1888 pour qui « les parlers populaires se perdent les uns dans les autres par des nuances insensibles », ce qui ne pouvait que poser des questions nouvelles sur la manière dont fonctionnaient les lois phonétiques dans le détail. La relativité des lois phonétiques fut encore mise en évidence par les travaux fondés sur l'examen de parlers populaires locaux de P. J. Rousselot en 1891 et L. Gauchat en 1905 qui, tous deux, contestent la réalité du fonctionnement aveugle des lois phonétiques.

A l'orée du XX^e siècle on peut dire que les nombreuses discussions soulevées à propos des néo-grammairiens et le renouvellement des matériaux et des méthodes dû à l'importance nouvelle de la dialectologie, portant son attention davantage sur les faits actuels que sur l'histoire, allaient obliger la linguistique, et tout particulièrement la linguistique romane, à s'avancer sur des chemins nouveaux et variés, sinon divergents.

III. — La linguistique romane au vingtième siècle

La linguistique romane au XX^e siècle semble perdre de son importance dans le courant de la linguistique générale emportée de plus en plus vers la systématisation et le formalisme sous l'influence de diverses tendances qui appartiennent au mouvement général des idées : le renouvellement de l'idéalisme dans le premier tiers du siècle, la *Gestalttheorie*, la psychologie des profondeurs, le néo-empirisme, le behaviorisme et, plus récemment, la théorie mathématique des ensembles. En réalité la linguistique romane continue de mériter son titre

de *praeceptrix linguisticae* suivant le mot de L. Spitzer, non seulement parce que la linguistique générale va le plus souvent chercher ses matériaux dans son domaine, mais aussi parce que la richesse et les possibilités de ce domaine aussi bien en synchronie qu'en diachronie permettent de contrôler et de corriger au besoin les excès de la systématisation par le respect dû aux faits.

La linguistique romane du XIX^e siècle finissant avait sans doute rompu l'équilibre naturel entre esprit et matière par le positivisme excessif de la méthode historico-comparative de type naturaliste. Là contre, le XX^e siècle connaîtra, en schématisant, une triple réaction : celle de l'idéalisme de Vossler, celle du biologiste de la langue Gilliéron, celle du sociologue de Saussure.

1. L'école idéaliste ou esthétique de Karl Vossler.
— Les conceptions de l'école idéaliste reposent sur une orientation psychosociologique, qui s'efforce de mettre en évidence l'influence de l'esprit et de la culture sur la langue ou, par une démarche inverse, l'importance des faits linguistiques pour la connaissance de la mentalité, de la personnalité, du génie d'un peuple. Ces conceptions sont tributaires des doctrines du philosophe italien G. B. Vico, de celles du linguiste allemand W. von Humboldt et elles s'apparentent à celles de l'esthéticien italien Benedetto Croce, elles-mêmes appuyées sur la théorie de l'intuition d'Henri Bergson.

Karl Vossler, considéré comme le chef de l'école idéaliste, a appliqué ses théories, exposées dès 1904, à l'histoire de la langue française littéraire dans un ouvrage paru en 1913, *Frankreichs Kultur im Spiegel seiner Sprachentwicklung* refondu en 1929 sous le titre : *Frankreichs Kultur und Sprache. Geschichte*

der französischen Schriftsprache von den Anfängen bis zur Gegenwart. Un exemple fera saisir l'esprit et la méthode de Vossler : constatant que vers le XIV^e et le XV^e siècle, se développe en France une mentalité commerciale étroite et calculatrice et qu'à la même époque l'usage de l'article partitif, rare en ancien français, se développe largement, il en conclut que l'article partitif a réussi à s'affirmer en français en tant que conséquence de l'esprit matérialiste de l'époque.

De la démarche inverse qui va de la langue à la mentalité, les exemples nous sont donnés par les études de E. Lewy, *Zur Wesengestalt des Französischen* (1922), et de E. Lerch, *Spanische Sprache und Wesenart* (1932), pour ne citer que des travaux se rapportant aux langues romanes. Cette démarche est à son apogée vers 1930 surtout chez les romanistes allemands de la *Wesenskunde* qui, se fondant sur la langue, veut reconstruire le « français éternel » ou l'« homme espagnol typique ».

On a facilement souligné le peu de rigueur de ces méthodes, les généralisations indues qu'elles s'autorisent, ainsi que le fait que les seules langues littéraires ne peuvent permettre de préciser véritablement la corrélation entre l'esprit du peuple et sa langue. Avec plus de méthode et plus de rigueur technique, des principes assez proches de ceux de Vossler ont été mis en pratique par l'école linguistique espagnole avec notamment R. Menéndez Pidal et Amado Alonso. Par ailleurs, l'Italie a donné naissance à la néo-linguistique, qui s'inspire des idées de B. Croce et qui a été fondée vers 1925 par M. Bartoli. A ce mouvement se sont rattachés beaucoup de linguistes italiens, en particulier G. Devoto et G. Nencioni.

L'école idéaliste a eu le mérite d'attirer l'attention

sur la stylistique qui est, pour elle, la linguistique pure, la discipline capitale d'où l'on doit partir pour descendre ensuite à la syntaxe et de là à la morphologie et à la phonétique. C'est un renversement complet de l'attitude des romanistes du siècle précédent. De ce point de vue, on peut penser que découle en partie la stylistique de C. Bally, élève, par ailleurs de F. de Saussure.

2. **Psychologie et sociologie dans la linguistique romane.** — Ce n'est pas seulement dans l'école idéaliste à tendance esthétique que les préoccupations psychologiques ou sociologiques ont joué un grand rôle au cours du développement de la linguistique romane au ^{XX}^e siècle.

Les préoccupations *psychologiques* jouent un rôle important chez un phonéticien tel que Maurice Grammont dont le *Traité de phonétique* est de 1933 : pour lui, dans le domaine de la phonétique générale le psychique intervient aussi bien que le psychologique, comme on peut le voir dans son livre *Le vers français. Ses moyens d'expression. Son harmonie* (1904). Le passage de la pensée à l'expression linguistique intéresse au premier plan la psychologie. A cet intérêt ressortissent : *La pensée et la langue* (1922) de Ferdinand Brunot, *l'Essai de la structure logique de la phrase* (1926) de A. Secheyne, les sept volumes (de 1927 à 1940) de *Des mots à la pensée. Essai de grammaire de la langue française* de J. Damourette et E. Pichon.

Bien que se rattachant au structuralisme issu de l'enseignement de F. de Saussure, la *Psychosystématique* de Gustave Guillaume est étroitement rattachée à la psychologie. Mais G. Guillaume, au lieu d'essayer de saisir la relation entre le linguistique et le mental au niveau de la pensée élaborée, expri-

mée dans le discours, comme le fait la « grammaire psychologique », entreprend de le faire au niveau le plus profond des schèmes dynamiques et des mécanismes psychologiques simples, en s'en tenant le plus souvent aux seuls faits de synchronie. Il pose ainsi le principe d'un structuralisme psychologique. G. Guillaume a publié en 1919 *Le problème de l'article et sa solution dans la langue française*, en 1929 *Temps et verbe* réédité en 1965 avec *l'Architectonique du temps* dans les langues classiques paru en 1945. Bien que fort discutée à l'étranger, la psychosystématique ou psychomécanique de G. Guillaume est certainement une des plus intéressantes tentatives structuralistes modernes. Les disciples de G. Guillaume sont nombreux en France : G. Moignet, B. Potier, Stefanini, etc. L'Université Laval, à Québec, publie actuellement à l'initiative de Roch Valin l'ensemble des *Leçons de linguistique de G. Guillaume*.

Aux préoccupations psychologiques, l'école de linguistique française formée essentiellement d'Antoine Meillet (1866-1936) et de Joseph Vendryes (1875-1960) ajoute des préoccupations *sociologiques*. La même orientation sociologique se retrouve dans le monument que constitue *l'Histoire de la langue française des origines à 1900* commencée par F. Brunot et continuée par Charles Bruneau. En lexicologie la conception de G. Matoré (*La méthode en lexicologie. Domaine français*, 1953) est purement sociologique. Quant à la linguistique marxiste, elle ne peut que faire droit aux données sociologiques. Son principal représentant en France est Marcel Cohen (*Pour une sociologie du langage*, 1956). On notera que l'influence du linguiste officiel de la linguistique soviétique Marr, finalement désavoué en 1950 par Staline, a été pratiquement nulle sur la linguistique romane.

3. La géographie linguistique. — Si les diverses écoles qui obéissent plus ou moins à des préoccupations psychologiques ou sociologiques ont détourné la linguistique des principes d'explication mécanistes et rigides au profit de conceptions plus humaines, à la fois plus souples et plus riches, de la réalité de la langue, la géographie linguistique, elle, cependant d'esprit fort positif, a introduit, essentiellement, le mouvement et la vie dans la linguistique romane. La vie, pour la géographie linguistique, consiste aussi bien dans l'activité de l'esprit humain (qui se manifeste, par exemple, dans la « pathologie » verbale) que dans les imbrications sociologiques (comme le montrent la mouvance des faits onomasiologiques et l'importance des *realia*) ; aussi bien dans la multiplicité des faits (l'application de la quatrième règle de Descartes, le dénombrement de faits vrais est à la base de sa méthode) en synchronie que dans leur organisation structurale, en synchronie encore à travers l'espace, et en diachronie grâce à la géologie et à la stratigraphie linguistiques.

Le fondateur de la géographie linguistique fut le linguiste suisse Jules Gilliéron (1854-1926) promoteur et coauteur avec E. Edmont de l'*Atlas linguistique de la France*, le premier de ces nombreux atlas linguistiques dont la publication continue de nos jours, des atlas régionaux de France mis en train par Albert Dauzat, aux atlas qui couvrent toute la Romania, en Espagne, en Suisse, en Wallonie, en Italie et en Roumanie, apportant des matériaux nouveaux et précieux pour une recherche qui peut ainsi se développer dans des directions aussi différentes que l'ethnologie ou la structuration du signifié pour ne prendre que deux exemples parmi bien d'autres.

L'originalité de la géographie linguistique ne réside aucunement dans une interprétation des faits linguistiques en fonction d'un substrat idéologique quelconque mais bien dans une méthode nouvelle de récolte des faits et de leur observation. En faisant entrer l'espace dans la linguistique, par la méthode des atlas linguistiques, elle bouleverse le statisme des théories linguistiques d'une part, et élargit, d'autre part, les bases des constructions structurales. Les travaux de Gilliéron ont prouvé la rigidité et le simplisme des schèmes néo-grammairiens : la notion de lois phonétiques cède le pas à celle de tendances phonétiques ; l'étymologie cesse de dépendre exclusivement de la phonétique et donne sa place à l'histoire du mot ; les changements linguistiques apparaissent conditionnés par de multiples facteurs : pathologie verbale, contaminations, télescopages, étymologie populaire, homonymie, paronymie, homophonie, hypertrophie sémantique et polysémie. La dialectologie inaugurée à la fin du siècle précédent prend ses véritables dimensions avec l'intégration de la mouvance spatiale : les notions de faisceaux d'isoglosses, de sommets et de murs de traits linguistiques permettent de résoudre la querelle des limites dialectales ; l'histoire ainsi que la structure des langues littéraires se trouvent considérablement éclaircies par la lumière projetée par l'examen des faits linguistiques des parlers populaires. La méthode des *Wörter und Sachen* est issue de la géographie linguistique : elle renouvelle l'étude des mots par l'attention accordée aux *realia*. La méthode des atlas linguistiques entraîne la science de l'onomasiologie qui étudie les diverses dénominations de l'objet : plante, animal, concept, laquelle par un retournement de la démarche appelle la science de la sémasiologie capable de donner

l'élan à une sémantique élargie et de préciser l'étude du signifié en permettant un examen structuré grâce aux données de l'espace linguistique ; récemment les essais d'un P. Guiraud pour établir des structures lexicologiques auraient été impossibles sans les données et les acquis de la géographie linguistique.

Il serait fastidieux de citer tous les noms des savants qui, plus ou moins, ont illustré l'un ou l'autre des aspects de la géographie linguistique ou des problèmes issus de sa méthode : leur grand nombre montre bien que la géographie linguistique a été la plus importante nouveauté du demi-siècle actuel, dans le domaine roman tout au moins. Contrairement à ce qu'ont pu penser certains linguistes « structuralistes », la géographie linguistique et les sciences annexes qui en découlent sont loin d'avoir dit leur dernier mot, ne serait-ce que par ce qu'elles peuvent apporter précisément au structuralisme actuel, soit en lui fournissant des matériaux sans cesse plus nombreux et plus précis pour ses constructions, soit en lui servant de garde-fou contre l'excès de formalisation et d'abstraction qui le menace, par définition, pourrait-on dire.

4. Le structuralisme ou les structuralismes dans la linguistique romane. — La géographie linguistique, essentiellement illustrée par des travaux portant sur les parlers romans a, sinon théoriquement, tout au moins pratiquement, introduit dans la linguistique un certain nombre de notions. Celle de l'importance et de la nécessité du *corpus* visant à une certaine totalité, avec la constitution des atlas linguistiques. Celle de l'importance de l'examen synchronique par l'interprétation des faits synchroniquement recueillis. Celle de la différence entre la langue et la parole grâce à l'étude simultanée des

rapports entre la langue produit social commun et les diverses réalisations locales et individuelles enregistrées par les atlas. Celle de l'importance des rapports du signifiant et du signifié examinés à travers la mouvance géographique grâce aux méthodes onomasiologique et sémasiologique des enquêtes de géographie linguistique. Celle de l'idée que la langue est, avant tout, une pure forme, en étant, par le fait, obligée d'étudier sans cesse la combinaison dans le signe linguistique de la substance phonique et de la substance psychique : ce qui est le cas toujours dans l'examen de la mouvance géographique en morphologie. Celle enfin de structure et d'immanence puisque, c'est en s'efforçant de dégager, du corpus géographique amassé, les lois immanentes à telle ou telle structure, que la géographie linguistique a renouvelé la dialectologie.

Ce n'est donc pas nécessairement dans la *Gestalttheorie* de M. Wertheimer ni même dans la sociologie positiviste d'Emile Durkheim qu'il faut aller chercher des rapports, ou simplement des analogies, pour expliquer l'œuvre de F. de Saussure à partir duquel on fait à bon droit s'ébranler le mouvement *structuraliste*. F. de Saussure, en s'efforçant de dégager l'objet propre de la linguistique, en le séparant de tout ce qui n'est pas lui, a réussi à donner une plus grande rigueur aux grands principes de la distinction entre diachronie et synchronie, de la distinction entre langue et parole, du caractère arbitraire du signe linguistique, combinaison d'un signifiant et d'un signifié, du caractère de la langue, pure forme dont les signes résident essentiellement dans la combinaison de ces deux éléments et fonctionnent les uns par rapport aux autres dans un système qui leur est propre ; enfin, de la notion de

système précisément, système qui, en tant que tel, a ses lois propres et immanentes et forme par là même une structure.

Le *Cours de linguistique générale* de F. de Saussure fut publié seulement après sa mort en 1916 par ses disciples C. Bally et A. Sechehaye, et c'est seulement dans le second tiers du ^{xx}e siècle que les théories de Saussure connurent un grand succès. L'école dite de Genève, avec H. Frei, C. Bally, A. Sechehaye, en publiant les *Cahiers* de F. de Saussure s'est efforcée de développer l'enseignement du maître. Cette école considérait comme absolument anti-thétiques une linguistique synchronique structurale et la linguistique diachronique du siècle passé présentée comme entièrement atomistique.

Dans le domaine des langues romanes de nombreux ouvrages s'accordent à étudier la langue sous l'aspect de la seule synchronie notamment en syntaxe et en stylistique. En syntaxe, on citera : K. Sandfeld, *Syntaxe du français contemporain* (1928, 1936, 1943) ; G. Gougenheim, *Système grammatical de la langue française* (1939) ; C. de Boer, *La syntaxe du français contemporain* (1947) ; W. von Wartburg, P. Zumthor, *Précis de syntaxe du français contemporain* ; L. Tesnière, *Eléments de syntaxe structurale* (1959) pour le français ; M. Criado de Val, *Sintaxis del verbo español moderno* pour l'espagnol ; N. Drăganu pour le roumain ; pour les dialectes romans : L. Remacle, *Syntaxe du parler wallon de la Gleize* (1952 et 1956) ; C. Camproux, *Etude syntaxique des parlers gévaudanais* (1958). En stylistique, C. Bally, élève de F. de Saussure, donne son *Traité de stylistique française* (1919-1921) ; à côté de lui il faut citer, pour le français, F. Strohmeier, M. Cressot, J. Marouzeau, P. Guiraud ; pour l'italien, R. Roedel, G. Devoto, B. Migliorini et F. Chiap-

PELLI ; pour le portugais, M. Rodriguez Lapa, et pour le roumain, I. Iordan.

Si nous citons tous ces noms auxquels il faut ajouter ceux de G. Guillaume et de ses disciples de l'école psychosystématique dont il a déjà été parlé, c'est pour souligner que les études romanes ne sont point restées en dehors du mouvement linguistique qui met en avant les notions de synchronie et de structure, après avoir été à la tête de la linguistique du siècle passé. Mais il est un domaine où, plus que dans celui de la stylistique ou de la syntaxe, l'opposition fut nette, sinon parfois violente, entre l'ancienne linguistique romane et la nouvelle linguistique structuraliste, c'est celui de la phonétique évincée par la *phonologie*. On reprocha sévèrement à la phonétique historique traditionnelle, abondamment représentée dans les études de linguistique romane, ce que l'on appela son « atomisme », autrement dit l'étude des sons en tant qu'éléments isolés. On lui opposa la science nouvelle que l'on appela « phonologie » qui se définit comme une phonétique fonctionnelle et structurale : à l'étude des particularités physiologiques et acoustiques des sons, objet de la phonétique expérimentale, on substitua l'étude de la fonction distinctive des sons, appelés « phonèmes », en mettant en évidence les oppositions, les relations différentielles qui existent entre eux et leur permettent ainsi de distinguer les mots de la langue. Cela naturellement en synchronie.

Telle était, en tout cas, la position des fondateurs de la phonologie avec F. de Saussure et l'école de Genève. Mais, dès 1926, l'école de Prague, avec Nicolaï Troubetzkoï (mort en 1938) et Roman Jakobson, fit admettre la conviction que l'histoire d'une langue peut être étudiée de façon structurale si l'on compare entre elles les coupes successives

synchroniques. Tout en mettant l'accent (surtout des années 30 à 50) sur le principe que la langue est un système où presque tout se tient, la linguistique structurale a porté ses efforts vers une phonologie diachronique en plaçant au centre de son examen les changements advenus dans le système d'une langue. Cette phase descriptive est actuellement dépassée grâce à R. Jakobson et surtout à A. Martinet : la phonologie diachronique ne se limite plus à enregistrer les changements, mais elle s'efforce d'en trouver les raisons.

La linguistique romane a fortement contribué au développement de la phonologie et notamment de la phonologie diachronique. Pour les dialectes romans, le mouvement a été lancé par A. Martinet avec sa *Description phonologique du parler franco-provençal d'Hauteville* (1939). Pour le français A. Martinet également donne le départ avec ses *Remarques sur le système phonétique du français* (1933), immédiatement suivi par les études de Gougenheim puis par celles de B. Malmberg, de Haudricourt-Juilland, etc. La phonologie espagnole est illustrée par T. Navarro Tomas, E. Alarcos Llorach, la portugaise par H. Lüdtke. Pour le roumain on a les travaux de Puscariu et de Petrovici, pour le rhéto-roman ceux de Bender, Francescato, Salzmann, pour le catalan ceux de Alarcos Llorach, pour l'occitan ceux de R. Lafont. Pour l'italien il faut surtout citer la *Fonologia generale e fonologia della lingua italiana* (1969) de Žarko Muljačić.

Une autre école structuraliste, celle du Cercle linguistique de Copenhague qui a élaboré les principes de la *glossématique*, formulés en 1943 par L. Hjelmslev, son principal représentant, concerne assez peu la linguistique romane. K. Togeby a appliqué les principes de la glossématique dans son

ouvrage : *Structure immanente de la langue française* (1951), où la théorie nuit parfois à la réalité. A E. Alarcos Llorach on doit, d'autre part, sa *Gramática estructural. Segon la escuela de Copenhague y con especial atención a la lengua española* (1951).

Les structuralistes américains que l'on range sous l'étiquette « école américaine » pour leurs traits communs dont les plus importants sont de promouvoir une linguistique strictement et uniquement descriptive (ce qui était rendu nécessaire par leur objet premier : les langues amérindiennes) et de procéder à une description purement objective des langues étudiées ont peu apporté à la linguistique romane. Les fondateurs de la linguistique américaine ont été E. Sapir (mort en 1939) et L. Bloomfield (mort en 1941). S'éloignant de l'orientation psychologique de ses débuts, Bloomfield en est venu à créer une linguistique purement mécaniste. L'école dite de Yale en est la représentante. Un de ses membres R. A. Hall Jr a publié en 1948 une esquisse structurale du français (*French*) où il est parfois difficile à un Français de reconnaître sa langue. Ces dernières années, les théories de N. Chomsky ont connu un grand développement parmi les linguistes, romanistes compris. Elles concernent surtout des spéculations théoriques sur le langage que l'on essaie d'appliquer à la description des langues particulières.

