THE ROMANCE LANGUAGES

Edited by
MARTIN HARRIS
and
NIGEL VINCENT

Routledge Language Family Descriptions

THE ROMANCE LANGUAGES

Other published works in the series:

The Celtic Languages
The Slavonic Languages
The Germanic Languages
The Semitic Languages

Forthcoming:

The Indo-European Languages
The Dravidian Languages
The Uralic Languages
The Turkic Languages

THE ROMANCE LANGUAGES

EDITED BY

MARTIN HARRIS
AND
NIGEL VINCENT

London and New York

First published 1988
by Routledge

Reprinted 1997, 2000
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN
270 Madison Ave, New York NY 10016

Transferred to Digital Printing 2006

Routledge is an imprint of the Taylor & Francis Group

Typeset in Times by Leaper and Gard Ltd, Bristol

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication Data
A catalogue record for this book is available from
the Library of Congress

ISBN 0–415–16417–6

Publisher's Note
The publisher has gone to great lengths to ensure the quality of this reprint but points out that some imperfections in the original may be apparent

Contents

Abbreviations

abl.	ablative
acc.	accusative
adj.	adjective
Am.	American
Ann.	Annobonese (fa d'ambó)
AR	Arumanian
art.	article
attr.	attributive
aug.	augment
Auv.	Auvernhat
BP	Brazilian Portuguese
Byz.	Byzantine
Cab.	Caboverdiano
Cal.	Calabrese
Camp.	Campidanese
Cat.	Catalan
Cha.	Chabacano
CLat.	Classical Latin
compl.	complement
cond.	conditional
conj.	conjugation
dat.	dative
decl.	declension
def.	definite
dem.	demonstrative
det.	determiner
dir.	direct
disj.	disjunctive
Dom.	Dominican
DR	Daco-Rumanian
ELeng.	East Lengadocian
Eng.	English
EP	European Portuguese
f.	feminine
fam.	familiar
Fr.	French
fut.	future
Gasc.	Gascon
GBi.	Guinea-Bissau Crioulo
gen.	genitive
ger.	gerund
Gk.	Greek
GR	Gallo-Romance
Gua.	Guadeloupéen
Guy.	Guyanais
Hai.	Haitian
imp.	imperative
imperf.	imperfect
impers.	impersonal
ind.	indicative
indef.	indefinite
indir.	indirect
inf.	infinitive
IR	Istro-Rumanian
It.	Italian
Lat.	Latin
Lem.	Lemosin
Leng.	Lengadocian
Log.	Logudorese
Lou.	Louisianan

Luc.	Lucano
m.	masculine
Mar.	Martiniquais
Mau.	Mauritian
Mil.	Milanese
Mod.	Modern
n.	neuter
Nap.	Naples
neg.	negative
nom.	nominative
Nuor.	Nuorese
obj.	object
Occ.	Occitan
OLat. etc.	Old Latin etc.
Osc.	Oscan
p.	person
pal.	palatal
Pal.	Palenquero
Pap.	Papiamentu
part.	participle
perf.	perfect(ive)
pers.	personal
Pied.	Piedmontese
p.-in-p.	prospective-in-past
pl.	plural
pluperf.	pluperfect
pol.	polite
Port.	Portuguese
pos.	position
pp.	past participle
pred.	predicative
prep.	preposition
pres.	present
pret.	preterit
Pri.	Principense
pron.	pronoun
pros.	prospective
Prov.	Provençal
PWR	Proto-Western-Romance
refl.	reflexive
Réu.	Réunionnais
R-R	Rhaeto-Romance
Rum.	Rumanian
Sard.	Sardinian
Sey.	Seychellois
sg.	singular
Sic.	Sicilian
SIt. etc.	Southern Italian etc.
Sl.	Slavic
SLu.	Saint Lucian
Sp.	Spanish
STo.	São Tomense
su.	subject
subj.	subjunctive
suff.	suffix
Sur.	Surselvan
SVO etc.	subject–verb–object etc.
Tusc.	Tuscan
VC	verb–complement
Vegl.	Vegliote
Ven.	Venetian

Preface

It may not come amiss, in a book where matters diachronic and synchronic are inextricably intertwined, to begin by saying a few words about its genesis before considering its structure. The immediate origin of the volume may be traced to the fact that a number of the contributors had previously been invited to write chapters for Bernard Comrie's encyclopaedic compilation *The World's Major Languages* (Croom Helm, 1987). This invitation had the effect of forcing each of us to think hard about which aspects of the structure, history and sociology of our chosen languages should, and/or could, be included within the inevitable length restrictions. At the same time, it was clear that the format of Comrie's volume, while perfectly understandable and justifiable in its own terms, was likely to misrepresent the Romance language family as a whole for two distinct reasons. In the first place, the individual chapters were not always long enough to allow a full treatment of certain key aspects. Second, a number of languages were excluded which, albeit minor on politico-demographic grounds, were nonetheless of major significance in providing the evidence necessary to a successful reconstruction of the historical evolution of the family and to a proper appreciation of its current typological diversity. We therefore decided to go ahead with a project which took as its goal the characterisation of one particular language family, where the material for inclusion was determined principally on linguistic and historical grounds internal to the family itself, and for the most part independent of external political considerations. To this end, the chapter lengths in the present volume are roughly twice those to be found in Comrie's and the number of Romance languages has been expanded from the original five — French, Spanish, Portuguese, Italian and Rumanian — to encompass Catalan, Occitan, Sardinian and Rhaeto-Romance. Of course, there are still exclusions. Some might argue for a separate chapter on Galician or Istro-Romance, or for more detailed coverage of 'dialects' such as Andalusian or Sicilian or Brazilian Portuguese. With even more space, it might have been possible to give way to these, and other conceivable and similar, demands, but one must always stop somewhere on the continuum from

idiolect to language family. As things stand, we believe that the present volume achieves a reasonable coverage of the linguistic diversity of both ancient and modern Romania.

Since in other respects the model of the chapters prepared for *The World's Major Languages*, with their mixture of the synchronic and diachronic, their relative theoretical neutrality and their adherence to a discursive style of narrative, appeared satisfactory, there was every incentive to keep the same nucleus of authors, who have in several cases incorporated some of the material from their earlier pieces. Naturally, however, new recruits had to be found for the new chapters. In addition, we decided to prepare a new chapter on Latin, where the emphasis was on looking forward to the ways in which that language has changed in the course of its development and fragmentation into the Romance languages. At the same time, it also seemed a good idea to gather into one introductory chapter facts of distribution and numbers of speakers, thus leaving each single chapter to be organised according to mainly structural linguistic criteria. Finally, in view of the recent rapid expansion of interest in the topic of pidgin and creole languages, we felt it would be useful, and indeed genuinely innovatory in a work of this kind, to add a chapter surveying and synthesising current research into Romance-based pidgins and creoles.

One respect in which we have departed from the model of Comrie's volume is in our attempt to impose a degree of uniformity of coverage on the central descriptive chapters. Not only do they fall into the same major sections, namely phonology, morphology, syntax and lexis, but within each section we have tried to ensure that a similar range of core topics has been covered. Thus, anyone wishing to look up say the sequential constraints on clitic pronouns or the historical evolution of tonic vowels in all the languages represented should not be disappointed. We have, however, deliberately refrained from going further in this direction. Thus, the reader will find that the individual chapters differ in the selection of non-core topics dealt with, in the mix of diachrony and synchrony, in the relative balance between the sections and in the types of theoretical approach adopted. While some may lament such heterogeneity, our own feeling is that it enhances the value of the work by demonstrating how differences of perspective may lead to varying assessments of significance with respect to broadly similar bodies of data.

How, then, might this book be used? A number of possibilities spring to mind. On the one hand it might serve as a work of reference for points of information ranging from the very particular — e.g. how many phonemically distinct laterals are there in Catalan? — to the more general — e.g. what are the main sources of loan vocabulary in Spanish? — and so on. Another reader might wish for a general overview of the history and structure of a particular language, either for linguistic purposes such as typological investigation or as background to a study of the history,

literature or whatever of the region concerned. Yet again a third type of reader might wish to check out a given phenomenon such as basic word order in all the members of the family.