Conclusion. — Le xx^e siècle a accordé de plus en plus à la linguistique une importance telle qu'elle est devenue une des grandes sciences de la vie moderne. Il faut reconnaître que la conception structuraliste nouvelle est, pour une très grande part, à l'origine de cet élargissement. Le rôle de la linguistique romane n'en a pas été pour autant diminué grâce notamment à la richesse du matériau

que les études romanes apportaient et ont continué de fournir aux linguistes. Ces études ont fortement contribué au développement de la linguistique du siècle soit par les résistances que sa masse opposait aux conceptions nouvelles, soit, au contraire, par les apports acquis qu'elle intégrait en elles.

C'est très probablement à l'importance de la linguistique romane que l'on doit l'attitude qui tend à prédominer de nos jours parmi les linguistes : s'efforcer d'accorder la systématisation avec le respect des faits. C'est cette attitude qui explique, par exemple, la création de la phonologie historique dont Jakobson énonçait les principes en 1931. C'est elle qui inspire E. Coseriu dans son ouvrage *Sincronia, diacronia e historia* (1958). C'est elle que l'on retrouve dans les *Eléments de linguistique générale* (2^e éd., 1961) d'A. Martinet. C'est la même attitude qui inspire l'un des maîtres actuels de la linguistique romane W. von Wartburg, à qui l'on doit, entre autres choses, la publication du monumental *Französisches Etymologisches Wörterbuch*, quand il écrit dans *Problèmes et méthodes de la linguistique* (2^e éd., 1963) en attribuant à J. Gilliéron le mérite d'avoir préparé les voies à l'étude de « la structure globale d'une langue dans son processus de transformation progressive » : « De cette manière la science du langage deviendra, dans une nouvelle phase de son évolution, une véritable histoire structurale. » Et l'on ne saurait mieux conclure ici qu'en citant encore W. von Wartburg :

« Comme dans les premières décennies du XIX^e siècle, aujourd'hui la linguistique se trouve de nouveau au début d'un grandiose et nouveau développement..., son avenir réside dans la combinaison de la nouvelle direction de recherche structurale-fonctionnelle avec les perspectives et les connaissances qui résultent de la recherche historique. »

CHAPITRE II

ORIGINE DES LANGUES ROMANES

On a dit, dans l'introduction de cet ouvrage, que l'on entend par langues romanes les idiomes qui continuent d'une façon ininterrompue le latin des anciens Romains. Se pose aussitôt à l'esprit une question : à quoi est donc due la diversité des langues romanes ? La question se pose d'autant plus nécessairement que l'Empire romain où devaient naître et se développer les langues romanes fut partout romanisé par les soldats des légions romaines et les vétérans introduisant avec eux le latin comme langue officielle, langue de l'administration, de l'école, du commerce. Plus tard suivra l'emploi du latin comme langue de la liturgie chrétienne.

La romanisation de l'Empire romain s'étant pratiquement effectuée partout suivant le même processus, on a pensé que l'origine des diverses langues romanes devait résider dans la chronologie de cette romanisation. Du début de la conquête politique de l'Italie vers 500 av. J.-C. jusqu'à la conquête de la Dacie tentée à partir de 85 apr. J.-C. et effectuée sous Trajan seulement en 105-106 apr. J.-C., on a estimé que le latin importé à des dates diverses en divers territoires était lui-même divers, d'où la diversité des langues romanes. Il se peut que le latin importé en Sardaigne vers 238 av. J.-C. ait été quelque peu différent du latin importé en Gaule du Nord en 50 av. J.-C. ou de celui apporté en Dacie en 105-106 apr. J.-C., mais pour que la thèse en question soutenue par G. Grö-

ber fût recevable sans plus, il faudrait supposer qu'après chacune de ces importations les liens aient été rompus entre le pays nouvellement romanisé et Rome même. Contre cette thèse on a fait valoir justement que le contact avec Rome n'avait jamais été rompu ; que, par ailleurs, la romanisation des pays intéressés s'était le plus souvent étendue sur une durée qui dépassait largement une zone chronologique réduite à l'époque de l'arrivée des Romains dans un nouveau pays (la romanisation de l'Espagne a duré, par exemple, environ deux siècles) ; et qu'enfin l'évolution du latin pendant les siècles de romanisation a été peu considérable. Si bien qu'il est admis que, si la thèse de Gröber peut expliquer certaines différences peu importantes, elle ne peut certes pas donner la clé de la diversité et de l'origine des langues romanes.

I. — Le latin vulgaire

Les véritables raisons de l'origine des diverses langues romanes doivent donc être cherchées ailleurs. On en voit la raison centrale, sinon essentielle, dans le latin vulgaire en opposition au latin classique.

Les anciens ont souvent fait allusion à l'existence de ces deux latins. Ils les opposent l'un à l'autre en appelant celui-ci *sermo eruditus*, *perpolitus*, *urbanus*, celui-là *sermo cotidianus*, *inconditus*, *militaris*, *plebeius*, *proletarius*, *rusticus*, *usualis*, *vulgaris*. Cette opposition d'épithètes laisse facilement entendre que ce que nous appelons latin classique était le latin littéraire et que l'abondance des autres épithètes désigne le latin parlé. Bien que le latin littéraire influençât sans cesse le latin parlé grâce à son prestige social, administratif, esthétique, c'est naturellement le latin parlé que la romanisation

introduisit dans les divers territoires de l'Empire romain.

Il est nécessaire d'écarter une idée fausse qui considérerait le latin vulgaire comme une sorte de latin décadent formant une couche linguistique plus récente que le latin dit classique. L'idée que le latin vulgaire se serait constitué, par exemple, par un écart progressif d'un latin primitif dit classique, à la suite d'innovations qui se seraient manifestées, par exemple, au III^e siècle apr. J.-C., n'a pas résisté à l'examen de la critique linguistique. Par exemple, un des traits caractéristiques du latin vulgaire qui a eu une importance considérable pour le vocalisme des langues romanes : la fusion de *i* bref et de *e* long, de *u* bref et de *o* long, se trouve déjà en syllabe atone chez Plaute et en syllabe tonique sur des inscriptions de la première moitié du II^e siècle av. J.-C. pour le changement du *i* en *e* ; quant à celui du *u* en *o* il est attesté en syllabe atone à partir de la seconde moitié du I^{er} siècle av. J.-C., et en syllabe tonique vers le milieu du I^{er} siècle apr. J.-C. Bien d'autres traits du latin vulgaire se retrouvent dans le latin archaïque : la syntaxe de Plaute, par exemple, offre bien des exemples de tours absolument rejetés par le latin classique qui se retrouvent dans le latin vulgaire et dans les langues romanes. Dans les faits il s'est passé ceci. Lorsque Ennius et les écrivains de son temps eurent créé le latin littéraire, la distinction s'est établie entre *sermo urbanus* et *sermo plebeius*. Le premier, langue de Rome, grâce au prestige de la capitale est devenu le signe de l'*urbanitas* (le latin de l'*Urbs*). Mais dans la langue parlée de la campagne et du peuple qui en était issu, les archaïsmes et les dialectalismes continuèrent de vivre d'une façon plus ou moins obscure formant l'usage que l'on appela la *rusticitas*, usage

jugé incorrect, ce qui ne l'empêchait pas d'apparaître dans certaines circonstances de la parole. C'est ainsi, par exemple, que Cicéron n'hésite pas à employer la forme *oricula* pour *auricula* dans ses *Epistulae ad familiares*, cette forme due sans doute au parler « rustique » des nourrices et servantes, étant sentie alors comme plus expressive entre familiers.

Si donc le latin vulgaire n'était autre que le latin parlé et non point un état intermédiaire entre latin classique et langues romanes, état qui expliquerait la successive dégradation du latin classique en langues romanes, comment concevoir que ce latin parlé se soit par la suite effrité en idiomes divers ? Une des raisons essentielles est, sans doute, que ce latin parlé tout en étant un n'était pas cependant homogène. L'idée d'un latin vulgaire homogène, idée reçue par les néo-grammairiens, n'est en réalité qu'une création de l'esprit. H. F. Muller a soutenu l'existence d'un latin vulgaire uniforme dont l'unité aurait été maintenue par les relations interprovinciales, par le développement du commerce, par la communauté de religion créée par le christianisme maintenant l'unité après la chute de l'Empire et les invasions germaniques. C'est seulement sous le règne de Charlemagne qu'auraient cessé d'agir les forces unificatrices et que, à la suite de la réforme linguistique carolingienne pour rétablir le latin classique, l'abîme se serait creusé entre la langue écrite et la langue populaire. Muller faisait reposer linguistiquement son opinion sur les diplômes mérovingiens. A cette thèse on a opposé le fait que la langue des diplômes mérovingiens est une langue aulique artificielle, que l'époque mérovingienne, loin d'avoir favorisé les relations interprovinciales et le commerce, fut au contraire une époque de décadence culturelle et de progression barbare, que

le rôle linguistique du christianisme est à double effet (unificateur pour la langue des doctes, diversificateur pour le langage des simples fidèles) et surtout qu'il faudrait supposer une sorte de coup de baguette magique pour faire surgir dans le bref temps d'une ou deux générations tant de parlers divers d'un latin parlé unique.

En réalité le latin vulgaire ou parlé n'a jamais été homogène d'une façon absolue malgré son unité d'ensemble. Il fut très certainement différencié à la fois verticalement et horizontalement. Verticalement le latin parlé variait avec les couches sociales : le latin familial du patricien n'était point celui du paysan ou celui du gladiateur. Le patricien ne parlait pas le même latin dans les débats du sénat et dans le cercle familial. Nous savons que la maîtresse de Catulle, *Claudia*, était appelée *Clodia* dans l'intimité. Horizontalement il existait de nombreux provincialismes : si certains de ces provincialismes sont restés confinés dans leur territoire particulier, d'autres ont conquis droit de cité à Rome même. Enfin, suivant l'époque et le lieu, des préférences ont certainement existé pour choisir entre le matériel linguistique qu'offrait le latin parlé et privilégier ici ou là une partie de ce matériel. C'est ainsi, par exemple, que, à côté de l'emploi en latin élevé de *edere*, de *pulcher* et du comparatif synthétique, un latin moins distingué usait de *comedere*, *formosus* et du comparatif analytique avec *magis*, mais aussi de *manducare*, *bellus* et du comparatif analytique avec *plus*. Sans doute l'examen détaillé de l'usage de ces diverses expressions pourrait en préciser pour chacune les conditions d'emploi et l'histoire ; mais il est un fait, c'est que chacune de ces expressions appartenait bien au latin. De même que l'existence du français s'accommode fort bien de l'existence

des mots et expressions régionales ou que l'occitan ne cesse pas d'être de l'occitan parce que, en telle région, mollet se dit *pompil* et, en telle autre, *botel*, de même le latin parlé n'en existait pas moins dans son unité, même si l'on disait en Gaule *manducare* et *plus bellus* tandis que l'on disait en Espagne *comedere* et *magis formosus*. De même s'il arrivait que la mode installât dans le latin parlé des mots ou expressions du latin noble, comme le fait a été prouvé dans le domaine du franco-provençal, par P. Gardette, l'unité du latin vulgaire n'était point entamée par ce caractère hétérogène inattendu.

Ainsi dès les plus anciens temps du latin existait ce que l'on appelle le latin vulgaire, un latin parlé à la fois un et varié. On conçoit que ce dernier ait pris plus d'importance chaque fois que les circonstances historiques accordaient plus d'importance aux orientations démocratiques qui, amoindrissant l'influence de l'*urbanitas*, donnaient plus de liberté aux tendances centrifuges de la langue. On a souvent signalé ce fait pour l'époque qui suit le III^e siècle apr. J.-C., et naturellement, pour celle qui commence avec la chute de l'Empire romain. C'est donc, en bref, dans l'existence du latin vulgaire à côté du latin classique dès les plus anciens temps, d'une part, et d'autre part, dans la diversité de ce latin, qui n'en restait pas moins un pour autant qu'il faut sans doute rechercher la clé première de l'origine des langues romanes. Cette clé, c'est donc, avant tout, la qualité même du latin.

II. — Les strats

Comment ce latin parlé, à la fois un et divers, en est-il arrivé à prendre les formes particulières que nous lui trouvons dans les actuelles langues romanes ? Faut-il penser qu'il y est arrivé par une

évolution immanente qui, de nouveautés en nouveautés, s'enchaînant les unes aux autres, à l'intérieur de son propre système, auraient fini par modifier ce système lui-même pour donner les diverses structures linguistiques que nous reconnaissons dans les langues romanes d'aujourd'hui ? Le soutenir tel quel, serait sans doute caricatural, ne serait-ce que par une certaine contradiction entre « évolution immanente » et « nouveautés ». A l'opposé, faut-il admettre que l'évolution du latin parlé aboutissant à le scinder en diverses langues « néo-latines » serait due à l'action externe d'idiomes divers qui, entrés en contact avec le latin, à des époques et dans des conditions diverses, auraient plus ou moins indûment introduit certains traits de leur propre structure dans la structure du latin, ébranlant par là même cette structure et entraînant des changements en chaîne qui auraient fait du latin ce qu'il est maintenant dans lesdites langues néo-latines ?

La seconde explication appelée souvent « théorie des strats », présentée parfois avec trop de facilité par les tenants de la linguistique historique, a été quelquefois combattue avec trop d'esprit de système par les linguistes dits structuralistes. S'il est vrai, comme le dit A. Martinet, que « toute langue change à tout instant », ces changements peuvent provenir de points de départ fort divers, dont les uns sont certainement immanents à la langue même, mais dont les autres ne sont pas moins certainement externes à la langue. Il est évident que, puisqu'il s'agit de langue, il convient d'accorder, dans l'évolution du latin en langues néo-latines, une priorité assurée aux faits linguistiques sans pour autant exclure les données historiques. A. Martinet (1) a

(1) A. MARTINET, *Eléments de linguistique générale*, Armand Colin, 1963, p. 217.

fort bien mis les choses au point dans les lignes suivantes :

« Les difficultés qu'on éprouve à identifier toutes les circonstances qui ont pu influencer sur la genèse d'un changement linguistique ne sauraient détourner les chercheurs d'une analyse explicative : il convient simplement de toujours donner la priorité à cet aspect de la causalité des phénomènes qui ne fait intervenir que la langue en cause et le cadre permanent, psychique et physiologique, de toute économie linguistique : loi du moindre effort, besoin de communiquer et de s'exprimer, conformation et fonctionnement des organes. En second lieu, interviendront les faits d'interférence d'un usage ou d'un idiome sur un autre. Sans faire fi jamais des données historiques de tous ordres, le diachroniste ne les fera intervenir qu'en dernier lieu, après avoir épuisé toutes les ressources explicatives que lui offrent l'examen de l'évolution propre de la structure et l'étude des effets de l'interférence. »

Le substrat. — On donne ce nom à la couche (strat) de faits linguistiques (assurés ou présumés) qui subsistent d'une langue disparue sur un territoire donné dans des conditions telles que leur influence paraît probable dans le nouvel état de la langue qui a remplacé la langue d'origine. Plus simplement on appelle « langue de substrat » la langue évincée : ainsi on parlera du substrat gaulois à propos des langues néo-latines parlées sur le territoire de l'ancienne Gaule. Il est, par ailleurs, à peu près certain que les langues de substrat ont recouvert elles-mêmes, en plusieurs territoires, des langues encore plus anciennes dont on devine les traces. On peut donner à ces dernières l'appellation de sous-substrat : c'est ainsi que l'on a pu déterminer un sous-substrat préindo-européen dans les Alpes et les Pyrénées. On ne sait pas grand-chose évidemment des rapports qui ont pu exister entre substrat et sous-substrat, mais il est raisonnable de penser que langue de substrat et langue de sous-substrat ont coexisté côte à côte pendant un certain

temps comme le gaulois, langue de substrat, a coexisté en Gaule avec le latin pendant un certain temps. Il est même arrivé que langue de sous-substrat et langue de substrat vivaient encore côte à côte quand s'est établie la langue qui devait les évincer l'une et l'autre : c'est ainsi que les inscriptions découvertes à Ensérune montrent qu'au II^e siècle encore apr. J.-C. étaient utilisés côte à côte, en Languedoc, le latin, le gaulois et une langue plus ancienne (ibère ?) qui reste pour l'instant inconnue même si l'on a pu déchiffrer son alphabet.

On conçoit que le maniement des données du substrat soit difficile et périlleux : on connaît mal ou très peu les langues en question (témoignages épigraphiques le plus souvent, rarement littéraires en ce qui concerne les témoignages anciens ; matériel onomastique et vestiges lexicaux décelés par la méthode historico-comparative, en ce qui concerne les témoignages modernes) ; on ne sait pas dans quelle mesure tel ou tel substrat était lui-même homogène et nous ne savons pas grand-chose sur la géographie linguistique des substrats. En plus de ces motifs qui doivent rendre prudent quand on invoque le rôle du substrat dans l'évolution du latin en langues néo-latines, on a voulu opposer aux explications par le substrat le décalage chronologique entre l'époque des contacts linguistiques et celui des changements dont ces contacts seraient la cause. Cette objection ne saurait être retenue à la lumière de faits qui se passent encore sous nos yeux : le français parlé dans les grandes villes des pays occitans n'est pas le français de Paris : maintien du *e* final atone qui a, pour conséquence, de garder le français régional méridional parmi les langues paroxytoniques, tandis que le français de France est devenu une langue oxytonique ; maintien

des divers *e*, muets en français de Paris, etc. Bien que la plupart des habitants des grandes villes du midi de la France aient perdu l'usage de la langue d'oc, certains depuis quatre générations, ils maintiennent dans leur français actuel les habitudes linguistiques de leur langue d'oc d'origine, langue de substrat. L'influence du substrat, bien que difficile à préciser dans le détail, ne saurait ainsi être niée : l'observation confirme les simples données du bon sens.

Un certain nombre d'hypothèses ont donc été fondées sur le substrat pour expliquer certains des changements survenus dans le latin.

A un substrat *italique* osco-ombrien on rattache l'assimilation amenant *-nd-* à *-nn-* et *-mb-* à *-mm-*. Cette assimilation qui caractérise le centre et le sud de l'Italie se retrouve dans la péninsule Ibérique en catalan, aragonais et au-delà des Pyrénées en Gasconne. Le même phénomène se rencontrant en picard, wallon, en sarde, on pourrait penser qu'il s'agit alors d'une assimilation spontanée. Mais, en ce qui concerne le nord-est de la péninsule Ibérique, on a constaté que le centre de diffusion de ce phénomène se trouve sur le territoire des deux villes actuelles de Huesca et de Llerida qui furent colonisées par des colons osques. On a donc pu admettre pour cette partie de l'Hispanie l'influence d'un substrat osque.

A un substrat *étrusque* on a attribué la transformation des consonnes sourdes intervocaliques en aspirées, phénomène qui caractérise les parlers toscans, d'où l'expression *gorgia toscana* employée pour le désigner. Le romaniste G. Rohlfs s'est montré sceptique à l'égard du substrat étrusque comme on s'est montré sceptique à l'égard du substrat grec ancien, soutenu par lui, dans les dialectes de l'Italie

méridionale dont on attribue la « grécité » à l'influence, postérieure, de la domination byzantine.

On reconnaît la présence d'un substrat *ligure* dans l'Italie du Nord — Piémont, Lombardie, Emilie — et en Provence : on lui attribue la faiblesse de *-l-* intervocalique, qui apparaît au maximum dans le dialecte de Gênes où *ala* devient *aa*.

Dans la péninsule Ibérique et débordant au-delà des Pyrénées on a eu sans doute un substrat *basque* et un substrat *ibère*. Les Basques sont les anciens Vascones apparentés aux Aquitains ; leur langue est certainement préindo-européenne apparentée aux langues caucasiennes. On estime que les Ibères sont venus de l'Afrique du Nord et sont d'origine libyque ou chamitique. On attribue au substrat ibérique la prothèse de *a* devant *r* (que l'on trouve par exemple dans beaucoup de parlers gascons où l'occitan général *lo riu* est représenté par l'*arriu*) et le passage de *f* à *h*.

L'influence du substrat *gaulois* est certainement la plus forte qui puisse se rencontrer en territoire roman. Les éléments celtiques sont considérables dans la toponymie de la France et de l'Italie du Nord. Bien que la question ait été longuement débattue et que la solution reste encore en suspens pour certains, on attribue au substrat gaulois le passage de *u* latin à *ü*. On lui attribue également le passage de *ct* à *it*.

L'étude du matériel onomastique et des vestiges lexicaux a montré l'existence d'un substrat méditerranéen préindo-européen. On l'a signalé notamment pour la Sicile (à côté d'un substrat grec), la Sardaigne, pour les pays où est attesté un substrat ligure, le ligure présentant deux strats, un méditerranéen, un indo-européen, dans les pays rhétiques. Ce substrat méditerranéen s'étend des Pyrénées au

Caucase. En France, il semble bien qu'il ait joué un rôle important dans les pays de langue d'oc. C'est ce substrat méditerranéen qui est souvent considéré comme un sous-substrat.