As regards intended level of readership, we have tried to ensure that the book is both sufficiently clear and expository for it to be used for general reference or as a textbook on undergraduate or graduate courses in Romance linguistics. Yet we also cherish the hope that it will offer information and occasional insights of value to scholars in linguistics and allied disciplines. To aid the reader, we have included both an analytical index and a system of cross-reference within the main body of the text. Each chapter concludes with a select list of essential reference works and further reading.

At this point, it is perhaps worth adding a word of advice aimed particularly at readers of the whole book or at those who use the index and the cross-references to follow up a given theme or area of data across a number of chapters. The multiple authorship, which distinguishes this work from the usual run of Romance linguistic manuals, brings with it obvious advantages in terms of breadth and depth of coverage, but also some attendant complications which derive from the contributors' differing perspectives and ranges of interest. The latter are easier to cope with than the former. It is clear, for example, that the more developed treatment of lexis in the Spanish chapter is not due so much to any intrinsic features of the history and structure of that language's vocabulary as to John Green's authorial decision to give special prominence to this topic in his chapter. Interested readers may then use the model of extended coverage Green provides and the data from the corresponding sections of other chapters to construct for themselves more detailed accounts of, say, loanwords in French or word-formation in Rumanian. Similarly, one author may have used a technique which could have potentially interesting results when applied to data from another language. Thus, Wheeler offers on page 206 a brief description of a method for compiling an inventory of typically Catalan core vocabulary. It would be intriguing to see this applied more generally. In the field of syntax, Haiman's detailed analysis of word-order in Rhaeto-Romance has obvious and as yet inadequately explored implications for the analysis of northern Italian dialects and their relation to the standard language. These are questions which are only briefly hinted at in the Italian chapter, though there is some interesting related discussion *vis-à-vis* Sardinian (pp. 338ff). More generally, all the relevant sections in the different chapters could be read consecutively, and the patterns of recurrence and diversity assembled into a single account of word order in Latin and Romance. All these, and — we hope — many more, represent ways in which the contents of the present volume will serve not just to document the results of current and previous work but also to stimulate new research into the extraordinarily rich data, both synchronic and diachronic, which the Romance languages offer.

Elsewhere, the differences between chapters lie not in emphasis but in the contrasting ways contributors have chosen to solve a problem of presentation. For example, in languages like Occitan or Sardinian, where for social and historical reasons there is no recognised standard or norm, how is one to state the facts? For Occitan phonology Wheeler opts for 'a basically diasystemic and somewhat abstract approach' (p. 246–7). Jones prefers to supplement his general overview of Sardinian phonology with a detailed treatment of the dialect of one village. Or compare with both of these the problem in the case of Rhaeto-Romance where there is a complex pattern of standardisation based on the recognition of different dialects at different periods in different places. This has necessitated a presentation in parallel of a number of systems in Haiman's chapter.

Despite the foregoing, the enterprising reader may still be puzzled, if not by differences of emphasis and presentation then by apparent contradictions between chapters. Yet it is in the nature of linguistic inquiry that analysts will at times disagree about how to interpret a given example or construction. In the case of a closely related group like the Romance languages, there is the additional complication that linguist A might propose a particular analysis of a phenomenon in language *x*, while linguist B might offer a different account of an apparently identical phenomenon in language *y*. Thus, to take a case in point, Vincent in his chapter on Italian (pp. 306ff) gives a semantically based treatment of the distribution of the various complement types in that language. Green, by contrast, suggests that complement selection in Spanish (pp. 117–18) is for the most part lexically arbitrary. Do such instances represent structural differences between the languages or theoretical differences between the linguists? The question matters both synchronically and diachronically. Vincent (pp. 65ff) also gives a semantic account of complementation in Latin, suggesting therefore that the change between Latin and Italian involves the development of a new pattern of semantic motivation which goes hand-in-hand with the emergence of a new set of formal exponents. On Green's view, either the lexical arbitrariness has to be projected back to the Latin stage, or else the history of Spanish — and indeed of Italian, if Green's analysis can be successfully generalised — is one of successive loss of semantic motivation. As it happens, in this instance the complete resolution of these differences is not possible, since much of the necessary detailed linguistic analysis has not yet been carried out. Once again, our hope is that allowing the different perspectives to stand side by side will serve to highlight areas for potential future research.