Le superstrat. — On parle de superstrat, lorsque, à l'inverse de ce qui s'est passé dans le cas du substrat, une nouvelle langue a exercé son influence sur la langue des autochtones qui s'est maintenue et a été adoptée par les nouveaux venus dans le cas où ceux-ci sont demeurés sur le territoire de la langue d'origine. En ce qui concerne les langues romanes on peut parler des superstrats germanique, arabe et slave. Si l'influence du superstrat arabe sur les langues ibéro-romanes et sur le sicilien, et celle du superstrat slave sur le roumain sont restées relativement limitées, celle du superstrat germanique a eu une importance beaucoup plus grande et plus généralisée.

Bien qu'ayant occupé presque toute l'Espagne jusqu'au XI^e siècle, époque de la Reconquista, bien que l'Espagne méridionale ait été occupée pendant huit siècles de 711 à 1492 et malgré le fait que les Mozarabes, chrétiens arabisés, aient été bilingues, le superstrat arabe se manifeste surtout dans le lexique. Encore convient-il de noter qu'une grande partie de ce lexique s'est étendue à toutes les langues romanes dans le domaine d'un certain nombre de sciences. On peut dire que le superstrat arabe a plus un caractère proprement culturel et de civilisation qu'un caractère linguistique. L'invasion slave en pays daco-roumain à partir du VI^e siècle semble avoir exercé son influence surtout à l'époque du proto-roumain. Elle se manifeste, elle aussi, surtout dans le domaine du lexique sans modifier profondément la structure du parler roman. Toutefois, le superstrat slave s'est fait sentir sur le système pho-

nologique ainsi que sur la morphologie du roumain ; la formation des mots doit notamment au slave un certain nombre de préfixes et de suffixes très vivants.

En ce qui concerne les Germains, on remarquera que l'influence linguistique du superstrat germanique a été beaucoup moins grande que l'importance des invasions proprement dites. C'est d'une façon indirecte, mais très importante, que l'influence germanique a agi sur la naissance des langues romanes. Les invasions germaniques ont fractionné le territoire linguistique roman et ont isolé la Dacie et la romanité dans la péninsule Balkanique, du reste de la Romania. Par l'occupation des hautes terres de la Suisse elles ont séparé le gallo-roman du rhéto-roman. Par celle de la Gaule et de l'Hispanie elles ont brisé beaucoup de liens entre ces territoires et l'Italie. C'est en interrompant les communications directes à travers la Romania, que les invasions germaniques ont favorisé largement le développement des différences existant dans le latin vulgaire des diverses provinces promouvant ainsi l'éclatement en divers idiomes.

L'influence linguistique du superstrat germanique a été spécialement étudiée par Walter von Wartburg. Ont été mis au compte de cette influence : le fait que le superstrat lombard aurait empêché la formation d'une frontière linguistique nette entre Bologne et Florence ; le cas particulier des dialectes normanno-picards qui dans la France d'oïl ne palatalisent pas *k* devant *a* (*chat* y est *kat*) ; la délimitation entre l'occitan et le franco-provençal ; la formation des deux langues romanes de France : le français et l'occitan. Ces théories n'ont pas été unanimement acceptées. Pour l'Italie, les observations de Schür, Merlo et Bolelli parlent contre

l'influence du superstrat lombard. Pour la France, A. Brun a soutenu que la division entre français et occitan repose essentiellement sur un dualisme préhistorique. La thèse de Brun reste, certes, trop hypothétique (tant que notamment on n'aura pas poussé plus avant les recherches toponymiques et de vestiges lexicaux concernant le Massif central, qui peut être également considéré comme un massif de sources occitanes), mais on lui a reconnu deux avantages sur la thèse du superstrat germanique de W. von Wartburg : tout d'abord, elle met l'accent sur le fait que la ligne qui divise la France en deux langues, d'est en ouest et non dans d'autres directions, est non seulement linguistique mais également juridique, ethnique, culturelle et agraire.

« Ces faits non exclusivement linguistiques diminuent considérablement l'influence du superstrat germanique, écrit B. E. Vidos (1), et nous prouvent que la ligne de démarcation entre nord et sud en France n'est point due aux établissements éventuels des Francs et des Burgondes, mais doit être beaucoup plus ancienne et doit plonger ses racines beaucoup plus profondément dans l'histoire du pays. »

Le second avantage de la théorie de Brun est que, alors que la thèse du superstrat germanique considère la question de la formation d'une limite linguistique entre le français et l'occitan comme résolue précisément par l'action de ce superstrat, la thèse du dualisme préhistorique nous invite à reporter à plus tard la solution de la question. Il est permis de penser qu'il y a, en effet, encore beaucoup de travail et de découvertes à faire dans ce domaine.

Si on ne peut mettre au compte du seul superstrat germanique la division de la France en langue

(1) B. E. VIDOS, *Manuale de linguistica romanza*, Firenze, 1959, p. 251.

d'oc et langue d'oïl, il est certain, par contre, que l'influence germanique a contribué grandement à faire du français la langue la plus germanisée des langues romanes. On trouve des éléments germaniques dans toutes les langues romanes, mais la longue période de bilinguisme roman-germanique que connut la France du Nord durant la période mérovingienne et une très grande partie de la période carolingienne (jusqu'aux dernières années du IX^e siècle) a fait que le nombre des emprunts au vocabulaire germanique en français est le plus grand de toutes les langues romanes. De plus, l'influence germanique se retrouve dans la phonétique (adoption du *h* aspiré), dans la morphologie (importance des suffixes d'origine germanique), dans la toponymie (beaucoup plus germanique dans le nord de la France que dans les pays d'oc) et dans la syntaxe, tout au moins dans celle de l'ancien français, en ce qui regarde, par exemple, la préférence pour l'ordre déterminant-déterminé : ainsi la toponymie nous offre *Neuchâtel*, *Chaumont*, etc., au nord pour *Castelnau*, *Moncaut/Moncaup* en pays d'oc. (Même de nos jours la préférence pour la postposition du déterminant en oc est, dans la langue courante, un trait syntaxique qui distingue cet idiome du français dans de très nombreux cas où celui-ci use de l'antériorité : « la grand-ferme » pour « la boria granda », « un bel arbre » pour « un aubre bèl », etc.)

A propos des strats, on signalera que l'on parle aussi d'*adstrat* pour désigner l'influence réciproque qu'exercent deux langues vivant l'une près de l'autre. Mais l'influence de l'*adstrat* concerne plutôt l'histoire des langues romanes postérieurement à l'époque de leur formation. De plus, bien souvent, il est délicat de trancher entre superstrat et adstrat.

III. — Groupements humains et communications

Ni la qualité du latin vulgaire, ni l'existence de strats divers ne peuvent à elles seules expliquer l'origine des diverses langues romanes à travers l'Empire romain.

Très certainement en faveur d'une différenciation de la langue parlée ont joué : les différences chronologiques du latin importé dans les différentes provinces ; les différences de qualité de ce même latin importé dans telle ou telle région par des immigrants qui étaient moins des Romains que des Italiotes latinisés devenus citoyens romains ; les différences d'ordre socioculturel, la prédominance d'un groupe social déterminé dans la colonisation d'une province ayant pu favoriser l'implantation d'une certaine variété socioculturelle du latin ; le processus de latinisation qui a pu différer suivant cette même prédominance : colonisation surtout agricole, ou peuplement urbain, entraînant comme conséquence une latinisation surtout par voie orale dans le premier cas, davantage par voie scolaire dans le second cas ; c'est ainsi qu'on a pu soutenir qu'un des traits qui ont servi à distinguer une Romania linguistique orientale (italien, roumain) et une Romania linguistique occidentale (français, espagnol, occitan, etc.), à savoir la non-conservation ou la conservation du *s* final serait dû à la conformité avec la prononciation populaire dans le cas de l'Orient, à une prononciation culte dans le cas de l'Occident. Très certainement encore jouaient en faveur de la différenciation les divers substrats des provinces latinisées.

Mais ces diverses raisons n'auraient sans doute pas suffi à créer la différenciation linguistique actuelle de la Romania si un certain nombre de

facteurs d'unité n'avaient cessé de s'exercer à un moment de l'histoire. Ces facteurs dépendaient de l'unité politico-administrative de l'*Imperium* : action du latin écrit dû au prestige littéraire mais aussi et surtout à l'autorité et, davantage encore, à l'utilité de la langue de l'administration et de l'école ; action des relations nombreuses et constantes qui unissaient les provinces à Rome et entre elles et exigeaient l'emploi d'une langue commune, comprise de tous et utilisée par tous aussi bien dans le domaine de l'administration que dans celui du commerce. Ces facteurs en entraînaient un autre qui était au fond le plus important et le plus efficace pour le maintien de l'unité de langue : la création et le développement d'une *koinê* nécessaire dans toutes les provinces pour maintenir la communication entre colons, fonctionnaires, militaires, commerçants issus de différentes régions de l'Empire. On a constaté que les langues importées dans les pays de colonisation sont plus homogènes dans leur habitat nouveau que dans leur propre pays d'origine. Loin de pousser à la différenciation de la langue, l'étendue même de l'Empire romain tendait à en renforcer l'unité par les besoins mêmes de la communication.

On peut donc estimer qu'aussi longtemps que l'Empire demeura fortement organisé et uni, aussi longtemps que Rome demeura la capitale incontestée de l'Empire, les facteurs d'unité ont dû largement contrebalancer les facteurs de différenciation. Mais, lorsque l'importance de Rome commença à s'affaiblir, dès avant même le III^e siècle, par la généralisation de la citoyenneté romaine, la formation d'une nouvelle classe dirigeante issue des diverses parties de l'Empire, la proclamation d'empereurs d'origine provinciale, la formation de cou-

rants commerciaux qui ne passent pas par Rome, le rayonnement culturel et linguistique de la capitale s'en trouva sans doute pour le moins affaibli. La réforme de Dioclétien, dont le but était de lutter contre l'anarchie politique, régularisa au fond une situation effective. A partir de 284 l'Empire est divisé entre quatre préfectures elles-mêmes subdivisées. On a cru observer précisément que c'est à cette époque que remontent certaines particularités linguistiques qui donnent une image tripartite de la Romania : Romania orientale (péninsule Italienne, Sicile, Balkans), Romania insulaire (Corse, Sardaigne), Romania occidentale (Afrique, péninsule Ibérique, Gaule transalpine, Gaule cisalpine, Norique et Pannonie). La Romania insulaire continuait à distinguer les résultats de *e* long et de *i* bref, de *o* long et de *u* bref du latin classique. La Romania orientale amuissait *s* finale et gardait les sourdes intervocaliques. La Romania occidentale conservait l'*s* finale et sonorisait les consonnes sourdes intervocaliques.

La chute de l'Empire romain en 476 a sans doute marqué définitivement la fragmentation de l'unité du latin réduit à ne plus pouvoir développer sa fonction de langue de communication et de liaison de l'Etat. Les grandes invasions en ruinant l'Empire ont bouleversé la société ; en disloquant son organisation de fond en comble, elles ont modifié les relations entre groupes humains ; créant de nouvelles frontières, elles remodelaient ces groupes. Par ailleurs en abaissant le niveau culturel, sinon en anéantissant toute culture, elles abandonnaient à elle-même la langue de l'Empire déchu.

Les grandes invasions germaniques ont isolé la Romania de l'Est du reste de la Romania ; elles ont coupé le domaine rhéto-roman du domaine gallo-

roman, les territoires de la Gaule et de l'Espagne de ceux de l'Italie. On peut, à partir du *v*^e siècle, parler de balkano-roman, de rhéto-roman, d'italo-roman, de gallo-roman, d'ibéro-roman. Cela, d'une façon toute relative et sur la base des diverses langues romanes, la documentation faisant défaut pour préciser ces notions. En tout cas, on ne peut plus parler de latin vulgaire.

Le processus de diversification ne s'en tient d'ailleurs pas à ces grandes divisions. La langue, partout abandonnée à elle-même, ne s'arrête pas sur le chemin des différenciations. Celles-ci sont favorisées par tous les jeux linguistiques : jeu des substrats libérés que ne contient plus aucun modèle supérieur, auquel s'ajoute l'action des sous-substrats dispersés dans les cantons les plus reculés laissés à eux-mêmes ; jeu des adstrats germaniques divers dans la période de bilinguisme plus ou moins poussé suivant les régions ; jeu enfin, et probablement surtout, bien que la question soit difficile à étudier, des ébranlements locaux en chaîne à l'intérieur du système de la langue, à partir de tels ou tels changements plus ou moins fortuits qui finissent par modifier considérablement la structure linguistique en chaque point de l'espace où un groupe humain se trouve constitué en son isolement. A l'extrême, le latin vulgaire aboutit à ce pullulement de patois qui fait dire à ceux qui les pratiquent encore aujourd'hui : « *le patois change dans chaque village* ». Et il est, en effet, possible d'étudier, aujourd'hui, la structure particulière de tel ou tel parler tout à fait ponctuel.

Mais, en fait, ce processus extrême de diversification s'est constamment heurté à des forces qui le ralentissaient ou même en inversaient le cours sur nombre de points par leurs imbrications mêmes et

leurs contradictions, à l'intérieur du système, forces largement favorisées par des facteurs socioculturels qui, pour n'être plus ceux qui jouaient dans l'Empire romain, n'en continuaient pas moins à soutenir leur rôle, réagissant sur l'état d'anarchie linguistique. Parmi ces facteurs on a souvent mis en évidence celui du rôle tenu par les centres de gouvernement ecclésiastique, qui se substituèrent à l'administration romaine. H. Morf a, par exemple, souligné le rôle des circonscriptions épiscopales. Les limites s'en confondaient, d'après lui, avec les limites des communications, les habitants d'un diocèse se réunissant dans les centres ecclésiastiques secondaires et dans la cité de l'évêque où sa juridiction obligeait les habitants des villages à se rendre. Par ailleurs, les jours de fêtes chrétiennes monopolisaient les marchés grands et petits. Par le fait le *diocèse* constituait une unité administrative et culturelle qui a certainement eu des conséquences linguistiques, en contrecarrant l'atomisation naturelle des parlers de chaque petit groupe humain et en favorisant un mouvement inverse d'unification si bien qu'on a pu penser que les phénomènes linguistiques diffusés par les habitants appartenant à un même diocèse furent arrêtés aux limites du diocèse. La diversification linguistique de la Romania se serait ainsi ramenée aux limites des groupes humains soumis à une juridiction épiscopale.

Là se ramènerait pratiquement le rôle du *christianisme* dans l'apparition des langues romanes. Sans doute le fait que le christianisme occidental, à la différence du christianisme oriental qui a usé de diverses autres langues (gothique, slave, copte, etc.), a toujours utilisé le latin comme langue de l'Eglise, a contribué à maintenir une certaine unité « latine ». Mais cette unité ne se maintint que,

si l'on peut dire, pour l'usage interne des ecclésiaux. Et difficilement encore puisque le pape Zacharie (741-752) fut obligé, par exemple, d'accorder son approbation aux baptêmes donnés par les prêtres avec la curieuse formule : « In nomine de Patria, Filia et Spiritua Sancta ! » Le processus de différenciation entre latin et parlers populaires était trop avancé pour qu'il en fût autrement. Avant même ce qu'on appelle la « renaissance carolingienne du latin » qui devait définitivement trancher entre latin et langue populaire, l'Eglise avait dû tenir compte de cette évolution comme on vient de le voir. En 813 le Concile de Tours prescrivit officiellement aux prêtres d'user de la *rustica romana lingua*, afin d'être plus facilement compris du peuple. L'Eglise encourageait ainsi les parlers populaires en abandonnant l'usage du latin dans son ministère parmi le peuple. Dès lors le parler populaire devenait dans chaque diocèse la langue normale et la *rustica romana lingua* se trouva par le fait adaptée à chaque diocèse.

On a constaté enfin que les limites des diocèses correspondaient souvent à celles des anciennes *civitates* qui, elles-mêmes, concordaient *grosso-modo*, avec les limites des peuplades primitivement connues. Les groupements humains et les limites de leurs communications à l'époque de la cristallisation des diocèses chrétiens correspondaient donc souvent à ceux des *civitates* romaines et des *nationes, gentes* ou *tribus* préromaines. Ainsi une action fort complexe et très ancienne a dû se produire qui a pu arrêter la diversification linguistique de la Romania aux limites des circonscriptions ecclésiastiques. Il ne faudrait pas toutefois accorder plus d'importance qu'il ne convient à ces limites. Si certaines études de géographie linguistique ont montré le

bien-fondé de cette façon de voir, d'autres études en ont montré les limites. L'exemple du Gévaudan est caractéristique à ce sujet. Le diocèse « gabalitin » recouvre, encore aujourd'hui, à peu près exactement le territoire de la *civitas* des Gabales. Or le territoire du Gévaudan est nettement, du point de vue linguistique, subdivisé en cinq zones dialectales : une zone sud-est, une zone sud-ouest, une zone nord-est, une zone nord-ouest et une zone centrale où se trouvent les deux petites villes historiquement les plus importantes, Marvejols, d'origine gauloise, et Mende, le siège de l'évêché, d'origine probablement plus ancienne, cette zone centrale se caractérisant par la réunion de traits linguistiques appartenant à chacune des quatre autres zones. De plus une limite linguistique des plus importantes traverse le Gévaudan d'est en ouest, celle de la palatalisation de *c, g* devant *a*, la non-palatalisation caractérisant les zones sud-est et sud-ouest. Enfin il se trouve que les limites du Gévaudan sont très nettement délimitées par la géographie physique, ce qui n'a pu que renforcer les facteurs historiques socioculturels d'unité, qui ont été sans aucun doute très important dès l'époque de la formation des parlers romans, puisque l'évêque de Mende, dès les origines, fut également comte de Gévaudan. Ces facteurs existaient sans doute déjà dès l'époque de la *civitas* gallo-romaine et de la nation gabale pré-romaine, sous la domination gauloise. Il faut donc admettre que le groupement humain diocésain, pas plus que les groupements de la *civitas* gallo-romaine et de la nation gauloise des Gabales n'ont été seuls décisifs dans la détermination des idiomes romans du Gévaudan. C'est dire que les facteurs administratifs sont loin d'avoir joué le rôle prédominant qu'on leur a parfois attribué. Un exemple très

particulier, pris en Gévaudan même, le prouve clairement. La commune (paroisse ancienne) de Saint-Etienne-du-Valdonnez est, du point de vue linguistique, très nettement séparée en deux par un faisceau d'isoglosses qui constitue un véritable mur de séparation linguistique : le bourg même de la paroisse au pied du mont Lozère appartient à une zone linguistique, les hameaux situés sur les hauteurs vallonnées du mont Lozère à une autre zone linguistique (1).

Ainsi le groupement, constituant un noyau administratif, ne peut à lui seul expliquer la formation de l'idiome ; c'est que sans doute les communications ne se réduisent pas aux seules liaisons administratives et culturelles. Entrent en jeu à la fois les liaisons naturelles qui forment un réseau serré et continu, d'une part, et, d'autre part, les liaisons à grande distance. Pour les premières on notera, par exemple, que la zone dialectale sud-est du Gévaudan recouvre le bassin des eaux qui coulent vers l'est et le sud (gardons et Cèze), la zone nord-est, le bassin des eaux tributaires de l'Allier, la zone nord-ouest, le bassin des eaux tributaires de la Truyère, la zone centrale, le bassin des eaux tributaires du Lot. Pour les secondes, on rappellera le fait maintes fois signalé que l'importante route romaine de grande communication qui va de Lyon à Bordeaux forme un vaste cercle autour du Massif central suivant parallèlement la limite entre occitan et français. Il va de soi que le fait n'est pas fortuit et l'on admet que cette route a contribué à la subdivision de la Gaule en Nord et Sud tant du point de vue linguistique que non linguistique, même si le tracé de cette route a pu obéir à des éléments préexistants

(1) Cf. C. CAMPROUX, *Essai de géographie linguistique du Gévaudan*.

qui avaient déjà, d'une certaine façon, déterminé la double séparation linguistique et non linguistique.

On peut donc dire en résumé que l'origine des langues romanes est due à des facteurs nombreux et complexes, les uns linguistiques, les autres non linguistiques, qui ont joué dès les temps les plus anciens, qui ont été plus ou moins contenus par l'extension du latin grâce à l'Empire romain, et qui finalement, modifiés par le latin lui-même et le modifiant, ont repris force avec la chute même de l'Empire romain.

IV. — Apparition des langues romanes

Nous n'en sommes encore ici qu'au stade de la formation des langues romanes, c'est-à-dire au stade de la diversification linguistique de la Romania et non point à celui de la construction et de l'élaboration des langues romanes telles que nous pouvons les caractériser aujourd'hui.

D'après ce qui précède on se rend compte que l'on ne peut parler de naissance des langues romanes : les idiomes romans prolongent en fait le latin sans que la moindre solution de continuité puisse être décelée. A quel moment, dès lors, peut-on situer la prise de conscience de l'existence des idiomes romans ? A peu près certainement vers la fin du VIII^e siècle et le début du IX^e. Nous avons vu qu'en 813 le Concile de Tours demande aux prêtres de prononcer leurs homélies *in rusticam romanam linguam*. Vers la même époque apparaissent les *Gloses* : *Gloses de Reichenau* qui expliquent les mots latins les plus difficiles du lexique biblique par des mots plus populaires sous l'habit latin desquels transparaissent les mots romans ; *Gloses de Kassel* qui donnent les équivalents germaniques de mots

latins qui laissent très souvent transparaître les mots romans. De la même époque, on cite également l'*Indovinello veronese* (*Dévinette de Vérone*) qui serait le plus ancien document de tout le domaine roman où se révélerait nettement la langue vulgaire. En voici le texte jugé le meilleur, établi par A. Monteverdi (1) : *Se pareba boues - alba pratalia araba - albo uersorio teneba - negro semen seminaba*. Ce qui signifierait : « l'action d'écrire », les bœufs étant les doigts, le champ blanc le parchemin, la charrue blanche la plume d'oie et la semence noire les lignes tracées à l'encre. Monteverdi pense que l'auteur a usé ici du « vulgaire maternel », mais on peut se demander, à bon droit, si la « devinette » n'est pas composée dans ce latin semi-vulgaire qui devait être au Moyen Âge l'outil de communication entre condisciples pas trop sûrs de la grammaire et du lexique latin » (2).

En tout cas, ni l'*Indovinello veronese*, ni les gloses de Reichenau ou de Kassel, pas plus que les gloses en espagnol de San Millan de la Cogolla (province de Logrono) et de Silos en Castille qui sont du X^e siècle, n'attestent l'intention d'écrire directement en vulgaire roman. Il en va de même pour les mots et les phrases en occitan que l'on trouve de très bonne heure çà et là dans les documents du midi de la France (VII^e-IX^e siècle). Le plus ancien texte suivi, intentionnellement rédigé en vulgaire roman, que nous ayons, est le texte des *Serments de Strasbourg* (842). On a beaucoup discuté pour savoir en quel parler roman les serments sont rédigés, même si l'on parle communément de « plus ancien document du français ». Pour G. Paris ce texte est

(1) A. MONTEVERDI, A proposito dell' indovinello veronese, dans *Saggi neolatini*, Rome, 1945.

(2) G. TAGLIAVINI, *Le origine delle lingue neolatine*, Bologne, 1964, p. 460.

écrit en français du Nord plus particulièrement en dialecte picard. Pour d'autres romanistes et, en dernier lieu, pour A. Ewert il serait écrit en une variété dialectale du Sud-Ouest, peut-être en Poitevin. H. Suchier pensait qu'il s'agissait d'un texte rédigé en un parler franco-provençal et proposait même le dialecte de Lyon. Plus tard, A. Tabachovitz croyait pouvoir conclure que l'auteur du document avait été un bilingue d'origine lorraine. Pour R. A. Hall la langue des *Serments* serait un « préfrançais ». Dans un article de *Romania* en 1939, F. Lot estime que, « vu la composition hétéroclite de l'armée du plus jeune des frères » en présence, « le rédacteur du serment a usé d'une langue artificielle conventionnelle, d'une sorte de roman commun ». Dans un des articles des *Actes du VIII^e Congrès international des Etudes romanes*, en 1956, A. Castellani considérant que soit dans la composition de l'armée de Charles, soit dans la composition de sa chancellerie abondaient les éléments d'Aquitaine, propose de voir dans la langue du texte une langue de chancellerie de la cour d'Aquitaine, ce qui transférerait la variété dialectale du domaine français au domaine occitan.

Quoi qu'il en soit de ces opinions diverses, il faut savoir que des *Serments de Strasbourg* nous ne possédons le texte que dans une transcription datant approximativement de l'an 1000. Si bien que le plus ancien texte français est, en définitive, la *Séquence de sainte Eulalie*, composée vers 900, peut-être au couvent de Saint-Amand, au nord de Valenciennes dans un parler à la limite du picard et du wallon. Dans le sud de la France, le témoignage littéraire le plus ancien en occitan se trouve dans l'*Aube* bilingue (latin-occitan) notée sur les marges d'un manuscrit du x^e siècle. On a beaucoup discuté sur

la langue et le pays d'origine de la *Passion* dite de Clermont-Ferrand (fin du x^e siècle). Dans l'étude qu'il a publiée en 1962 sur ce texte, d'Arco Silvio Avalle déclare : « La *Passion* est sans aucun doute le produit d'un *Grenzgebiet* non seulement géographique mais en premier lieu culturel. » En tout cas la langue du manuscrit qui nous l'a conservée est nettement occitane : on peut en juger dès le premier quatrain :

Hora vos dic vera raizon
de Jesu Christi passïun :
los sos affanz vol remenbrar
per que cest mund tot a salvad.

On peut donc dire que, dès le x^e siècle, l'existence des langues romanes est donc bien assurée au-delà des simples attestations de caractère linguistique : la conscience de cette existence est suffisamment nette pour avoir donné naissance à des œuvres de caractère littéraire, notamment sur le territoire de l'ancienne Gaule au nord comme au sud. En Espagne le témoignage des *kharjas* (strophe finale des poèmes lyriques dits *muwassahat*) en langue romane montre qu'il en fut de même de bonne heure. Les langues romanes étaient prêtes pour la *Chanson de Roland* (aux environs de 1100), pour la poésie lyrique de Guilhem de Peitieu (1071-1127) et pour le *Cantar de mio Cid* (xii^e siècle). L'usage littéraire de l'italien apparaîtra nettement au xiii^e siècle notamment dans l'utilisation qu'en fait le troubadour Raimbaut de Vaqueiras. Mais dès lors nous n'en sommes plus aux origines des langues romanes : nous entrons dans la phase du développement qui en fera les langues que nous connaissons. On passe de la dislocation du latin à la reconstruction des langues néo-latines.

CHAPITRE III

LES LANGUES ROMANES

I. — Développement

1. Patois, dialecte, langue. — On a souvent défini les patois d'une façon erronée quand il s'agit des parlers populaires issus du latin. Ainsi le *Robert* reproduit la définition suivante du patois donnée par A. Brun :

« Une *langue* est un dialecte qui a réussi, un *patois* est un dialecte qui s'est dégradé... Le dialecte dégénère quand il ne s'écrit plus ; ne s'écrivant plus, il se diversifie en multiples variétés, dissolution qui se précipite quand les classes dites supérieures cessent de l'employer dans l'usage oral, pour adopter une langue commune. »

La définition d'A. Brun était sans doute inspirée par la situation de ce romaniste par rapport au provençal que le peuple de chez lui appelle, à tort, patois. Sa définition est en quelque sorte une réaction en faveur du provençal-patois. Nous avons vu précédemment que l'aboutissement du procès de dissolution du latin livré à lui-même était précisément l'atomisation du latin en une poussière de parlers locaux. Ce sont ces parlers locaux que l'on appelle patois en linguistique romane (cf. par exemple le titre de l'ouvrage d'Antonin Duraffour, *Glossaire des patois franco-provençaux*). Au point de vue linguistique tel parler local du Jura ou des

Vosges, même sans prestige aucun, même en voie de disparition complète, est un idiome aussi *parfaitement* roman que n'importe quelle langue nationale. C'est ainsi que Vendryès (cité également par le *Robert*) a justement écrit :

« Les parlers de nos campagnes, les patois, comme on les appelle, ont souvent des règles plus strictes que les langues apprises dans les grammaires. »

Littré se faisait des patois et des dialectes la même idée qu'A. Brun. Il écrit, en effet, à l'article « Dialecte » :

« Avant le *xiv^e* siècle, il n'y avait point en France de parler prédominant, il y avait des dialectes... Après le *xiv^e* siècle il se forma une langue littéraire et écrite, et les dialectes devinrent des patois. »

En réalité, les patois ou parlers populaires sont *coexistants* avec ce que l'on appelle les dialectes. La différence entre eux n'est point de nature ; elle réside seulement dans la différence quantitative et qualitative des traits linguistiques qui les caractérisent. Par rapport aux patois un dialecte est un ensemble de parlers dont aucun n'est exactement semblable à l'autre, mais qui ont un certain nombre de traits linguistiques en dénominateur commun. Certains de ces traits sont privilégiés qualitativement par l'importance qu'ils ont aussi bien au regard du linguiste qu'aux oreilles de ceux qui parlent. C'est ainsi que le nord-gévaudanais se distingue du sud-gévaudanais par la palatalisation de *c*, *g* devant *a* (*Kami* au sud, *tchami* au nord : chemin). A ce trait cependant qui forme limite dialectale s'en ajoutent plusieurs autres moins sensibles individuellement mais qui unis forment un faisceau qui dans le cas présent constitue un véritable mur linguistique. De part et d'autre de cette limite, les parlers locaux qui ont en communs dénominateurs ces

traits linguistiques se diversifient à leur tour jusqu'à l'émiettement, notamment dans le domaine de la phonétique mais aussi dans celui de la morphologie et du vocabulaire, par le fait qu'aucune des aires phonétiques, morphologiques et lexicologiques ne recouvre exactement le même domaine, certaines de ces aires s'étendant même au-delà du mur ci-dessus. Comme, par ailleurs, les limites de ces aires diverses se regroupent en faisceaux partiels soit au sud soit au nord du faisceau principal, les divers parlers se regroupent en dialectes de moindre étendue (auxquels on peut donner le nom de sous-dialectes). De la sorte, et non plus quant au rapport patois/dialecte, mais quant au rapport des dialectes entre eux, on pourra, d'après cet exemple, définir le dialecte, comme le fait Marouzeau dans son *Lexique de la terminologie linguistique* :

« Un dialecte se définit par un ensemble de particularités telles que leur groupement donne l'impression d'un parler distinct des parlers voisins, en dépit de la parenté qui les unit. »

On voit par ce qui vient d'être dit que peu importe pour la notion de dialecte, du strict point de vue linguistique, que le dialecte soit écrit ou non. Au temps où les troubadours d'origine gasconne écrivaient leurs poèmes dans l'occitan général des troubadours, le dialecte gascon n'en existait pas moins en fait comme dans la conscience des écrivains et des grammairiens. De nos jours, le fait que le plus important écrivain du Gévaudan, F. Remize, ait écrit dans le dialecte central du Gévaudan n'empêche nullement les gens des Causses gévaudanais d'avoir conscience que les parlers dont ils usent appartiennent à un autre dialecte ou sous-dialecte, conscience qui correspond aux faits linguistiques. Ce qui est vrai, c'est que lorsqu'on a affaire à un dialecte

écrit, moyen d'une littérature, on se trouve sur le chemin qui conduit du dialecte à la langue. On peut alors dire non pas comme le fait A. Brun « qu'une langue est un dialecte qui a réussi » mais « qu'un dialecte qui a réussi est une langue ». Ainsi la littérature écrite en dialecte wallon a pu faire parler de « langue wallonne ». Mais la « réussite » d'un dialecte est quelque chose de relatif. Si le dialecte toscan a réussi à devenir la langue italienne, on continue de parler de dialecte anglo-normand ou de dialecte picard malgré l'importance culturelle qu'eurent ces dialectes au Moyen Âge. Le dialecte reste essentiellement un langage parlé sur un espace relativement réduit, inclus dans un espace linguistique plus vaste à l'intérieur duquel une certaine parenté, comme le dit Marouzeau, l'unit à d'autres parlers voisins avec lesquels il constitue une sorte de confédération à quoi on peut donner le nom de langue. C'est ainsi que le grec ancien n'a jamais été qu'un faisceau de dialectes grecs.

On aperçoit de la sorte la difficulté de définir la notion de langue par rapport à celle de dialecte. Une langue peut être, en effet, comme le dit A. Brun, un dialecte qui a réussi. C'est le cas du français, de l'italien, de l'espagnol, issus des dialectes de l'Ile-de-France, de Florence et de la Vieille-Castille. Ce n'était pas le cas de l'ancien occitan en qui l'on s'accorde à voir une *koiné* formée de divers traits dialectaux ; ce n'est pas le cas de l'occitan moderne qui résulte d'un dia-système entre les divers dialectes d'oc. C'est encore moins le cas du sarde, par exemple, qui, bien que n'ayant pas « réussi », n'en est pas moins considéré, par les linguistes, comme une langue à part. L'exemple du sarde, auquel il convient d'ajouter, sans doute, celui du rhéto-roman, nous indique qu'on ne peut définir perti-

nemment la notion de langue par la présence ou l'absence de traits politiques ou culturels : unité et indépendance politique, fonction culturelle ou littéraire. Ces traits valent pour l'espagnol, l'italien, le français. Seule la fonction culturelle et littéraire vaut pour l'occitan. Le catalan est resté une langue bien qu'il ne soit plus celle d'un pays indépendant. Le galicien est considéré comme un dialecte espagnol alors que le portugais a rang de langue, bien que le galicien dans sa période la plus ancienne ait formé une unité avec le portugais et que la langue des poètes galiciens ne se distinguât pour ainsi dire pas du portugais d'alors.

Il n'existe donc point de critère précis et unique qui puisse permettre de cerner exactement la différence entre dialecte roman et langue romane, puisqu'une langue peut être déterminée à la fois par des critères proprement linguistiques mais aussi par des éléments d'ordre différent : facteurs historiques, politiques, culturels, littéraires. Il est indéniable cependant qu'il existe des langues romanes définies, chacune, par un faisceau de caractères qui lui sont propres, mais, si l'on voulait s'en tenir à un strict point de vue linguistique, il n'existerait que des dialectes romans, diversement bien ou mal traités par le destin des hommes qui les parlent, ou les ont parlés.

2. Les dialectes romans. — Il est difficile de distinguer et de classer les divers dialectes romans. La chose se comprend aisément quand on ne perd pas de vue la façon dont ces dialectes sont sortis du latin. En réalité il existe ce que l'on a appelé justement une « Romania continue » au sein de laquelle on peut seulement discerner certains contours. C'est ce qui explique les classifications

diverses que l'on a proposées des idiomes romans. Ainsi A. Monteverdi distingue : 1) les parlers daco-illyro-romans ; 2) les parlers italo-romans ; 3) les parlers gallo-romans ; 4) les parlers ibéro-romans. C. Tagliavini, lui, adopte la même répartition, mais il place le dalmate que Monteverdi range avec les parlers daco-illyro-romans, dans le groupe italo-roman, et le catalan placé par Monteverdi avec les parlers ibéro-romans, dans le groupe gallo-roman. P. Bec, pour sa part, reprend à peu près le même classement, mais il le complète par la reconnaissance d'un groupe occitano-roman intermédiaire entre le gallo-roman proprement dit (français et franco-provençal) et l'ibéro-roman (portugais et espagnol) : l'occitano-roman comprend dès lors l'occitan moyen, le nord-occitan, le gascon et le catalan. W. D. Elcock distingue cinq groupes : 1) gallo-roman (oïl et oc) ; 2) hispano-roman (espagnol, portugais, catalan) ; 3) italo-roman (italien, sarde) ; 4) rhéto-roman ; 5) balkano-roman (roumain). Sa division est d'ailleurs hésitante : il considère le catalan apparenté au gallo-roman, les dialectes italiens du Nord, intermédiaires entre l'italo-roman et le gallo-roman, et il tient le dalmate pour un dialecte de transition entre l'italo-roman et le balkano-roman. Quant à Walter von Wartburg, renonçant à des subdivisions peu certaines, il propose de partager la Romania en trois groupes seulement : un groupe français proprement dit, un groupe roumain et un groupe méridional ou méditerranéen comprenant toutes les autres langues romanes.

Etant donné la complexité du problème et en l'absence d'un consensus des spécialistes, on peut proposer la division suivante, en dix groupes, des dialectes romans, qui s'inspire, *grosso modo*, du

point de vue de W. von Wartburg : espagnol, portugais, italien, sarde, occitano-roman, rhéto-roman, dalmate, roumain, franco-provençal, français.

Les dialectes *espagnols* recouvrent le territoire politiquement espagnol excepté la zone catalane, la zone basque et la zone galicienne. La structure dialectale de l'espagnol est beaucoup plus pauvre et plus uniforme que celle du français et de l'italien. Le fait est dû à l'action de la *Reconquista* qui a considérablement avantagé le castillan. Au Moyen Age existait dans les pays sous domination arabe un dialecte roman, le mozarabe. Le mozarabe, parler dépourvu de prestige, fut rapidement adapté par ses usagers au castillan dans la progression de celui-ci vers le sud en même temps que se développait la *Reconquista*. On distingue, dans les dialectes espagnols, les dialectes du Nord ; l'asturien dans la province d'Oviedo, le léonais dans la province de Léon (bien que, dans le détail, les limites linguistiques et administratives ne coïncident pas), et l'aragonais, principalement dans la province de Huesca au pied des Pyrénées. Ces dialectes remontent au latin vulgaire ibérique et ils connurent un usage littéraire. Tout le centre de la péninsule est occupé par le dialecte castillan. C'est dans ce dialecte que fut écrit le poème du Cid (avec traces de parlers navarro-aragonais) vers 1140 et que fut traduit en 1241, par ordre du roi Ferdinand III, le *Forum Ludicum* (*Fuero Juzgo*), code fondamental de l'Espagne dérivé de la *Lex Visigothorum*. Au sud on a l'andalou et les parlers des îles Canaries. Les dialectes espagnols du Sud ne remontent pas directement au latin vulgaire mais proviennent de différenciations régionales du castillan. Les juifs espagnols expulsés d'Espagne à la fin du ^{xv}^e siècle importèrent dans les pays où ils trouvèrent refuge,

surtout l'Afrique du Nord, la Turquie, Rhodes, la Macédoine, une variété archaïque d'espagnol connu sous le nom de judéo-espagnol, partout en décadence actuellement.

Les dialectes *portugais* sont encore plus homogènes que les dialectes espagnols, bien que l'on distingue quatre zones dialectales : au nord le *falar minhoto* à l'ouest, et le *falar trasmontano* à l'est ; au centre le *falar do Castelo Branco e Portalegre* et le *falar beirão* à l'est, et le *falar de Baixo Vouga e Mondego* à l'ouest ; au sud, zone qui englobe la moitié du pays (Estremadura, Alentejo et Algarve), on distingue l'*alentejano* et l'*algarvio* ; la quatrième zone comprend les dialectes insulaires des Açores et de Madère, qui offrent beaucoup de traits communs avec les dialectes du Sud. La relative homogénéité de ces parlers s'explique par le mouvement du nord vers le sud de la Reconquête au royaume de Portugal. Au groupe des parlers portugais se rattachent les parlers galiciens situés sur le territoire de l'ancien royaume de Galice qui forme la partie nord-occidentale de l'Espagne. Au Moyen Age parlers portugais et parlers galiciens étaient intimement liés ; ils furent illustrés par la lyrique portugo-galicienne.

A l'opposé des parlers espagnols et portugais, les parlers du groupe *italien* sont excessivement variés. On a pu essayer de traduire une nouvelle de Boccace en près de 700 dialectes ! Les dialectes italiens sont très vivaces. Alors qu'en France les parlers d'oïl sont pratiquement en voie de disparition, les dialectes italiens, surtout ceux du Nord, sont encore souvent parlés même dans des milieux cultivés et il existe, par toute l'Italie, une importante littérature dialectale. On divise, en général, les parlers italiens en trois grandes zones linguis-

tiques : a) les dialectes septentrionaux qui comprennent les dialectes gallo-italiens (piémontais, ligure, lombard, émillien-romagnol), le vénitien et l'istrien ; b) les dialectes toscans avec le florentin en position centrale ; c) les dialectes centro-méridionaux qui peuvent être subdivisés en trois grands sous-groupes : groupe des Marches, de l'Ombrie, de Rome ; groupe des Abruzzes, des Pouilles septentrionales, de Molise, Campanie et Lucanie ; groupe salentin, de la Calabre et de la Sicile. Les parlers corses sont, en général, considérés comme apparentés au toscan, bien que certains linguistes et érudits corses modernes parlent d'une langue corse. Au nord le piémontais et, à un moindre degré, le ligure offrent des traits de transition entre les parlers italiens et les parlers occitans.

Le groupe *sarde* occupe le territoire de la Sardaigne à l'exception de l'îlot catalan d'Alguer, et de deux îlots génois (Carlo-forte et Calasetta). Des quatre dialectes parlés en Sardaigne, deux groupes sont rattachés par certains romanistes non pas au groupe linguistique sarde mais au groupe italien et plus précisément au dialecte toscan ; ce sont au nord les parlers de la région de Sassari et de ses environs immédiats et au nord-est ceux de la région de Gallura : la morphologie, la syntaxe et le lexique de ces parlers sont de provenance italienne. Les documents anciens de ces deux régions sont rédigés en une langue de pur type sarde ; c'est à partir du XII^e siècle que le parler fut envahi par le toscan. Le parler de Gallura est de type corse méridional : il ne s'agit pas d'un trait ancien mais du résultat d'une émigration corse que l'on a fixé à partir de la fin du XVI^e siècle. Les deux principales variétés de dialectes proprement sardes sont le logoudorien dans le centre de l'île et le campidanien au sud. On

considère le logoudorien comme le sarde par excellence : c'est le parler employé sous la dénomination de *sardo illustre* en littérature.