While we obviously cannot claim that everyone who consults this volume will find their queries answered, we believe that it will take its place on the shelves of both libraries and individuals as a reliable and up-to-date guide to the history and structure of the Romance languages and to the way they are currently being investigated.

In conclusion, we would like to acknowledge the cooperation of the individual contributors not only in preparing the versions of their own chapters on time, but for putting up with queries and requests for changes, and for cross-reading other chapters and providing often invaluable comments and corrections. Thanks too to numerous friends and colleagues who read drafts and offered advice, and to an anonymous referee for a number of helpful observations. We are particularly grateful to Christy MacHale for drafting the maps, to Jenny Potts for her skilful and informed copy-editing, to Mark Barragry on the production side and last, but certainly not least, to our editor, Jonathan Price.

Martin Harris
University of Essex

Nigel Vincent
University of Manchester

1 The Romance Languages

Martin Harris

The Romance languages, whose history, structure and present-day distribution are the subject of the chapters which follow, share a common source: their development in each case may be traced back to Latin. Latin for its part developed from a form of Italic spoken originally in a number of small communities in Latium (Lazio) in central Italy, probably settled by Proto-Latin speakers around 1000 BC. The Italic branch of Indo-European appears to have been brought to the peninsula towards the end of the second millennium BC, and included Oscan (spoken over much of southern Italy at least until the time of the Pompeii disaster, as graffiti clearly testify), Umbrian (spoken in the north Tiber valley) and a number of other more or less well known varieties in addition to the Latin group of dialects. The label 'Latin' may be said to refer initially to this group of related dialects (including, for instance, Faliscan, spoken around what is now Cività Castellana, some fifty miles from Rome on the north bank of the Tiber), but it soon came firstly to designate the speech of Rome — attested since the sixth century BC — and then to be used as an increasingly broad cover term for a range of related varieties differing along temporal, geographical and social dimensions (see below). Latin was, as we have seen, bordered to the south and east by cognate tongues, while to the north its principal neighbour was the non-Indo-European Etruscan. Farther north still, by the fourth century BC — the time at which Rome was establishing her dominance in central Italy — the Po plain had been settled by speakers of varieties of Celtic (p. 3), a separate Indo-European family, but one which bears a number of striking structural parallels to Italic. In the extreme south, on the other hand, Greek was a recurrent source of external but still Indo-European influence.

As Roman military, political and economic influence spread during the period of expansion of the Roman Empire, firstly within Italy and then beyond, the Latin language also flourished, coming to be spoken in much of western and central Europe, and western north Africa, only Greek (spoken in the eastern Mediterranean and the Middle East) being a serious linguistic rival within the imperial boundaries. In particular, in addition to

those areas of Europe which are currently Romance-speaking, much of southern Britain, the rest of what is now Belgium, Holland, much of Germany, Austria, Yugoslavia and Albania, and a fairly narrow coastal strip in what is today Morocco, Algeria, Tunisia and Libya were also Latin-speaking to a greater or lesser extent. The subsequent retreat of Romance from part of the territory it once occupied — which has of course in recent centuries been vastly more than offset by its expansion overseas — came about mainly in the period after the end of the Empire in the west during the fifth century AD (earlier north of the Danube; cf. p. 23), largely to the profit at that time of Germanic and Slavic languages within Europe and, from the later seventh century onwards, of Arabic in north Africa.