Le groupe *occitano-roman* occupe une place centrale parmi les dialectes romans. On a pu parler à son propos de carrefour des langues romanes. On le divise en catalan, gascon et occitan proprement dits. Le catalan ancien différait assez peu de l'ancien occitan ; il en différait, en tout cas, beaucoup moins que le gascon. Le domaine catalan comprend : 1) en France, le Roussillon et la Cerdagne ; 2) la république d'Andorre ; 3) en Espagne, une bande du territoire de l'Aragon, le Principat ou Catalogne proprement dite (Barcelone, Gérone, Tarragone, Lérida), une grande partie des régions de Valence et d'Alicante, les îles Baléares ; 4) en Sardaigne le territoire de la ville d'Alguer où le catalan fut importé à la suite de l'annexion de cette région à la couronne d'Aragon en 1353. On distingue généralement deux zones de parlers catalans : une zone orientale et une zone occidentale. A la première appartiennent le catalan oriental proprement dit (Barcelone, Gérone, une partie des provinces de Tarragone et de Lérida), les parlers des Baléares, le roussillonnais, le parler d'Alguer ; à la seconde appartiennent le léridan (Andorre, Cerdagne espagnole, part des provinces de Lérida, Gérone, Huesca, Saragosse, Castellon de la Plana, Tarragone), le valencien (part des provinces de Castellon de la Plana, de Valence et d'Alicante). Le catalan normalisé par le grammairien Pompeu Fabra est la langue officielle de la république d'Andorre ; il fut également la langue officielle du gouvernement de la république, puis de la *Generalitat* de Catalogne jusqu'à l'établissement du régime de Franco en Espagne. Il est redevenu avec le nouveau régime

de Juan Carlos la langue officielle de la Catalogne redevenue autonome.

Le gascon se subdivise en de nombreux sous-dialectes dont certains sont très proches de l'occitan languedocien et dont d'autres accumulent les traits originaux proprement aquitains. Parmi ceux-ci le béarnais a une très forte personnalité ; il est resté langue écrite jusqu'au XVIII^e siècle et son emploi officiel n'a entièrement disparu qu'à la Révolution française. Le territoire du gascon recouvre l'ancienne Aquitaine dont César nous dit déjà qu'elle différait *par la langue*, les coutumes et les lois des autres parties de la Gaule et dont il fixait les limites à la Garonne : « *Gallos ab Aquitanis Garunna flumen dividit.* » Le territoire du gascon est en gros limité par l'océan Atlantique, les Pyrénées, la Garonne de son embouchure à Toulouse, puis par le cours de l'Ariège dont il s'éloigne progressivement vers l'ouest avant de joindre les Pyrénées. Il pénètre en Espagne avec le val d'Aran qui est de parler gascon.

L'occitan proprement dit continue la langue des troubadours du Moyen Age, appelée langue d'oc par Dante, limousin très souvent par les troubadours eux-mêmes, provençal par les romanistes du siècle dernier. Le territoire proprement dit de l'occitan jouxte au sud/sud-ouest, le catalan et le gascon. Au nord la limite actuelle part de la rive sud de l'estuaire de la Gironde, suit ce fleuve vers l'amont, puis remonte vers le nord en laissant Angoulême à l'ouest, englobe le Limousin, l'Auvergne, rejoint le franco-provençal à Châteldon, coupe le Rhône au-dessous de Valence, passe légèrement au-dessous de Grenoble (dans la vallée du Drac on parlait encore naguère un dialecte d'oc), rejoint les Alpes où les parlers d'oc empiètent légèrement sur le territoire politique italien, et se confond vers le sud

avec la frontière politique italo-française. Le romaniste J. Pignon a récemment montré que l'occitan était au XI^e et au début du XII^e siècle la langue d'une partie du Poitou, notamment la langue de Poitiers la capitale de Guilhem de Poitiers, duc d'Aquitaine, le premier troubadour connu. On divise l'occitan proprement dit en deux grandes zones dites zone nord-occitane et zone de l'occitan moyen. Le nord-occitan comprend le limousin qui occupe en gros le nord de la Dordogne, la Corrèze, la Haute-Vienne, la Creuse ; l'auvergnat (nord du Cantal, Puy-de-Dôme, Haute-Loire) et le provençal alpin (Ardèche, Drôme, Hautes-Alpes, nord des Basses-Alpes). L'occitan moyen comprend le provençal (Provence, comtat Venaissin, comté de Nice, parlers de Nîmes et d'Uzès sur la rive droite du Rhône) et le languedocien qui occupe ainsi une position centrale entre catalan, gascon, nord-occitan et provençal. L'occitan moyen (et notamment les dialectes languedociens, de Béziers-Rodez à Toulouse) constitue le groupe de parlers le plus conservateur et le moins éloigné de la langue classique du Moyen Age. Jusqu'au XV^e siècle, l'occitan fut utilisé comme langue juridique et administrative des pays d'oc. Il connaît depuis le XIX^e siècle une renaissance culturelle qui va s'amplifiant malgré le handicap que représente sa situation de langue minoritaire non reconnue, bien qu'une place lui ait été accordée, depuis peu, dans l'enseignement officiel.

Les dialectes *rhéto-romans*, ou roman alpin, occupent les vallées des Alpes comprises entre les Grisons et l'arc des Alpes Juliennes jusqu'à l'Adriatique y compris encore à la fin du XIX^e siècle, la ville de Trieste. Avant l'immigration germanique, ces parlers étaient autochtones dans toute la région des Préalpes (provinces de la Raetia et du Noricum

notamment) et orientés vers la Gaule. Ces parlers se divisent ordinairement en trois groupes isolés : 1) le groupe occidental en Suisse dans les Grisons, formé par les dialectes romanches dont les deux principales variétés sont le surselvan et l'engadinois ; 2) le groupe central des dialectes ladins dans les Dolomites (Italie) ; 3) le groupe oriental du frioulan avec Udine comme centre. En 1938 le romanche a été reconnu comme quatrième langue nationale de la Suisse et enseigné comme telle.

Les parlers *dalmates* attestés à Veglia (ou Krk) et dans les villes de Zadar, Split, Dubrovnik et Antivari sont aujourd'hui éteints : on précise que le dernier sujet vegliote parlant dalmate est mort en 1898. Le dalmate a été absorbé du côté roman par le vénitien, du côté slave par le croate.

On distingue habituellement quatre grands dialectes *roumains* : 1) le dialecte istro-roumain avec quelques milliers d'usagers en Istrie orientale, non loin de Fiume (ou Rijeka), en Yougoslavie ; 2) le mégléno-roumain avec quelques milliers d'usagers également, au nord de Salonique (Grèce) ; 3) le macédo-roumain ou aroumain parlé par 350 000 personnes environ, la plupart dans le nord de la Grèce, les autres en Yougoslavie, Bulgarie et Albanie ; 4) le daco-roumain qui est essentiellement le roumain du nord du Danube. Celui-ci est parlé par environ 20 millions de personnes, dont la grande majorité habite le territoire actuel de la Roumanie ; le reste se trouve dans la partie du Banat appartenant à la Yougoslavie et dans quelques villages bulgares et hongrois aux confins de la Roumanie. Par ailleurs plus de 3 millions d'usagers du daco-roumain habitent dans les territoires dépendant de l'U.R.S.S., principalement dans la république de Moldavie, mais aussi dans la république d'Ukraine.

Le daco-roumain s'étend ainsi jusqu'au-delà du Dniester. La république de Moldavie pour des motifs essentiellement politiques que ne justifie pas la linguistique a reconnu officiellement l'existence d'une langue moldave qui n'est réellement qu'une forme hybride du roumain littéraire avec un léger coloris dialectal moldave.

Les parlers *franco-provençaux* sont au contact des parlers italiens (Piémont), des parlers germaniques (Suisse alémanique), des parlers d'oïl et des parlers d'oc. Ils sont ou étaient parlés dans la plus grande partie du Forez, du Dauphiné et de la Savoie, dans la partie sud de la Franche-Comté, en France, ainsi que dans le val d'Aoste en Italie et dans la Suisse romande (canton de Neuchâtel, Fribourg, Vaud, Genève, Valais). On a proposé de diviser les parlers franco-provençaux en deux grandes zones : *a*) une zone septentrionale (cantons de Vaud, Valais et majeure partie des cantons de Fribourg et de Neuchâtel), zone assez isolée entre Jura et domaine germanique où s'accroissent les différences dialectales ; *b*) une zone méridionale où se trouvent de grands centres comme Genève, Grenoble et surtout Lyon qui ont permis au Moyen Âge le développement d'une *langue de cité* ainsi que la formation d'une *scripta* relativement unifiée sur la base du parler lyonnais.

Les dialectes *français* ou d'oïl (ancienne forme d'où provient le moderne adverbe d'affirmation *oui*) occupent la partie septentrionale du domaine gallo-roman. Ils occupent la partie nord de la France politique sauf : à l'ouest, la Bretagne bretonnante (Finistère, Côtes-du-Nord, Morbihan) ; au nord, l'extrémité de la Flandre (Dunkerque, Hazebrouck) ; à l'est, l'Alsace et la partie germanique de la Lorraine. Hors de France ils recouvrent la partie sud

de la Belgique (Wallonie), les îles anglo-normandes (Jersey et Guernesey) et, en Suisse, une grande partie du Jura bernois. Il est difficile de délimiter exactement les dialectes français. On en reconnaît ordinairement une quinzaine qui tirent en général leur nom des anciennes provinces françaises et qui peuvent se regrouper par grandes régions : au nord, le picard et une partie du normand ; au nord-ouest, le normand de basse Normandie, le gallot (Bretagne non bretonnante), le tourangeau, l'angevin ; au nord-est, le wallon ; à l'est, le lorrain ; au sud-est, le franc-comtois, le bourguignon, le bourbonnais ; au sud-ouest, le saintongeais, le poitevin, le berichon ; au centre, le champenois, le francien. A ces dialectes continentaux il faut ajouter l'anglo-normand importé par Guillaume le Conquérant, en Angleterre en 1066, qui jouit d'un grand prestige littéraire et administratif du XII^e au XV^e siècle.

Les vicissitudes de l'histoire n'ont pas accordé le même sort aux dialectes romans. Certains, qui méritaient la dénomination de langue par leur originalité linguistique, n'ont pas survécu comme le dalmate. D'autres, au contraire, « ont réussi » et sont devenus, par des cheminements divers, les actuelles langues romanes.

3. Les langues romanes. — Le *dalmate*, avons-nous dit, s'est éteint à la fin du XIX^e siècle, en ce qui concerne le vegliote. Le parler de Raguse (aujourd'hui Dubrovnik) avait déjà disparu vers la fin du XV^e siècle. Les dialectes dalmates des autres villes se sont éteints beaucoup plus tôt. Le dalmate ne nous a laissé que des documents d'archives en dehors des données de la toponymie et des vestiges lexicaux qu'il a laissés dans le vénitien et le croate.

Le sort du *franco-provençal* a été un peu meilleur.

Il subsiste encore dans les « patois » franco-provençaux. Au Moyen Age il nous a laissé une œuvre littéraire non négligeable : les œuvres de Marguerite d'Oingt, ainsi que le témoignage d'une langue de cité notamment à Grenoble, à Lyon et dans le Forez. Une littérature dialectale est née au XVI^e siècle bien vivante encore au XVIII^e siècle, comme en témoignent les poèmes en dialecte de Saint-Etienne publiés par G. Straka.

Le *sarde* n'a jamais été une grande langue de culture. Par contre son importance comme langue véhiculaire et administrative a été relativement importante du XI^e au XIV^e siècle, époque à partir de laquelle le catalan puis l'espagnol l'évinceront comme langue de l'administration en attendant d'être remplacé de nos jours par l'italien. Le sarde reste aujourd'hui la langue ethnique d'environ un million de personnes. A l'heure actuelle semble se dessiner pour le sarde une nouvelle consécration culturelle.

A côté de ces trois langues romanes au destin malheureux ou modeste, les autres langues romanes nous offrent, d'une part, les langues qui ont réussi : italien, espagnol, portugais, roumain et français et, d'autre part, deux groupes de langues au sort moins heureux : le groupe occitano-roman et le groupe rhéto-roman.

La *langue littéraire italienne* occupe une place à part parmi les langues romanes par la façon dont elle s'est formée. Grâce à la très grande vitalité toujours actuelle des dialectes italiens, on peut se rendre compte aisément que l'italien présente un caractère toscan, et plus précisément florentin, évident. Ce caractère est dû, on le sait, à l'importance de l'œuvre de Dante, de Boccace et de Pétrarque. La renommée littéraire de ces trois

écrivains a fait obstacle à toutes les tentatives pour promouvoir un autre dialecte que le florentin comme langue littéraire, tentatives qui ont eu lieu du ^{xiv}^e au ^{xvi}^e siècle. Une étape fondamentale a été marquée dans l'histoire du développement de l'italien littéraire lorsque l'humaniste P. Bembo publia en 1525 ses *Prose della volgar lingua* dans lesquelles il prenait comme modèle de la langue littéraire non pas le florentin parlé du ^{xvi}^e siècle, mais le pur florentin du ^{xiv}^e siècle qui avait été illustré par les trois grands : Dante, Boccace, Pétrarque. La position de Bembo fut largement acceptée. Elle triompha définitivement lorsqu'en 1612 l'*Accademia della Crusca*, fondée en 1582, choisit comme modèle du premier vocabulaire normatif publié par elle, la langue des écrivains du *Trecento*, surtout celle de Boccace. Ce triomphe empêcha naturellement la naissance d'une langue nationale italienne au contact vivant des dialectes au cours des ^{xvii}^e et ^{xviii}^e siècles, l'Italie étant alors politiquement et économiquement divisée. La question de la *fiorentinità* ou de l'*italianità* de l'italien littéraire longuement débattue surtout depuis le ^{xvi}^e siècle fut définitivement résolue au ^{xix}^e siècle avec l'unification politique de l'Italie. Elle le fut grâce surtout à l'œuvre d'A. Manzoni entre 1830 et 1840 qui réussit à revivifier l'italien à la source vivante du florentin parlé des Florentins cultivés, en donnant l'exemple dans son roman *I promessi sposi*. Ainsi l'italien littéraire reposant indubitablement sur le fond florentin devint la langue de l'unité italienne. Dans la capitale de l'Italie unifiée, à Rome, déjà dès le ^{xv}^e siècle, l'italien florentin avait pris le pas dans l'usage littéraire sur le dialecte *romanesco*. Ce dialecte même n'est plus que le dialecte des environs de Rome, un nou-

veau parler romain s'étant formé qui est surtout de caractère toscan. Rome depuis l'unification de l'Italie et surtout depuis la première guerre mondiale est ainsi devenue le grand centre d'expansion de l'italien littéraire, langue nationale à fondement essentiellement florentin.

La formation de l'*espagnol* est d'un caractère fort différent de celle de l'italien littéraire. Elle est due à l'expansion graduelle de la suprématie politique du dialecte de la Vieille-Castille, primitivement resserré, au nord de l'Espagne avec l'aragonais et le léonais, entre le catalan à l'est et le galicien à l'ouest. Dans la seconde moitié du ^{xii}^e siècle la Castille étend son influence grâce à l'affaiblissement des royaumes de Navarre et de Léon. Le castillan s'étend au ^{xiii}^e siècle aux dépens de l'aragonais à l'est et du léonais à l'ouest. Lorsque commence la *Reconquista*, menée sous la prépondérance de la Castille, le castillan s'étendra vers le sud jusqu'à Cadix, refoulant et absorbant les dialectes romans mozarabes du Sud qui s'étaient formés sous la domination arabe. Au ^{xiii}^e siècle l'expansion du castillan fut favorisée par le prestige croissant de la cour d'Alphonse le Sage dans toute l'Espagne, prestige dû notamment à la création de la prose littéraire espagnole, grâce à l'activité littéraire d'Alphonse le Sage et de sa cour dont la principale occupation intellectuelle était la traduction de livres hébreux ou arabes. Avec la prise de Grenade en 1492 qui mit fin à la domination arabe, la suprématie du castillan était assurée. Il devenait la langue littéraire de toute l'Espagne, aux dépens même du galicien et du catalan, malgré la vitalité culturelle de ces deux langues, du moins jusqu'à la Renaissance catalane du ^{xix}^e siècle qui a remis en cause la suprématie de l'espagnol en Catalogne.

La formation du *portugais* littéraire moderne bien qu'en partie conditionnée elle aussi par le développement de la *Reconquista*, offre néanmoins un caractère assez différent de celle de l'espagnol. Tout d'abord, le portugais qui prend son origine dans le gallego-portugais qui, à une époque archaïque, liait intimement les parlers portugais et les parlers galiciens, s'est particularisé précisément en rompant cette unité primitive. Avec l'indépendance politique, le portugais est devenu une langue romane tandis que le galicien, sans indépendance politique, est considéré comme un dialecte espagnol. Ensuite, ce n'est pas, comme dans le cas de l'espagnol, le dialecte du Nord de la primitive Lusitanie qui est devenu le portugais littéraire. Sans doute les parlers du Nord, au cours de la Reconquête, à la suite des victoires du roi dom Alfonso Henriques, ont-ils absorbé, peut-être en s'identifiant avec eux, les parlers romans dont, par ailleurs, on ne sait pas grand-chose, qui s'étaient formés durant la domination arabe du VIII^e au XII^e siècle et qui sont connus sous le nom de *romanzo moçarábico*. Mais, à partir des XV^e et XVI^e siècles, la langue littéraire portugaise offre des particularités, inconnues dans la province de Minho au nord, et qui par contre, sont caractéristiques des provinces du Centre (Beira) et du Sud (Estramadura). Le portugais n'est donc pas, comme c'est le cas pour l'espagnol, l'italien et le français, le résultat de la prédominance d'un seul dialecte. Les dialectes des diverses provinces portugaises ont leur part dans sa formation, notamment celles où se trouvaient Coimbra et Lisbonne, au centre et au sud.

La formation du *français* est le résultat de la prééminence accordée à l'ancien dialecte francien de l'Ile-de-France et plus précisément au parler de

Paris. Cette prééminence ne fut définitivement acquise, d'ailleurs, que relativement tard. Bien que, dès 1182, le poète artésien Conon de Béthune ait été raillé à la cour de Paris pour son accent picard, pendant un assez long temps le francien fut concurrencé par le normand (qui pouvait se prévaloir du support de la culture anglo-normande de Grande-Bretagne) et surtout, du point de vue littéraire, par le picard. La *scripta* picarde fut illustrée par une riche littérature bourgeoise. Froissart (1333-1410) écrira encore en picard au XIV^e siècle. L'action politique de la royauté française devait faire triompher le parler du roi, soutenu, par ailleurs, de bonne heure, par l'action poétique et religieuse de l'abbaye de Saint-Denis, « cœur idéal de la monarchie ». Quelle qu'en soit l'origine, la *scripta* francienne s'imposa au cours du XIV^e siècle et l'on peut dire qu'au XV^e siècle elle est devenue la langue littéraire française. L'édit de Villers-Cotterêts n'a probablement pas eu grande influence en ce qui concerne l'existence des « langues » écrites régionales d'oïl, mais il eut une grande importance en imposant le français comme langue de l'administration dans tout le royaume français auquel le développement de l'imprimerie imposa le dogme de l'unité de la langue, fortement soutenu par la politique centralisatrice de la royauté. Le français du XVI^e siècle cependant est encore une langue qui jouit d'une très grande liberté et dans la phonétique et dans la syntaxe. L'unification ne sera définitivement acquise qu'au XVII^e siècle avec le triomphe de la notion du « bon usage ».

Le *roumain* littéraire est apparu à une date relativement tardive qu'explique l'histoire des populations de parlers roumains. Le premier texte de datation certaine est une lettre d'un boyard de Valachie

écrite en 1521. A partir du milieu du xvi^e siècle, l'essor de la langue littéraire doit beaucoup à l'activité du diacre Coresi et de ses disciples. Dans la seconde moitié du xvii^e siècle et au début du xviii^e, la production des textes, surtout religieux, mais aussi profanes, dont la langue repose sur le daco-roumain, connaîtra une très grande vitalité. C'est, d'ailleurs, à cette époque, que le slavon cesse d'être la langue officielle des populations roumaines. Cette époque peut être considérée comme le siècle d'or de la culture roumaine ancienne en même temps que le moment classique du vieux roumain. C'est au xix^e siècle que se forme définitivement le roumain moderne notamment dans la deuxième moitié de ce siècle, par l'action conjuguée de la renaissance de la nation roumaine sous l'influence des idées de la Révolution française et de la prise de conscience des origines latines du peuple roumain, favorisée par le catholicisme des roumains de Transylvanie. Cette formation doit énormément à l'action d'un certain nombre de grands Roumains, patriotes et idéalistes, qui ont pris conscience du retard de leur civilisation par rapport à l'Occident et qui veulent ardemment combler ce retard. Restant solidement appuyés sur l'apport folklorique et sur la langue des vieilles chroniques roumaines, ces hommes ont recours à l'italien, au français, à toutes les ressources de la culture occidentale. Contrebalançant les effets du cosmopolitisme par la roumanisation due aux sources populaires, le roumain moderne est devenu la belle langue nationale que nous connaissons.

L'*occitano-roman* et le *rhéto-roman* connaissent des destinées moins heureuses. Parmi les dialectes rhéto-romans le *frioulan* parlé par environ 600 000 personnes est l'instrument d'une littérature

de valeur, particulièrement en poésie. Le *romanche* dont le plus ancien texte littéraire provenant du pays des Grisons remonte au xii^e siècle a été reconnu nous l'avons dit, comme quatrième langue nationale de la Suisse en 1938. Mais il convient de noter qu'il n'est point comme l'allemand, le français et l'italien, langue officielle, ce qui constitue pour lui un net désavantage. Le romanche a pénétré dans l'enseignement primaire et secondaire et il est le moyen d'une littérature non négligeable. Malheureusement, le romanche n'a pu parvenir à se constituer en langue unitaire et l'on enseigne le surselvan dans la région catholique des Grisons et l'engadinois dans la zone protestante.

En ce qui concerne l'*occitano-roman*, il faut mettre à part le *catalan*. Jusqu'à la fin du xiii^e siècle, et plus tard encore (Auzias March et Jordi de Sant Jordi sont du xve siècle) d'une façon moins classique, les poètes catalans écrivent dans la langue des troubadours proprement occitans. Le souvenir de cette communauté sera encore vivant au xix^e siècle où le catalan est souvent appelé *llenga llemosina*. C'est Ramon Llull (1235-1315) qui écrira dans un catalan, encore fortement occitan, les premières œuvres poétiques catalanes. C'est également lui qui crée la prose philosophique et narrative catalane. Au xiv^e siècle le catalan est la langue officielle de la maison royale d'Aragon et c'est le catalan de la chancellerie royale qui est à la base du catalan littéraire médiéval, dans lequel sont écrites les diverses chroniques catalanes. Avec l'union à la Castille en 1479, le catalan s'efface devant l'espagnol et se trouve réduit à l'état de dialecte sinon de « patois ». Au xix^e siècle, la langue catalane connaît une puissante renaissance appuyée sur le rayonnement économique de Barcelone. Le catalan fut unifié grâce à

l'œuvre de Pompeu Fabra. En 1931 le catalan devint la langue officielle de la République catalane proclamée par Macià et resta la langue officielle du gouvernement autonome de la *Generalitat de Catalunya* jusqu'à la prise du pouvoir en Espagne par Franco en 1939. Malgré les vicissitudes politiques le catalan aujourd'hui fait preuve d'une très grande vitalité, non seulement comme langue d'une très abondante et très riche littérature, mais comme langue largement véhiculaire par la volonté même du peuple catalan.

L'*occitan*, proprement dit, a connu la destinée la plus mouvementée de toutes les langues romanes littéraires. Il fut, au Moyen Age, incontestablement une très grande langue de civilisation, sinon la plus importante langue de civilisation par l'influence de sa poésie lyrique et des valeurs éthiques que cette poésie répandit, aussi bien vers le sud, dans la péninsule Ibérique, en Italie et en Sicile, que vers le nord, en Grande-Bretagne et dans les pays germaniques. La *scripta* occitane générale, déjà formée dès les premiers textes administratifs du milieu du XI^e siècle (cf., par exemple, les serments de fidélité faits à Guilhem IV seigneur de Montpellier), non seulement fut utilisée comme langue littéraire, mais fut d'emblée utilisée comme langue du droit et des sciences, et resta *langue de cité*, juridique et administrative jusqu'à la fin du XV^e siècle. La langue de cité résista longtemps au prestige de la langue du roi de France qui commence à se manifester dès le rattachement des terres du comté de Languedoc au royaume de France à la suite de la guerre contre les Albigeois. Le troubadour Bernard Sicart de Marvejols, contemporain de la Croisade contre les Albigeois, s'irrite déjà « d'entendre des gens courtois saluer humblement les Français du nom de *Sire* »

(au lieu de l'occitan *Senher*). Au moment de l'édit de Villers-Cotterêts, le prestige du français était déjà bien établi en face de l'occitan survivant. Du point de vue littéraire, la langue d'oc connaît au XVI^e siècle une première renaissance, mais socialement parlant, bien que toujours très vivante populairement, elle se trouve rabaissée à l'état de « patois ». La Révolution et surtout les guerres de l'Empire contribuent à la misère sociologique de l'occitan. Le Romantisme et les idéaux populaires de 48 introduisent dans ce processus de « misération » l'idée de la dignité du langage du peuple. Le XIX^e siècle connaît alors la renaissance des parlers d'oc avec Mistral et le Félibrige. Au XX^e siècle cette renaissance se poursuit avec la Société d'Etudes occitanes (continué par l'Institut d'Etudes occitanes) et l'œuvre du grammairien Louis Alibert. Une place très mesurée est accordée à la langue d'oc dans l'enseignement par le gouvernement central. L'occitan moderne connaît une nouvelle consécration culturelle (et même véhiculaire) et prétend à la dignité d'une culture ethnique consciente, malgré la querelle momentanée du « mistralisme » linguistique.

II. — Typologie

Il n'est pas question de relever ici tous les phénomènes qui caractérisent les langues romanes actuelles. On se contentera de signaler quelques grands traits généraux par quoi diffèrent ou se rapprochent nos principales langues.

1. Les sons. — En ce qui concerne le *vocalisme*, les voyelles finales du latin sauf *a*, et mis à part quelques cas particuliers, ont disparu en français, occitano-roman et rhéto-roman.

Le vocalisme roumain se distingue de celui des autres langues romanes par la présence d'un *i* dit guttural. Le français et l'occitan (mais non le catalan) se distinguent, eux, par la présence d'un *ii* palatal qui a remplacé le *u* latin (ce *ii* se retrouve également dans les parlers gallo-italiques de l'Italie du Nord).

La phonologie du latin présentait un système assez simple comprenant les cinq voyelles *i, e, a, o, u* qui pouvaient être soit longues, soit brèves. Du latin au roman primitif s'est créé un système nouveau qui a remplacé la quantité par la qualité phonologique. Ce système comprenait sept voyelles *i, e* fermé, *e* ouvert, *a, o* ouvert, *o* fermé, *u*. Ce système est commun à la plupart des langues romanes à l'exception du roumain, des parlers de l'Italie méridionale et du sarde.

Un des faits qui ont le plus contribué à changer le visage de ce système est le phénomène de la diphtongaison, qu'il s'agisse de diphtongaison primitive ou de date secondaire. L'italien, l'occitano-roman, l'espagnol, le portugais ont été beaucoup moins sensibles que le français et le roumain aux phénomènes de diphtongaison. L'italien connaît la diphtongaison romane en syllabe libre seulement : *piede, buono* en face de *porta*. L'occitano-roman ne connaît pas la diphtongaison romane spontanée (sauf localement pour *o* ouvert : *couor/couar*), mais seulement la diphtongaison conditionnée : *lieit, nueit, uelh, buou* (le catalan se distingue ici par l'aboutissement différent de la diphtongaison de *e, o* ouvert devant yod : *lit, nit, ull* et par la non-diphtongaison conditionnée par (w) : *bou*). L'espagnol offre la diphtongaison romane en syllabe ouverte et fermée, mais ne diphtongue pas devant un élément palatal : *pie, bueno, puerta, tierra*, mais

ojo. Le portugais est caractérisé par l'absence de diphtongaison spontanée ou conditionnée : *pé, bom, noite*. Le français a largement connu la diphtongaison à date ancienne en diverses conditions avant de ramener ses diverses diphtongues (ou triphth.) à la monophtongaison (ou aux groupes *wa, wi, wê*) ; la diphtongaison de *o* (ouvert ou fermé) libre aboutit à *œ* ouvert ou fermé (meule < *mōla*, neuf < *nōvu*, neveu < *nepōte*, fleur < *flōre*) ; *e* ouvert aboutit à *yè* ou *yé* (fiel < *fēl*, pied < *pēde*) ; *e* fermé passe à *ei* puis à *oi* et à *wa* (toile < *tēla*) ; la diphtongaison du *a* aboutit à *ê/è* (mer < *mare*, aimer < *amare*). Caractéristiques du français également sont les diphtongaisons (ou triphth.) des types : *main* (manu), *doyen* (decanu), *paix* (pace), *voix* (voce), *voisin* (vecinu), *puits* (puteu), *saint* (sanctu), *coin* (cuneu), *juin* (juniu). La réduction de la plupart des anciennes diphtongues fait du français moderne une langue pauvre en émissions polyphthongues. Le roumain, au contraire, notamment par suite d'un important phénomène de réfraction vocalique pour *e, ie, o* accentués possède vingt et une diphtongues et huit triphthongues, et on en relève un nombre plus important dans la langue populaire.

Le traitement de la diphtongue latine *au*, la seule diphtongue du latin parlé, distingue le roumain et l'occitan proprement dit, qui conservent la diphtongue, du français, du catalan, de l'espagnol et de l'italien où elle aboutit à *o* (occ. *causa* ; roum. *cauză* en face d'it., esp., cat. *cosa* ; fr. *chose*). En portugais *au* est restée à un stade intermédiaire ou : *cousa*. La diphtongue s'est conservée également à l'est et à l'ouest de la Rhétie, au sud de l'Italie et en Sicile. En sarde elle s'est simplifiée en *a* : *casa*. À l'initiale elle devient *u* en italien (lat. *audire* > *udire*). En français *au* et *eau* de formation

secondaire (*autre, chevaux, beau, etc.*) sont également devenus *o* au XVI^e siècle.

Un trait important qui distingue le caractère pris par les langues romanes est le traitement subi par le vocalisme proto-roman. L'italien conserve ce vocalisme et ses oppositions phonologiques sont toujours pertinentes. En portugais est également pertinente l'aperture de *e* et de *o* (*sède* : siège/*séde* : soif ; *còrte* : tranchant/*córtte* : cour). En espagnol *é* et *è*, *ó* et *ò* romans ont été confondus en (*e*) et (*o*) moyens : l'aperture de ces voyelles n'est donc pas phonologiquement pertinente. En occitan, en dehors des cas particuliers, l'aperture de (*e*) est pertinente ; quant au (*o*), *ó* fermé roman est devenu *u* en occitan proprement dit. En français l'état roman en question a été bouleversé par divers phénomènes et notamment par la diphtongaison généralisée des voyelles toniques libres. En roumain, la distinction de timbre existe mais elle offre un caractère phonétique (variante libre, variante combinatoire) et non phonologique.

Certains traits concernent plus spécialement certaines de nos langues. La nasalisation ne se trouve qu'en français et en portugais ; une semi-nasalisation est normale dans un certain nombre de parlers occitans. L'inflexion de *ay* en *ey* (qui peut se réduire ensuite à *è*) se rencontre en catalan, espagnol, occitan occidental, français. La neutralisation et la fermeture des voyelles atones se trouvent en portugais et en occitan. La neutralisation vocalique est connue du français, du portugais et du catalan ; elle ne se réalise qu'en position atone, soit en finale, soit en contrefinale, soit en prétonique. Ce (*e*) neutre, contrairement à ce qui se passe en portugais et en catalan, se labialise en français en ce que nous appelons *e* muet.

Le traitement du *consonantisme* révèle une égale complexité tout en laissant également apparaître certaines lignes générales. En ce qui concerne les consonnes initiales, le roumain et l'italien sont conservateurs. En italien les consonnes initiales demeurent sauf : 1) *k* palatal et yod (*j, g* latins) qui deviennent *tš, dž* (*città, genero*) ; 2) les groupes *kl, gl, pl, bl, fl* avec *l* palatalisé qui réduisent ce *l* à *y* (*chiave, ghianda, pieno, biada, fiamma*). En roumain les consonnes initiales demeurent également en dehors des cas de palatalisation : *k* palatal et *g* palatal aboutissent comme en italien à *tš, dž* ; dans les groupes qui offrent le second élément *l*, seuls *kl* et *gl* ont palatalisé *l* (*cheie, ghindă* en face de *plin, blestemă, floare*) ; *l* après s'être mouillé devant *i* ou combiné avec *ie* disparaît *iertă* (libertaire), *iepure* (lèporem). Contrairement à l'italien qui garde l'élément bilabial des groupes *kw, gw*, le roumain perd cet élément : la gutturale reste devant *a* (*quam > ca*) mais se palatalise dans **quero* (l. cl. *quaero*) > *cer*, **sanguem > sînge*.

Le traitement des consonnes initiales est beaucoup plus simple en portugais qu'en espagnol. *K* palatal (*c + e, i*) aboutit à *s* en portugais et à *th* (anglais *thin*) en espagnol. Le *y* (*i* consonne, ou *g + e*) aboutit en portugais à *ž* dans tous les cas (écrit *j, g*) tandis qu'on a en espagnol tantôt *y* (*yace, yerno*), tantôt le son appelé *jota* (*joven, juego*), tantôt une aspiration effacée par la suite (*hermano, helar*). Les groupes *kl, pl, fl* avec *l* palatalisé aboutissent à *l* mouillé en espagnol (*llave, lleno, llama*) mais à *š* en portugais (*chave, cheio, chamma*). Ceci ne s'est produit, d'ailleurs, que devant l'accent ; l'analogie explique les exceptions. Le portugais comme l'espagnol conservent les groupes *kw, gw* devant *a* tandis que devant les

autres voyelles la prononciation efface le *w* ; même devant *a* l'élément labial a tendance à disparaître dans la prononciation populaire du portugais. Le *v* et le *b* sont confondus dans la prononciation en espagnol, mais distincts en portugais. Les traits suivants sont spéciaux à l'espagnol : *s*- dans les mots empruntés aux Morisques aboutit à la jota : *jabon* (saponem) ; le passage de *s* à *th* (angl. *thin*) est dû au « ceceo » des Andalous : *zahorra* (saburra) ; enfin le passage de *f* à *h* (*hijo*, *harina*) qui se retrouve au-delà des Pyrénées dans le gascon est sans doute dû à des influences euskariennes.

L'occitan et le français palatalisent *c* + *e*, *i* et le résultat commun est *s*. De même *y* et *g* + *e*, *i* aboutissent dans les deux langues à *ž*, mais l'occitan ne palatalise pas *k*, *g* devant *a* (sauf en nord-occitan) à l'encontre du français (*cat*/*chat* ; *gamba*/*jambe*). *Kw*, *gw* perdent leur élément bilabial en français comme en occitan (sauf pour celui-ci en gascon et en catalan). Les groupes *cl*-, *gl*-, *pl*-, etc., sont conservés dans les deux langues. Enfin le français se distingue par l'existence du *h* dit aspiré dans les mots d'origine germanique étendu analogiquement à certains autres mots (*haut*, par ex.). L'aspiration disparue dès le xvi^e siècle se survit dans le fait que le *h* aspiré empêche l'élision (*la hache* contre *l'apia* en occitan).

En ce qui regarde les *consonnes intervocaliques*, le roumain ne connaît pour les occlusives aucun affaiblissement, qu'elles fussent sourdes (*foc*, *muta*, *ripa*) ou sonores (*cruda*, *lega*) ; seule la fricative *v*, représentant *v* ou *b* latins, s'est effacée entre voyelles : *cal* (caballum) ; *la* (lavare). En italien les sourdes intervocaliques restent dans les paroxytons (*fuoco*, *sete*), tendent à s'affaiblir, mais la tendance n'est pas générale, dans les proparoxytons

(*segale*, *redina*, *povero*) ; devant l'accent *c* s'affaiblit en *g* en principe (*pagare*), mais *t* et *p* restent intacts (*maturò*, *nipote*) (les formes comme *padella*, *arrivare* sont des formes du Nord). Quant aux sonores, elles restent intactes (*legare*, *nuda*, *cavallo*) ; toutefois *g* tend à s'effacer devant l'accent (*reale*, *fraore*).

L'espagnol, le portugais et l'occitano-roman se comportent à peu près de la même façon pour les occlusives intervocaliques. Les sourdes deviennent sonores (*pagar*, *mudar* dans les trois langues ; également *riba* en esp. et occ.). Les sonores se maintiennent ou s'affaiblissent : *g* se maintient après l'accent (esp. *llaga*, port. *praga*, occ. *plaga*) mais peut tomber devant voyelle accentuée (esp., port., occ. *liar* ; aussi *ligar* en occ.). Le *d* a eu un son fricatif, puis est passé à *z* en occitan ; en espagnol le *d* fricatif persiste parfois après l'accent (esp. *vado*, *nuda*) mais tend à s'effacer (esp. *prea*, *feo*, *creo*) ; il tombe toujours dans les proparoxytons ; devant l'accent on a aujourd'hui *caer*, *oir*. Le portugais maintient, comme à l'initiale, la distinction entre *b* et *v* tandis que l'espagnol et l'occitano-roman en partie (catalan, gascon et languedocien) confondent les deux phonèmes : la distinction orthographique dans ces langues est la conséquence de préoccupations étymologiques.

Le français est la langue qui a le plus altéré les occlusives intervocaliques. Les sourdes passent aux sonores et sont ensuite traitées comme elles : le latin *pacare* devenu **pagare* aboutit à *payer* comme *plaga* aboutit à *plaie* ; l'effacement de *c* est total devant voyelle vélaire (sûr < securum) ; *t* et *d* s'amuissent également (*muer* de *mutare* comme *suer* de *sudare*) ; *p* et *b* aboutissent également à la fricative sonore labio-dentale *v* : on a *trouver* de *tropare* comme on a *couver* de *cubare* et *laver* de *lavare*.

Le *s* intervocalique est resté dur en roumain. En italien il y a hésitation dans la prononciation du toscan littéraire (sourde dans *casa, cosa, naso, asino, arnese, difesa, posi, -oso*, etc. ; sonorisé dans *rosa, tesoro, bisogno, chiesa, usura, cortese*, etc.). La prononciation sonore est celle de l'Italie du Nord, la prononciation sourde celle du Centre et du Sud. En français, en occitano-roman, en portugais il y a eu sonorisation. Il en avait été de même en espagnol ancien, mais au *xvii*^e siècle le *-s-* est redevenu sourd en espagnol moderne.

Un trait particulier au roumain est celui qui transforme, à l'intervocalique et plus généralement à l'intérieur des mots, les groupes *kw, gw* en *p* (*apă* < *aqua*, *iapă* < *equa*, *limbă* < *lingua*). On retrouve un traitement semblable en sarde où latin *lingua* donne *limba* et *aqua* : *abba*. En français et en occitan, à l'intervocalique comme à l'initiale l'élément bilabial disparaît. Pour l'espagnol, le portugais et l'italien, le traitement à l'intervocalique est également le même qu'à l'initiale : conservation (en principe) de l'élément bilabial devant *a* et perte de cet élément devant les autres voyelles en espagnol, portugais ; conservation de l'élément bilabial en italien non seulement devant *a* mais aussi devant les autres voyelles.

Le traitement des liquides intervocaliques permet de souligner également quelques traits particuliers à chaque langue. En français, en occitan, elles sont conservées en général sauf en gascon où le *n* s'efface (*lua* : *luna*). L'espagnol a conservé également les liquides intervocaliques (sauf dialectalement *luna* donne *lua* comme en gascon, en Andalousie) tandis que le portugais connaît l'effacement de *-l-* et de *-n-* : *ceo* (*caelum*), *voar* (*volare*), *cor* de *coor* (*colorem*), *ter* (*tenere*), *femea* (*femina*). L'italien littéraire

conserve également les liquides intervocaliques. Dialectalement *-l-* passe à *-r-* en Italie du Nord ; ce traitement que l'on rencontre également en occitan dans la région du Lozère par exemple est caractéristique du roumain : *soare* (*solem*).

En ce qui concerne les consonnes doubles, l'italien se distingue des autres langues romanes par la conservation stricte des deux consonnes du latin, surtout après l'accent. Pour les groupes intérieurs de consonnes, caractéristique est le traitement du groupe *ct* : en italien *ct* s'assimile en *tt*, en roumain *ct* passe à *pt*, ailleurs on a *yt* qui reste en portugais, dans une partie de l'occitano-roman, en français (où la diphtongue produite subit des transformations secondaires) et qui évolue jusqu'à *tš* en espagnol et dans une partie de l'occitan : it. *fatto, notte* ; roum. *fapt, noapte* ; port. *feito, noite* ; fr. *fait, nuit* ; occ. *fait/feit, nueit/neit* et *fach, nuech* ; esp. *hecho, noche*. Le groupe *cs* subit un traitement à peu près symétrique : it. *tasso* (*taxum*) (toutefois devant l'accent on a *mascella* de **maxella*) ; roum. *coapsă* (*coxa*) (devant l'accent *măseă* de **maxella*) ; port. *teixo* ; fr. *cuisse* (*coxa*) ; occ. *cueissa* ; esp. *tejo*.