Despite the military and political collapse of the Western Empire and the subsequent loss of territory by Romance, Germanic and Slavic actually made less headway within Europe than might have been expected. In Iberia, for instance, the incoming Visigoths were already Latin-speaking before they arrived, and retained many aspects of the civilisation they found there, with themselves now in a dominant role; the continued use of Romance is therefore hardly surprising. In northern Gaul, to take a second example, a Catholic Frankish kingdom under Clovis emerged by the end of the fifth century, in which Latin was established from the outset as the language of both religion and administration and where a Romance vernacular — with a significant Frankish overlay — rapidly began to develop. The persistence — or reintroduction — of Romance in the area of present-day Rumania, on the other hand, is more difficult to account for, and is discussed in some detail below (pp. 2–3). The survival or otherwise of Romance when political mastery passed into other hands can be ascribed partly to the extent and profundity of earlier Latinisation and partly to the density and pattern of settlement of the newcomers; the use by the Christian church as its 'official' language of Latin/Romance (what some scholars call 'Late Latin' is the same as what others call 'Early Romance') is certainly also a relevant factor. It is with the subsequent fate of Romance in those areas where it did persist that this chapter is primarily concerned.

Of course, as was indicated earlier, even when Roman power was at its height there was not one single homogeneous form of Latin used by all speakers throughout the Empire: social and regional variation, particularly in the spoken language, would have been apparent at all times. There would, for example, have been considerable differences between the speech of Cicero and that of his slaves, or between the Latin spoken by a Roman provincial governor and that of his subjects, and the question of how such varieties should be distinguished and denominated is discussed below (pp. 26–7). During the period between the collapse of the Empire in the west and the emergence of the first Romance vernacular texts in various parts of Europe, one must envisage a situation in which this ever-present variation within Latin was accentuated as the language developed

in ever more divergent ways in different localities. There are three main reasons for this. The first is simply the general tendency towards linguistic fragmentation inherent in the language acquisition process, counterbalanced at all times by the need to communicate with others within a shared speech community. Given the loss of a single uniform education system, and given the increasing separation of various groups of Romance speakers from one another, particularly after the rise of the Moslems in the eighth century shook the cohesiveness of the Western Romance world, this shared speech community must have grown progressively smaller for most speakers; thus the pressures offsetting fragmentation weakened and dialectalisation proceeded apace.

Secondly, there were already during the Empire incipient divergences between the Latin of various provinces, partly at least because of the different language or languages which were spoken (and often continued to be spoken for centuries) in various regions before Latin became the predominant language. Thus there are, for example, considerably more words of Celtic origin in contemporary French and north Italian dialects than in Spanish, standard Italian or (even more so) Rumanian, reflecting the Celts' domination before Rome's expansion both of Italy north of the river Po (*Gallia Cisalpina*: 'Gaul this side of the Alps') and of most of present-day France (*Gallia Transalpina*: 'Gaul beyond the Alps'). One representative example may be found in derivatives of a Latinised Celtic word RUSCA 'bark' (cf. Welsh *rhysg* 'rind'), surviving with various meanings ranging from 'peel' and 'skin' through 'bark' to 'cork' and '(cork) bee-hive' in Gallo-Italian dialects, throughout Gaul, and in Catalan. Within Iberia there seem to have been several languages spoken in various parts before the arrival of the Romans, including (in addition to Celtic) both Basque, a non-Indo-European language still spoken in the western Pyrenees on either side of the Spanish–French frontier, and also another language or language family, Iberian, of unknown provenance and genetic relationship. Very often, the precise source of a word peculiar to all or part of Iberia is unclear; for this reason, those lexical items found in Ibero-Romance that are clearly of long standing and which are apparently neither of Latin nor Celtic origin — such as Sp. and Port. *cama* 'bed' (cf. pp. 118, 165) — seem best labelled simply 'pre-Romance'. In much of central and southern Italy, most of the 'substrate' languages were themselves, as we have seen, of the Indo-European Italic group closely related to Latin itself, although there are limited traces of the influence of Etruscan. From the eighth century BC, there was significant settlement by speakers of Greek in southern Italy and Sicily, with some borrowing of lexical items into early Latin (p. 75), but while one or two Greek-speaking communities survive to this day (p. 19), the effects of this on local Romance dialects appear to have been minimal. As for possible pre-Roman influences on Rumanian, these are lost in the mists of time, partly because the present-day location of Rumanian is very

probably not identical with that where Latin was first learnt (p. 23) and partly because we know virtually nothing of any pre-Roman languages in this entire area (cf. pp. 412–13).

The third reason for the increased linguistic divergence following the break-up of the Empire lies in the languages of the conquerors, whether immediate or subsequent. Thus one expects, for instance, to find most words of Germanic origin in French, particularly in those dialects — such as Walloon — nearest to the eventual Romance–Germanic frontier, with fewer in those other Romance-speaking areas where Germanic settlement was less dense. In Spanish (and to a lesser extent Catalan and Portuguese), one finds a substantial Arabic element (p. 119), reflecting the occupation of significant parts of the peninsula by Arabic speakers for nearly eight centuries, while in the case of Rumanian the constant contacts with Slavic and other non-Romance languages have led to a substantial non-Romance lexical element in the language even in everyday vocabulary (pp. 413–14).

The previous paragraphs have discussed the problem of linguistic divergence as though it were exclusively a lexical phenomenon: this is of course far from the case. Much has been written about the extent to which phonological, morphological and syntactic differences can be attributed to substrate or to adstrate factors, but in very few, if any, cases is general agreement reached. The pronominal use of *on* 'one' (< HOMO 'man') in French (p. 221) is a structure once regarded as certainly of Frankish origin; but while the parallel with modern German *man* 'one' (cf. *Mann* 'man') is indisputable, the direction of any influence between Germanic and Romance — and indeed whether such influence need be postulated at all — remains a contentious issue. Very often too, one finds that exactly the change or pattern under discussion is to be found also in some other Romance variety, or indeed in a totally different language family, in a situation in which the postulated external influence is wholly lacking. Such an instance is the passage of prevocalic initial [f] in Castilian (and Gascon) via [φ] and [h] to Ø ('the loss of initial [f]'), a development once confidently attributed to the influence of Basque or a Basque-related substrate, but parallelled in part at least in a number of southern Italian dialects, where a comparable cause cannot of course be adduced. In short, most non-lexical divergences, in Romance and elsewhere, seem best attributed to internal linguistic evolution, though of course the 'selection' of one change rather than another may be unconsciously favoured by structures found in other languages still actually in use in a multilingual community.

All of these reasons, then, led to the emergence of a number of linguistically distinct areas within the Romance-speaking world. The process of fragmentation, however, went much further. As we have seen, a language as it evolves is subjected always to two conflicting pressures simultaneously: the pressure towards convergence or homogeneity, which facilitates

communication within a perceived speech community, and the pressure towards divergence or heterogeneity (p. 3), caused by the very nature of the language acquisition process, which ensures that no one generation or individual learns their native language in exactly the form in which it has been internalised by their elders. Enhanced social mobility, a high level of education and greater frequency and range of travel and communication strongly favour the former pressure, as the recent retreat of many non-standard dialects clearly indicates; social and geographical immobility on the other hand, with very limited possibilities for education and travel, favour dialectalisation, with each community developing a form of language peculiar to itself, as part of a strong local identity. This process of course never goes so far as to prevent communication with those in the near vicinity; local dialects range along a spectrum, even in districts perceived as being on either side of a major dialect division. (Consider, for instance, the gradation from French to Tuscan via a whole set of French, Franco-Provençal and Gallo-Italian dialects spoken in adjacent parts of France, Switzerland and Italy.) Nevertheless the particular social context of the period between the end of the Roman Empire and the beginning of the Middle Ages did bring about marked linguistic divergence, the dialectal consequences of which remain, albeit now often rather marginally, in all of the present-day Romance-speaking areas within Europe.