La séquence d'un yod *a*, dans toutes les langues romanes, relativement altéré les consonnes intervocaliques. *N + y* reste partout sauf en roumain où il se réduit à *y* : lat. *vinea* ; it. *vigna* ; fr. *vigne* ; occ. *vinha* ; port. *vinha* ; esp. *viña*, mais roum. *vie*. *L + y* demeure en italien, en occitan (sauf à date récente dans certains parlers) et en portugais ; se réduit à *y* en roumain et en français moderne ; aboutit à la *jota* en espagnol : lat. pop. *folia* ; it. *foglia* ; occ. *fuelha* ; port. *folha* ; esp. *hoja* ; mais fr. *feuille* (anciennement *feulye* ; définitivement depuis la Révolution *feuye*) ; roum. *foaie*. *R + y* persiste seulement en roumain, laisse tomber le *r* en italien,

transpose le *y* dans les autres langues : lat. *area* ; roum. *arie* ; ital. *aia* ; fr. *aire* ; occ. *aira* ; port. *eira* ; esp. *era* (avec monophthongaison secondaire de *ei* à *e* générale en castillan).

Les labiales suivies de *yod*, en italien conservent le groupe labiale plus *yod* en redoublant même la labiale : lat. *sapiat*, *rubeum*, *cavea*, *vindemia* donnent *sappia*, *robbio*, *gabbia*, *vendemmia*. En roumain, il y a transposition du *y* pour *by* : *aibă* (habeat), *roib* (robeum). La transposition en portugais concerne *py*, *by*, *vy* : *saiba*, *ruivo*, anc. port. *gaiva*. En espagnol seul le *y* de *py* est transposé : *sepa* mais *rubio*, *gavia*, *vendimia*. En occitan pas de transposition mais on a la consonnification du *y* dialectalement pour *py* : *sapia* et *sacha*, mais *gabia*, *vendemia*. En français il y a toujours consonnification du *y* en *š*, *ž* avec disparition de la labiale : *sache*, *cage*, *vendange*.

T + y entre voyelles en français seulement dégage un *yod* en avant : *puteus* donne *puits* en fr., mais it. *pozzo*, esp. *pozo*, port. *poço*, occ. *potz*, roum. *puļ*. Devant l'accent en italien *t + y* aboutit à *dž*, et à *tš* en roumain devant *o*, *u* accentués : roum. *tăciune* (titionem), it. *ragione* (rationem). L'italien est seul à offrir une affriquée prépalatale continuant *k + y* : *brac(h)iu* *y* donne *braccio* pour *brazo* (esp.), *bras* (fr.), *braç* (occ.), *braļ* (roum.).

En ce qui concerne la *syllabe finale*, on notera que seul l'italien n'admet aucune consonne à la finale. L'espagnol, le portugais et l'italien conservent sous forme de *-o* le *-u(m)* final latin ; l'occitan et le roumain en général présentent la consonne finale après chute de cette voyelle ; le français offre la forme la plus usée en amuissant la consonne devenue finale, sauf réfections.

2. Morphologie. Syntaxe. — En morphologie et en syntaxe on citera simplement quelques traits communs à tel groupe de langues et quelques traits spécifiques d'une seule langue. L'espagnol, le portugais et l'occitano-roman ont en commun les traits suivants : imparfait de l'indicatif *-abam* pour conjugaison I, *-eam/-iam* pour II et III ; emploi de l'article défini comme démonstratif ; impératif négatif rendu par le subjonctif tandis que le français et le roumain se servent de l'indicatif comme pour l'impératif positif. Dans les cas suivants pour l'occitano-roman, offrent le même usage que l'espagnol et le portugais seulement le catalan, le gascon et une partie du languedocien : introduction de l'objet direct par *a* ; absence de morphème de partitif ; plus-que-parfait en *-ra* à sens conditionnel. L'espagnol, le portugais et le gascon/catalan offrent la disparition de la désinence *-(i)unt* (*venden*). L'espagnol et le portugais connaissent en commun la disparition des infinitifs proparoxytons (*perder*). L'espagnol et une partie de l'occitano-roman (catalan, gascon, languedocien en partie) emploient un subjonctif présent à valeur de futur. L'italien et une partie de l'occitan (gascon) ont des imparfaits en *-ava*, *-eva*, *-iva*. L'italien, l'occitan et le français ont maintenu un pronom *en* (*nde* en italien ancien et dialectal, *ne* dans l'italien standard moderne). L'occitan et le français ont maintenu de plus un pronom *y* (*i*). L'italien et le roumain ne connaissent pas la formation du pluriel par *s* des autres langues romanes : en italien et en roumain en général le pluriel masculin est en *-i* et le pluriel féminin en *-e* (en *-i* aussi en roumain). De plus ces deux langues offrent une flexion verbale en *-i* à la 2^e personne sg. dans tous les verbes.

Certains faits caractérisent une seule langue. Le

portugais a développé un infinitif personnel (tu es parti *sans me voir* : *sem me veres*). L'italien a considérablement développé l'emploi du verbe réfléchi pour rendre l'aspect passif ; il utilise l'infinitif pour l'ordre négatif au singulier. L'occitan a donné une large place aux prétérits à infixes -g-. Mais c'est le roumain qui présente les traits spécifiques les plus originaux par rapport à l'ensemble des autres langues romanes : il pratique l'enclise de l'article (loup : *lup*, le loup : *lupul*) ; il a conservé une certaine flexion (nominatif-accusatif *casă*, gén. datif *case* ; nom. acc. *lupul*, gén. dat. *lupului*) ; il a également maintenu le genre neutre ; il remplace l'infinitif par le subjonctif (*scio scribere* : *stiu sa scriu*), etc. Le français moderne également s'éloigne des autres langues romanes, au moins autant que le roumain, par des traits différents mais non moins importants du point de vue systématique. On citera seulement à titre d'exemples dont l'importance est indéniable le fait que la suffixation est ouverte dans toutes les langues romanes sauf en français ; la déchéance de la flexion verbale qui a amené la création d'un pronom personnel morphème, et plus généralement la simplification des marques (comparez, par ex., fr. *les lois étaient justes*, qui ne se distingue plus que par une seule marque, du singulier (*la*) *loi était juste* avec it. *le leggi erano giuste/la legge era giusta*, roum. *legile erau juste/legea era justă* ou occ. *las leis eran justas/la lei era justa*).

3. Le lexique. — En ce qui concerne le *lexique*, les langues romanes offrent le même mélange de faits semblables chez toutes et de faits plus ou moins spécifiques à chacune. Elles possèdent un fond commun latin mais avec des inflexions locales : l'italien a davantage retenu du lexique du latin clas-

sique ; l'espagnol et le portugais ont conservé un certain nombre de types lexicaux classiques disparus ailleurs même en Italie ; le français et l'occitan ont conservé à peu près le même fond latin, le français étant cependant plus novateur que l'occitan ; quant au roumain le fond latin de la langue générale est bien le même que celui du reste de la Romania mais il présente certains traits particuliers qu'il partage avec l'italien. Il a conservé seul quelques lexèmes latins.

Les substrats, superstrats, adstrats colorent différemment le lexique des langues romanes. Le substrat préroman du français est différent du substrat espagnol, portugais et occitan. Le roumain possède également un substrat local très ancien. Le substrat celtique est, en général, commun à toutes les langues romanes, excepté le roumain qui, par contre, se caractérise par la présence d'un superstrat slave et d'un adstrat hongrois. L'espagnol, le portugais et, en partie, l'occitan connaissent un substrat lexical « basque ». Les emprunts au germanique se retrouvent surtout dans les langues romanes de l'Ouest, mais ils sont plus nombreux en français : l'occitan ne connaît pas, par exemple, des mots de grande fréquence comme *haïr* du germ. *hatjan* ou des mots concrets d'usage familial comme *saule* (*salha), *hêtre* (*haistr) qui sont en occ. *asirar*, *sausage*, *fau* (lat. ad-irare, salicem, fagum). Les arabismes sont surtout nombreux en espagnol, moins fréquents en portugais, mais on en retrouve dans les autres langues romanes sauf en roumain (en dehors des mots empruntés au français) : ce sont des termes de civilisation se rapportant aux sciences (alambic, algèbre, zénith, chiffre, etc.). Ces termes se retrouvent aussi bien en français qu'en italien comme en espagnol et en portugais tandis que les termes de civilisation se

rapportant à l'administration et aux techniques (*alcade, albañal*) sont localisés dans la péninsule Ibérique. Le roumain connaît un certain apport grec et turc.

Les emprunts sont nombreux en divers sens. L'italien a réintroduit de bonne heure un fort contingent de mots latins qui ne se distinguent pas de la masse lexicale par suite du conservatisme phonétique de la langue. Il en va de même pour l'occitan dont les formes du lexique savant de la langue anciennes s'adaptent directement à la phonétique. Par contre, en français, les emprunts savants forment une masse bien distincte, créant souvent des doublets avec les formes populaires. Le français moderne se caractérise par la perte d'un grand nombre de termes du lexique populaire du fond primitif en face d'emprunts sans nombre soit au latin du Moyen Age soit au latin classique. Nombreux également sont les emprunts d'une langue romane à l'autre : l'italien a beaucoup emprunté à ses propres dialectes de même que l'espagnol et le portugais ont largement fait appel aux divers parlers ibériques, tout comme le roumain a fait appel à ses parlers populaires. De plus, l'italien a emprunté des mots de Cour à l'espagnol et des termes divers à l'occitan et au français en grande quantité. L'espagnol a accepté de nombreux gallicismes dès le Moyen Age, des italianismes au XVI^e siècle, tout comme le français, et naturellement un certain nombre de termes de l'hispano-américain. Le roumain a non seulement puisé dans le trésor de ses parlers populaires, mais il a, en même temps, largement subi l'influence française.

Il convient de noter, enfin, que le lexique français, plus que celui de toute autre langue romane, se caractérise par une lexicalisation très poussée qui, avec la restriction de la suffixation, limite

l'existence des « mots possibles » (en occitan sur un radical *carga* : charge, on peut former cinq fois plus de dérivés et composés qu'on ne peut le faire avec le français *charge*). D'autres traits du lexique français sont la faiblesse de la motivation entre signifiant et signifié d'où bris des familles de mots (*eau/aqueux* en face d'occ. *aiga/aigós*), l'abondance des homonymes, la faiblesse de l'autonomie du mot (d'où grandes possibilités de calembours et jeux de mots). Ces traits rendent le lexique du français beaucoup plus solidaire de la syntaxe et de la structure générale de la langue que n'importe quel autre lexique des langues romanes.

Faut-il conclure après ce rapide coup d'œil sur quelques traits de nos langues romanes ? En dehors d'un essai de classement que ces traits suggèrent (groupement italo-roumain à l'est, groupement ibéro-occitan tempéré par un groupement gallo-occitan à l'ouest), classement qui n'a pas, d'ailleurs, obtenu le *consensus omnium* des romanistes, on peut simplement dire que les langues romanes sont animées de diverses inspirations locales qui en font l'originalité sans que se perde le souvenir, ni la saveur, des sources originales. Elles forment ainsi une grande famille dont les enfants sont majeurs.

CHAPITRE IV

EXPANSION

On distinguera l'expansion géographique des langues romanes et une expansion interne aux langues elles-mêmes.

I. — Expansion géographique ou externe

Le latin n'a jamais conquis tout l'*Imperium romanum* : la partie orientale resta au grec. Mais, même ailleurs, le processus de latinisation n'arriva pas partout au même terme. Au moment où allaient se produire les invasions barbares, cette latinisation était naturellement moins forte dans les régions excentrées de l'*Imperium*, entrées plus ou moins tardivement dans la communauté romaine. A la suite des invasions, la Romania se trouva avoir perdu : la Grande-Bretagne, les régions flamande et rhénane, la zone alémanique d'Alsace et du nord de la Suisse, la région des Alpes bavaoises, les Balkans et la région danubienne (sauf des groupes importants qui devaient constituer la Roumanie), l'Afrique du Nord. Quant à l'Armorique, elle fut perdue également à cette époque-là, envahie par des Britanniques chassés de chez eux par les envahisseurs germains, à moins que l'on n'accepte la thèse

de Falch'un pour qui le breton continue directement l'ancien gaulois.

Le visage de la Romania, après les invasions, n'a guère varié en Europe. C'est hors d'Europe que les langues romanes vont connaître à l'époque moderne une expansion considérable, dans le domaine géographique. Préciser la qualité de cette expansion est difficile, car la nature de l'expansion peut aller de l'implantation (la langue romane devient la langue maternelle de la majorité des habitants du pays où elle s'est transportée), au simple rayonnement culturel (la langue est étudiée, en dehors de son domaine, par une partie d'une population alloglotte) en passant par l'importation en territoire alloglotte formant des flots romans, et par la superposition d'une langue romane à d'autres langues en vue d'exercer certaines fonctions (vie économique, administrative, religieuse, scientifique...).

La Romania nouvelle est due surtout à l'expansion de l'espagnol, du portugais et du français. L'italien a eu essentiellement une influence culturelle.

Le portugais est vraiment implanté au Brésil ; il y remplit toutes les fonctions sociales d'une langue : langue officielle, langue culturelle, langue courante. Le portugais parlé par environ 95 millions de Brésiliens n'offre que quelques traits locaux qui ne nuisent pas à l'unité de la langue dans son usage « illustre ». Mais le portugais est parlé un peu partout, autour du monde, dans des pays où il se trouve soit implanté, soit simplement superposé à d'autres langues. Il offre souvent alors des parlers de niveaux socioculturels séparés parfois assez largement, et ces parlers différents coexistent. Le domaine linguistique portugais comprend ainsi : le *saramaca tongo* (djoe-tongo) dans la région du haut Surinam (Guyane hollandaise), le portugais et le

créole portugais de l'archipel du Cap-Vert, le créole portugais du Sénégal, le portugais et le créole portugais de la Guinée portugaise et des îles de São Tomé et du Prince, le créole portugais de l'île espagnole d'Annobon (fâ d'Ambu), le portugais et le négro-portugais d'Angola, Mozambique, Zanzibar, Mombassa, Mélinde, le malayo-portugais (créoles de Java, Malacca, Singapour), le portugais de Timor (archipel de la Sonde), le créole portugais de Ceylan, l'indo-portugais (Goa, Damao, Diu), le créole portugais de Macao en Chine. On estime le nombre des usagers naturels du portugais dans le monde à environ 108 millions. Il convient d'ajouter à ce nombre, d'un point de vue scientifique et culturel, les 2 millions de Galiciens qui vivent en Espagne.

Les estimations sur le nombre d'individus qui ont l'*espagnol* comme langue maternelle varient assez : le chiffre de 190-200 millions n'est sans doute pas exagéré. L'espagnol occupe le troisième rang dans le classement des langues selon le nombre de leurs usagers naturels, après le chinois, l'anglais (certains statisticiens placent le hindi-urdu avant l'espagnol ; par ailleurs, le russe est placé au troisième rang si l'on tient compte du fait qu'il est la langue commune de tous les allophones de l'Union soviétique). L'espagnol est la langue officielle de toute l'Espagne (il n'est pas la langue naturelle des Catalans, des Basques, des Galiciens). Hors d'Espagne c'est la langue officielle du Mexique, de toute l'Amérique centrale, de Cuba, de la république Dominicaine et de Puerto Rico dans la mer des Caraïbes, de toute l'Amérique du sud, sauf le Brésil, ainsi que des territoires espagnols d'Afrique (Ceuta, Melilla, Ifni, Sahara espagnol). L'espagnol est abondamment parlé dans les Etats du Sud des Etats-Unis (Arizona, Texas, Colorado, Californie, Nou-

veau-Mexique). On le parle dans les Philippines, mais il y disparaît peu à peu devant l'anglais et devant la principale langue indigène le *tagalog*. Le créole de Curaçao dans les Antilles, le *papiamentu*, offre un lexique à 85 % espagnol, 5 % appartenant au néerlandais et le reste étant d'origine soit africaine, soit anglaise et portugaise. Un autre parler créole espagnol se trouve sur la côte colombienne atlantique. On appelle judéo-espagnol l'ensemble des parlers employés par les juifs expulsés d'Espagne par Fernand et Isabelle (les « Rois catholiques ») à la suite du décret d'expulsion de 1492. Le judéo-espagnol se situe tout autour du Bassin méditerranéen. Employés par des communautés repliées sur elles-mêmes (les Sefardies), ces parlers ont conservé un grand nombre de traits de l'espagnol du XVI^e siècle : distinction des sifflantes sourdes et sonores en phonétique, en morphologie les formes primitives *so* et *esto* pour *soy*, *estoy*, dans le lexique des mots tels que *fadar* : réussir, *adobar* : préparer. On dit le judéo-espagnol en décadence partout. Néanmoins des journaux paraissent en judéo-espagnol.

En Espagne la langue académique subit des influences locales notamment celle de l'andalou, caractérisé par sa phonétique au consonantisme relâché. A Madrid, sans parler de l'influence d'un langage familier qui fait beaucoup d'emprunt à l'argot gitan (*caló*), la langue de tous les jours subit peu à peu une influence andalouse, surtout dans les quartiers périphériques. La situation de l'espagnol en Amérique est assez complexe étant donné les nombreux facteurs ethniques, sociaux, géographiques de différenciation régionale, dus à la grande expansion de l'espagnol primitif. Les dialectologues sont arrivés à distinguer jusqu'à vingt-trois zones

de différenciations dialectales. Mais on peut regrouper ces zones en quatre aires principales : la zone mexicaine, la zone caraïbe (Amérique centrale, Antilles, côte vénézuélienne), la zone andine (Venezuela, Chili, Colombie, Equateur, Pérou, Bolivie), le Rio de la Plata (Argentine, Uruguay et sud du Paraguay). Les causes secondes de la différenciation sont naturellement diverses. L'adstrat indien d'abord : on a remarqué que les langues indiennes de la cordillère des Andes sont plus « consonantiques » que celles des plaines et l'on constate que l'espagnol andin a tendance à réduire l'expression vocalique (à Quito et même à Mexico) tandis que, par contre, les consonnes sont plus faiblement prononcées à Caracas ou aux Antilles. Ailleurs c'est le caractère archaïque et/ou rural qui a fait développer le « voseo » comme dans la zone du Rio de la Plata. Ailleurs encore se fait sentir l'influence d'un peuplement alloglotte (les italianismes du Rio de la Plata). Au Mexique l'influence de l'anglais est notable dans un certain nombre de calques (*esperar por* : *to wait for*, *viaje redondo* : *round trip*, etc.). Certains traits de l'espagnol d'Amérique s'étendent à toute l'Amérique hispanophone, tel le « seseo » (prononciation *siego* pour *ciego*). D'autres sont très répandus comme le « yeísmo » (prononciation *vaya* pour *valla*). Le lexique est dans son ensemble le lexique espagnol. Il convient de ne pas attacher trop d'importance aux développements locaux comme *camión* désignant l'autobus à Mexico tandis que le *materialista* est le conducteur de camions transportant des matériaux ! ou au fait qu'un même objet est rendu par des termes différents suivant les lieux : un *veston* peut être un *saco*, un *paltó*, une *americana* ou une *chaqueta*. L'emprunt de mots aux parlers indiens, pour désigner des

termes de civilisation, ne peut être pris en compte parmi les causes de différenciation : ces termes enrichissent simplement l'espagnol général. L'accord n'est pas toujours parfait entre l'espagnol d'Amérique et l'espagnol péninsulaire, même pour les néologismes : un magnétophone est en Espagne un *magnetófono/magnetofón* tandis qu'on emploie en Amérique *grabador/grabadora*. L'Académie de la langue, l'Institut de culture hispanique s'efforcent de lutter contre la tendance naturelle à la différenciation de l'espagnol des deux côtés de l'Océan.

Il est difficile d'évaluer exactement le nombre d'individus dont le français est la langue normale : on peut estimer qu'il y a environ 80 millions de francophones véritables dans le monde. En Europe le français est actuellement la langue usuelle de tous les Français, mais elle n'est la langue d'origine (ou langue ethnique) que des régions de l'ancienne langue d'oïl auxquelles on peut ajouter les régions appartenant au franco-provençal qui ont de bonne heure subi l'influence prédominante du français de Paris. Ailleurs le français est devenu, de nos jours, la langue courante, quoique encore à des degrés divers, d'environ 12 millions d'Occitans, de près d'un million et demi d'Alsaciens-Lorrains, de quelque 150 000 Flamands, de plus de 100 000 Basques et d'à peu près 300 000 Corses. En Belgique 45 % de la population est de langue française (l'ancien dialecte d'oïl, le wallon, y est toujours parlé et relativement cultivé ; c'est le seul dialecte de l'ancien français à avoir conservé une personnalité bien vivante). En Suisse 1 000 000 de citoyens suisses (22 % de la population), originaires de parlers franco-provençaux, ont le français comme langue naturelle, les parlers d'origine n'ayant plus rang que de « patois ». La situation linguistique est assez

complexe dans la région autonome du val d'Aoste (100 000 habitants) en Italie, qui a le français comme langue officielle. Le français est également « la langue officielle de fait » des quelque 340 000 habitants du grand-duché de Luxembourg dont la langue naturelle est un dialecte germanique fortement mêlé de vocables français. C'est le français qu'utilisent les tribunaux et c'est lui qui sert à rédiger les textes officiels. Il est également enseigné dès l'école primaire.