From the early part of the Middle Ages, however, at least in the western part of the Romance-speaking world, the first signs of a new phase of linguistic evolution could be discerned, namely the gradual emergence in a particular area of one dialect more favoured for various reasons than any other; from these favoured varieties, at different speeds in different territories, a series of national languages has developed. The precise timing and result of this development, which affected written forms of the language markedly sooner than the everyday spoken idiom, depended on a whole variety of historical factors, in particular the establishment or otherwise of a nation-state in a given region and the policy, explicit or otherwise, of the linguistically dominant group towards those whose native form of speech was other than theirs; these factors are considered in detail below. At this stage we will simply contrast by way of example the fates of Portuguese and Occitan, the former now a major world language and the latter having little official status even in those areas of rural southern France where it is still in use. Portuguese, originating from the Galician dialect spoken in the north-west of the Iberian peninsula, came to be the language of an area which since the mid-twelfth century has been — apart from a brief period from 1580 to 1640 — politically independent of Spain, and has flourished accordingly, whereas Occitan, despite the high standing of medieval 'Provençal' (p. 16) as the literary language of the troubadours and the fact that Occitan dialects were once spoken over more than a third of France, could not compete with the strong desire which developed in the

highly centralist post-revolutionary France for there to be a single national language. We return to this theme at several points in what follows.

We shall now look in turn at each of the branches of the Romance family of languages.

The Romance of Iberia

Within the Iberian peninsula, the major early division, apparent (with the usual caveat about dialect gradation) as early as the ninth century, was between Catalan on the one hand — which had and has close affinities with Occitan north of the Pyrenees (p. 16) and whose speakers were within the Frankish domains for several centuries — and the other dialects of Spain and Portugal, collectively referred to as Hispano-Romance. This latter group includes both the dialects of the Christian north (limited in the tenth century roughly to the northwestern third of the peninsula) and those of the Arab-dominated south, collectively known as Mozarabic. In linguistic terms we may observe that the eight centuries from the first Arab invasion in 711 near Gibraltar (< Arabic *gebel al-Tariq* 'mount of Tariq') to their final expulsion from Granada in 1492 can be characterised as a period involving firstly the gradual divergence of the dialects of the Arabised south from those of the north and then, slowly at first but later with greater speed, the recapture of Mozarabic- or Arabic-speaking territory by speakers of 'Christian' northern dialects. Simultaneously with these developments, we find at first the familiar process of linguistic fragmentation between the Christian kingdoms and then the gradual emergence of two of the resultant dialects, Castilian and Portuguese, to become in due course the national languages of Spain and Portugal.

More specifically, we may observe that as the Reconquest got underway, there was a range of Hispano-Romance dialects, traditionally grouped, largely because of the political divisions of the time, into four, these being, from west to east, Galician, Leonese, Castilian, and (Navarro-)Aragonese, with Catalan still further to the east. Speakers of each of these dialects gradually reoccupied territory to the south, but the central Castilian-dominated swathe gradually grew broader, to the point of cutting off the southward expansion of Catalan, Aragonese and Leonese at points close to Alicante in the east and Badajoz in the west, with a substantial strip further west christianised by speakers of Galician-Portuguese, who reached and recaptured the Algarve by the mid-thirteenth century, at which point modern Portugal may be said to have taken shape. At first, it was largely Mozarabic that these incoming dialects replaced (albeit possibly with some residual influence from Mozarabic on the dialects of Andalusia), but in much of the southern third of the country, from the latter part of the twelfth century onwards, it was often non-Romance languages, in particular Arabic and Berber, which gave way to Castilian, coming in from the

north and now very much in a dominant position. Furthermore, along those eastern and western flanks of Spanish territory initially reconquered by speakers of Aragonese and Leonese respectively, Castilian gained ground fairly rapidly, a process helped no doubt by the fact that the differences between the dialects at that time were significantly less than those found now between standard (Castilian) Spanish and those forms of Aragonese and Leonese which continue to be spoken today.

There are three questions arising from this greatly truncated account which need a brief response. Firstly, why did the Christian dialects of the north have such an easy task in defeating both the Mozarabic forms of Spanish spoken in the reconquered areas and also the non-Romance languages of the occupiers? Secondly, why, within the central group of dialects, did Castilian fairly early emerge as dominant? And thirdly, why and to what extent have Galician and Catalan escaped Castilian hegemony?