Hors d'Europe, le français est, avant tout, la langue du groupe des Canadiens français estimés à environ 6 millions et demi ; le domaine du français au Canada comprend essentiellement la province de Québec (4 millions et demi de francophones). Les groupes les plus importants de francophones se trouvent ensuite dans la province d'Ontario et au Nouveau-Brunswick. D'autres groupes moins importants sont dispersés dans le reste du Canada. Un grand nombre de Canadiens se sont établis aux Etats-Unis où ils forment d'importantes minorités dans le Massachusetts, le Maine, le New Hampshire, le Rhode-Island, le Vermont, le Connecticut, la Californie. Dans certaines villes (Manchester au New Hampshire) ils forment même la majorité. Beaucoup de ces immigrés francophones passent à l'anglais sauf dans quelques agglomérations d'Etats du nord-est des Etats-Unis. En dehors du Canada et de ces agglomérations des Etats-Unis, le français est encore, en Amérique, la langue des îles de Saint-Pierre-et-Miquelon, de la république d'Haïti (4 768 000 hab.) où le français commun se superpose au parler créole, des Antilles françaises qui constituent les deux départements français de la Guadeloupe et de la Martinique et où également le français commun se juxtapose à un parler créole de même type que celui

d'Haïti. De même en Guyane française à côté du français commun vit un français créole. Le français est encore parlé en Louisiane bien qu'il y soit menacé de disparition : une élite restreinte de vieilles familles créoles françaises de la Nouvelle-Orléans parlent un français commun assez pur ; des groupes ruraux venus d'Acadie (aujourd'hui Nouvelle-Ecosse et Nouveau-Brunswick, provinces canadiennes) conservent encore un parler rural fort proche du français d'Acadie ; enfin un créole négro-français (le gombo) est actuellement presque entièrement évincé par l'anglais. Un créole français est également parlé dans les régions qui firent jadis partie de l'empire français des Antilles : les îles Vierges, les îles du Vent. De même à Trinidad : ce pays bien que n'ayant jamais été possession française était peuplé d'une majorité de colons français. D'autres groupes de population francophone se trouvent ailleurs qu'en Amérique. Dans l'archipel des Mascareignes, français commun et créoles français coexistent à l'île de la Réunion département français et à l'île Maurice, autrefois dénommée île de France, territoire anglais aujourd'hui. Aux îles Seychelles, jadis françaises, anglaises depuis la conquête entre 1810 et 1814, la population (environ 50 000 hab.) emploie un créole français. En Inde, à Pondichéry notamment, subsiste un foyer de rayonnement de la langue et de la culture françaises. Enfin en Océanie le français est la langue des archipels désignés sous le nom collectif d'Etablissements français d'Océanie, de la Nouvelle-Calédonie et de la partie française des Nouvelles-Hébrides.

En dehors des pays proprement francophones, le français est également, sous des formes particulières, plus ou moins importantes, la langue seconde d'un certain nombre de pays d'Afrique qui sont, de ce

fait, plus ou moins bilingues. C'est le cas de l'Afrique du Nord, le rôle du français étant, en gros, plus important en Tunisie qu'au Maroc. Le français est la langue internationale, officielle et de l'enseignement (à côté, de plus en plus, de l'enseignement dans les langues locales en certains pays comme Madagascar et la Mauritanie) de dix-huit pays d'Afrique en dehors de l'Afrique du Nord. La situation du français dans l'Asie du Sud-Est est assez confuse à la suite des événements qui ont suivi la seconde guerre mondiale. Le français était naguère, au Sud-Vietnam, la langue d'usage et celle de journaux quotidiens de groupes d'origine française ou eurasienne. Le français était, de plus, seconde langue dans les Etats correspondant à l'ancienne Indochine française. Les derniers événements, dans cette région, auront fait perdre de l'influence au français : dans la République socialiste du Vietnam le français est, actuellement, enseigné dans l'enseignement secondaire à côté des trois autres langues : l'anglais, le russe et le chinois.

La position du français dans le monde est très forte. Il est la langue nationale ou officielle de trente pays ; un peu moins d'un tiers des délégations des nations l'emploie à l'Assemblée des Nations Unies. Le français est, à peu de chose près, à égalité avec l'anglais comme langue officielle des organisations et congrès internationaux (paradoxalement ce sont souvent les délégués français qui renoncent au français pour l'anglais !). Mais le sort du français est en grande partie dépendant de la volonté des Français dans le monde : on doit signaler, à ce point de vue, l'action d'un organisme comme l'Alliance française.

L'expansion de l'*italien* n'est pas comparable à celle des langues dont on vient de parler. Les élé-

ments italiens étaient importants dans la *lingua franca* ou *sabir* des ports de la Méditerranée dû à l'expansion commerciale de Gênes et de Venise. L'italien est langue de culture dans l'île de Malte et, à côté de l'anglais, langue officielle en Somalie. Des groupes importants d'italophones résident en Tunisie, en Amérique du Nord et en Argentine. Mais surtout l'italien a pour lui d'être la langue d'un pays qui jouit d'un grand prestige culturel.

On peut dire, en conclusion de ce coup d'œil sur l'expansion externe des langues romanes, qu'au moins deux d'entre elles, l'espagnol et le français, peuvent être estimées « langues mondiales ». Par le nombre de leurs usagers elles sont, certes, loin de s'approcher du chinois (680 000 000 d'usagers), mais le chinois n'est pas une langue homogène et son extension est localisée. L'espagnol et le français restent homogènes partout où ils sont utilisés dans le monde et ils sont utilisés largement ailleurs que dans leur aire d'origine. Ils réunissent tous deux les différents traits qui font d'une langue une langue que l'on peut dire mondiale : nombre relativement élevé d'usagers, audience étendue bien au-delà de ces usagers, large répartition à travers le monde, prestige culturel, politique, économique, reconnaissance officielle du statut de langue internationale. Le portugais, sans s'élever au même rang, a cependant largement dépassé le rang de langue vernaculaire et son expansion est en fonction désormais du rôle toujours plus grand que le Brésil est amené à jouer en Amérique du Sud.

II. — Expansion interne

L'expansion externe des langues romanes a dépendu et dépend surtout de facteurs extra-linguistiques. Il est un autre genre d'expansion qui a joué et joue un très grand rôle pour les langues romanes. On peut l'appeler expansion interne ou inter-langues ou interlinguistique. Ici, si le rôle des facteurs extra-linguistiques n'est pas négligeable, les facteurs linguistiques ont une grande importance par le fait qu'ils influent plus ou moins sur la structure même des langues, surtout sur la structure lexicologique mais aussi sur la structure syntaxique. On peut distinguer trois sortes d'expansion interne ou inter-langues : 1) à l'intérieur même de chaque langue romane, un fort mouvement d'expansion de la « latinité » ; 2) d'une langue romane à une autre langue romane, expansion d'une certaine « romanité » ; 3) hors des langues de la Romania, expansion d'une latino-romanité européenne.

1) Si les langues romanes étaient restées telles que les avait façonnées leur formation spontanée due aux lois phonétiques, par exemple, et si elles n'avaient vécu que sur leur propre fond d'héritage latin, elles auraient aujourd'hui un visage assez différent de celui qui est le leur. Mais, dès leur première apparition en tant que langues nouvelles, les langues romanes ont été largement ouvertes à l'expansion du latin en elles. Facilement remarquable du point de vue du lexique, le phénomène de ce que l'on appelle les latinismes atteint également la syntaxe et la sémantique. Dans toutes les langues romanes, sauf le roumain, le nombre des « mots savants » venus du courant latin est extrêmement élevé. En ancien occitan, par exemple, le vocabu-

laire des sciences est entièrement latin. Par l'intermédiaire des traités de grammaire ou de rhétorique occitans, ou par celui du Codi, rédigé en oc à l'Ecole de Droit de Montpellier, dès le Moyen Age le vocabulaire français de la rhétorique et du droit est entièrement latin. Il en va de même pour le vocabulaire de la médecine et de la théologie en oc d'abord (par suite de l'importance de l'Ecole de Médecine de Montpellier d'une part, et d'autre part par le fait que les communautés religieuses en pays d'oc, hérétiques ou non, usèrent de la langue vulgaire), en français ensuite. La langue de l'administration (chartes, *établissements* ou constitutions des villes) est tout imprégnée de latin si bien que dans les régions d'oc on a l'impression que la langue vulgaire de ces textes continue directement le latin des scribes, notaires, etc. La langue de la Chancellerie royale de Paris suit facilement le même chemin. Le français n'a pas attendu la Renaissance et ses humanistes pour se « relatiniser ». L'expansion du latin dans le français, comme dans l'occitan, est un phénomène qui remonte aux premiers bégaiements de nos langues modernes. Dès les débuts, l'inspiration est trouvée pour développer d'une façon permanente la langue moderne par le recours constant et direct au latin. L'exemple occitano-français a probablement été suivi de très près en Italie, comme on peut le déduire de l'estime en laquelle Brunetto Latini tenait la langue française. En tout cas il est certain qu'à la suite de Dante et de Boccace un fort contingent de mots latins ont été réintroduits dans l'italien ; mais, par le fait que la phonétique de l'italien est beaucoup plus conservatrice que la phonétique du français, ces « mots savants » ressortent beaucoup moins sur l'ensemble lexical d'origine. Il résulte, d'ailleurs, de ce fait, que le français

qui est généalogiquement la langue la moins latine parmi les langues romanes est celle qui, grâce aux emprunts, apparaît comme la plus latinisée. Si, par exemple, en occitan *aigós*, en italien *acquoso*, en espagnol *aguanoso*, *apos* en roumain se rattachent directement au mot chef de la famille *aiga* en occ., *acqua* en it., *agua* en esp., *apă* en roumain, en français *aqueux* apparaît nettement comme dérivé du latin *āquā* à côté de *eau*. Seul le roumain, séparé de la tradition romaine par les vicissitudes de l'histoire, n'a pas connu cette expansion latine dès les origines, le rôle du latin y ayant été tenu par le slave. Cependant le roumain s'est rapidement et assez largement ouvert à l'expansion latine à la fin du XVIII^e siècle grâce à l'influence des milieux catholiques de Transylvanie.

2) Les mouvements d'expansion d'une langue romane à une autre sont d'époques diverses. Dès l'époque carolingienne un important mouvement s'étend de la Romania gauloise vers les péninsules Italique et Ibérique. De nombreux mots espagnols (*ceniza*, *cereza*, *cerveza*, *guiar*, *meaja*, ...) et italiens (*arrivare*, *cervoglia*, *ciliegia*, *cinigia*, *guarire*, *mangiare*...) sont arrivés de la Romania gallo-romane du temps de Charlemagne. En dehors du vocabulaire, on admet, par exemple, que -ss- du subjonctif est arrivé en Italie de France au début du X^e siècle. L'expansion linguistique du français se fit, en grande partie, grâce aux itinéraires de pèlerinages (*stratae francigenae*, *camino francés*) et, en partie, grâce à l'influence de l'abbaye de Cluny. Au XII^e siècle, soit vers l'Italie, soit vers l'Espagne et le Portugal, fut grande l'influence linguistique de l'occitan et du français grâce au rayonnement de la civilisation des troubadours ainsi qu'à celui de l'épopée

française. Le nouveau futur roman synthétique se serait, par exemple, graduellement diffusé dans les autres langues romanes à partir de la France d'oïl ; la diffusion du suffixe -*ança* en ancien espagnol et en ancien portugais, -*anza* en italien, serait directement due à l'influence de l'occitan. Les langues, au-delà des Alpes et au-delà des Pyrénées, sont, dès le Moyen Âge, dans le domaine chevaleresque, ecclésiastique, administratif (à un moindre degré) et littéraire, toutes pleines de mots français et occitans.

Au XVI^e siècle c'est le français qui prend un bain d'espagnolisme et surtout d'italianisme, sans compter la place importante que prend l'occitan, dénommé alors gascon et dont l'influence est souvent confondue avec celle de l'espagnol ou de l'italien (suff. -*ada*, par ex.). Au point de vue de la morphosyntaxe et de la stylistique, au début du XVII^e siècle, Malherbe estimera qu'une tâche linguistique urgente est la dégasconisation du français parlé à la Cour.

A partir du XVIII^e siècle le mouvement d'expansion reprend largement du français vers les autres langues romanes. Au XVIII^e et encore plus au XIX^e siècle l'influence du français sur le portugais est la plus importante de toutes les langues étrangères, non seulement dans le domaine lexical mais aussi dans les domaines de la morphologie et de la syntaxe. De même la syntaxe espagnole s'est trouvée entachée de gallicismes nés aux XVIII^e et XIX^e siècles. L'influence du français est remarquable également dans l'espagnol d'Argentine (dans tout le Rio de la Plata), influence d'autant plus notable que de nombreux immigrants italiens n'ont pas réussi dans cette même région à faire prévaloir l'influence linguistique italienne. Cette prépondérance de l'influence linguistique du français est donc

due au caractère universel de la culture véhiculée par la langue. En ce qui concerne l'italien, les travaux de Tagliavini, Buck, Migliorini, Hope ont montré que l'influence du français, depuis le XVIII^e siècle, au cours des XIX^e et XX^e siècles, s'exerce non seulement par les emprunts, mais d'une façon plus significative par de nombreux calques. Quant au roumain, on jugera de l'influence linguistique du français par ces mots de Pușcariu déclarant que, lors de la nouvelle orientation du roumain au début du XIX^e siècle, « le mot européen pour nous, Roumains, était français ». Il suffit de parcourir quelques pages du *Dicționarul limbii române moderne* de l'Académie de la République populaire roumaine, publié sous la direction du Pr D. Macrea, pour se rendre compte de l'influence du français sur le roumain. On a calculé que le pourcentage des mots français dans le roumain était de 29,69 % tandis que celui des mots d'origine italienne, espagnole, anglaise, allemande et russe tous réunis atteignait à peine 3,81 %.

3) L'expansion de la latinité-romanité s'est exercée également en dehors des langues romanes. Le cas de l'anglais est bien connu dont une partie importante du vocabulaire est roman. Avant d'envahir les îles Britanniques, les Anglo-Saxons avaient déjà emprunté au latin des mots tels que *butter*, *cheese*, *kitchen* ; *mile*, *street*, *wall* ; *satur(day)*, *bishop*. Un certain nombre d'éléments latins se trouvaient pris dans les noms de lieux (*Lancaster*) qui furent adoptés définitivement. L'anglais connut, lui aussi, au Moyen Age, un important afflux de mots latins dus au christianisme, des mots se rapportant aux institutions et à l'enseignement de l'Eglise (*altar*, *angel*, etc.) naturellement, mais bien d'autres égale-

ment tels que *camel*, *circle*, *place*, *purple*, *silk*, *tiger*, etc. A partir du XIV^e siècle, à la suite de traductions du latin en anglais, des centaines d'expressions techniques ou savantes furent adoptées par l'anglais. Et cette expansion-là du latin continue de nos jours. A quoi s'ajoutent naturellement plusieurs milliers de mots français introduits pendant les trois siècles de bilinguisme qui ont suivi la conquête de l'Angleterre par les Normands. Tous ces éléments réunis font que l'influence linguistique de la latinité-romanité est incontestablement soutenue à travers le monde par la langue anglaise, qui diffuse dans le monde occidental et au-delà d'innombrables termes scientifiques marqués par leur origine latine (ou grecque).

Comme pour l'anglais, l'influence latine s'est exercée sur l'allemand très tôt : dès l'époque de César (cf. *Kaiser*) ; et le latin est resté, en Allemagne, la langue de la science et de la religion jusqu'au XVIII^e siècle. Dès les premiers siècles de l'ère chrétienne, l'allemand a reçu un fort contingent de mots latins se rapportant à la maison (*mauer*, *fenster*, *pforte* ; *küche*, *tisch*, *schüssel*, etc.), à l'agriculture (*frucht*, *pflanze*, *kirsche*, etc.), à l'ordre militaire (*wall*, *kampf*, *pfeil*, etc.), au droit et à la politique (*pacht*, *kerker*, *zoll*, etc.), au commerce (*pfund*, *strasse*, *kauf*(man), etc.), à la religion (*kreuz*, *münster*, *predigen*, etc.). A la Renaissance une seconde expansion latine a introduit en allemand quantité de termes tels que *universität*, *student*, *doktor*, *kurieren*, *zitieren*, *polizei*, etc.). Il convient d'ajouter qu'un très grand nombre de termes du lexique allemand n'est germanique que par la forme, sémantiquement le contenu idéologique est latino-roman : *ausnahme* (exceptio), *auslesen* (eligere), *anschauen* (intueri), *barmherzig* (misericors), *erziehen* (educere), *gegen-*

stand (objectum), *grossmut* (magnanimitas), *herabsehen* (despicere), *lebenslauf* (curriculum vitae), *mitleid* (compassio), *vaterland* (patria terra), *wohlwollen* (benevolentia), etc. Enfin un très grand nombre de mots français, bien plus important que celui de tous les mots empruntés aux autres langues, a été reçu en allemand surtout aux XI^e et XII^e siècles d'abord (termes de la civilisation chevaleresque), ensuite au XVI^e et surtout aux XVII^e et XVIII^e siècles (termes de la langue militaire, termes de parenté, mots appartenant au vocabulaire de la toilette, de la vie de société). A un moindre degré, l'allemand a fait des emprunts à l'italien à l'époque des Croisades et à la Renaissance.

Les langues slaves elles-mêmes ont largement accepté l'expansion romane. Pour ne parler que du russe, dès Pierre le Grand, la langue s'ouvre aux termes français. Au XVIII^e et au XIX^e siècle, la Russie participant de plus en plus largement à la vie européenne, le russe s'emplit de mots allemands, anglais, et surtout français appartenant au vocabulaire de la civilisation quotidienne. Le russe moderne participe largement du vocabulaire scientifique international d'origine gréco-latine. Des colonnes entières du *Dictionnaire russe-français* de L. V. Scerba et M. I. Matoussevitch (Moscou, Encyclopédie soviétique, 1969) ouvert au hasard sont remplies de ces mots qui appartiennent à la « langue européenne » et qui sont le produit de l'expansion de la romanité.

En résumé, on peut dire que l'influence linguistique latino-romane a profondément pénétré non seulement l'expression linguistique européenne mais, sans aucun doute, l'expression linguistique internationale moderne.

BIBLIOGRAPHIE SOMMAIRE

- AUERBACH (E.), *Introduction aux études de philologie romane*, Frankfurt, Klostermann, 2^e éd., 1961.
- BAL (Willy), *Introduction aux études de linguistique romane, avec considération spéciale de la linguistique française*, Paris, Didier, 1966.
- BEC (Pierre), *Manuel pratique de philologie romane*, Paris, Picard, t. I, 1970 ; t. II, 1971.
- BOURCIEZ (Edouard), *Eléments de linguistique romane*, Paris, Klincksieck, 5^e éd. révisée par l'auteur et par les soins de Jean BOURCIEZ, 1967.
- MILLARDET (Georges), *Linguistique et dialectologie romanes. Problèmes et méthodes*, Montpellier, Société de la Revue des Langues romanes, 1921-1923 (malgré la date de parution, ouvrage qui a gardé toute sa valeur).
- POP (Sever), *Encyclopédie de la philologie romane. Langues et dialectes de la Romania*, Louvain, Impr. Offset Dewallens, 1956-1957.
- WARTBURG (Walter von), *La fragmentation linguistique de la Romania*, Paris, 1967 ; trad. franç. de *Die Zergliederung der romanischen Sprachräume*, 1950.

REVUES

- Revue des Langues romanes*, Montpellier, depuis 1870.
- Revue de Linguistique romane*, Lyon, Strasbourg, depuis 1925.
- Romania*, Paris, depuis 1872.

TABLE DES MATIÈRES

INTRODUCTION. — Définitions	5
I. Définition de l'expression <i>langues romanes</i> , 5. — II. Définition de la place occupée par les langues romanes, 11.	
CHAPITRE PREMIER. — Développement historique des études romanes	15
I. La préhistoire de la linguistique romane, 15. — II. La linguistique romane au XIX ^e siècle, 22. — III. La linguistique romane au XX ^e siècle, 31.	
CHAPITRE II. — Origine des langues romanes.....	45
I. Le latin vulgaire, 46. — II. Les strats, 50. — III. Groupements humains et communications, 60. — IV. Apparition des langues romanes, 68.	
CHAPITRE III. — Les langues romanes.....	72
I. Développement, 72. — II. Typologie, 95.	
CHAPITRE IV. — Expansion	110
I. Expansion géographique ou externe, 110. — II. Expansion interne, 119.	
BIBLIOGRAPHIE SOMMAIRE	127