The first of these questions is relatively easily answered: Mozarabic could not compete in prestige with the speech of the newcomers, and given the 'religious crusade' nature of the Reconquest, this was clearly even more true of Arabic or Berber. Further, in Arab Iberia, Mozarabic had the status of a spoken patois, the languages of culture and administration having been Arabic and, to a significant extent, Hebrew. All in all, the victory of the northern forms can readily be explained, and the principal long-term effect of Mozarabic on Spanish and Portuguese was as a medium whereby a considerable number of lexical items of ultimately Arabic origin passed in due course into the two national languages of the peninsula.

The second question reflects simply the central role played by Castilian in the Reconquest within what is now Spain. After the recapture of Toledo in 1085, and in particular after the reunion of León with Castile in 1230, this pre-eminence increased, to the point where, when Mozarabic was abandoned, as we have already seen, in favour of the language of the incomers, it was in fact, except of course where the new ruling elite spoke Portuguese or Catalan, Castilian which was inevitably adopted. As elsewhere, the social prestige associated with the court, particularly during the reign of Alfonso X el Sabio (1252–84), reinforced the position established by military success.

The third question is more complex. We have already observed that the southern extension of the most northwesterly Hispanic dialect, Galician, developed into Portuguese, and we shall discuss the present fortunes of that language shortly. Perhaps as a consequence of the success of its offspring, perhaps because of its geographical remoteness, perhaps because of its own distinctiveness and the strong literary and cultural tradition dating back to the flourishing lyrical poetry of the Middle Ages (often written by non-Galicians, as for example Alfonso X), Galician has arguably survived more strongly than either Leonese or Aragonese, the other major medieval

dialects, although a variant of the former, spoken in Asturias around Oviedo and known as Bable, has recently enjoyed something of a resurgence. The fact remains, however, that from the sixteenth to the nineteenth century, Galician was frequently represented as a mere rural patois, and indeed one particularly low in social standing. It needed a literary revival in the nineteenth century and a linguistic revival in the twentieth to restore Galician's fortunes to some extent, although the long-term position is far from clear. Alongside this, one notes the current resilience of Catalan, the only language in Western Europe not the official language of a modern nation-state which can truthfully be said to be secure and whose present position is more fully discussed below.

Spanish (Castilian)

Spanish is not only the official language of Spain (including the Canary Islands and the north African enclaves of Ceuta and Melilla) but is spoken also in 19 republics in Central and South America and the Caribbean (Argentina, Belize — where English is also found — Bolivia, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Uruguay and Venezuela) as a result of the colonisation of that region — and indeed significant parts of what is now the USA — by Spanish speakers from the sixteenth century onwards, with constant reinforcement by incoming Hispanophones since that date. In addition, Spanish is spoken in the US-associated Caribbean territory of Puerto Rico and residually in the Philippines. Not to be overlooked either is the very substantial number of Spanish speakers within the USA itself, a number maintained and augmented primarily from Mexico, from Puerto Rico and from Cuba; officially estimated at some 14 million people, the real figure seems likely to be very significantly higher, perhaps approaching 20 million. Finally, one should note groups of Spanish speakers in Morocco and Western Sahara, and also in Equatorial Guinea, where Spanish is the official language. It is estimated that in all some 280 million people have Spanish as their native tongue or (in parts of both Spain and Latin America) as a second language alongside Catalan or an Amerindian language; it is worth noting that by far the largest number of Hispanophones — well in excess of 70 million — are to be found in Mexico, almost twice the total to be found in Spain. Recent figures suggest that some 12 million people, widely scattered but with a particular concentration in the central Andean region, speak a South American Indian language (of which over 1,500 have been listed), although many of these speakers also have at least some knowledge of Spanish (or, in the case of Brazil, Portuguese). Particularly resilient are Quechua, an official language of Peru alongside Spanish, spread far beyond its original homeland by the Incas in pre-Spanish times, and Guaraní, which shares official status with Spanish in Paraguay, partly at least as a

result of its widespread use by the early Jesuit missionaries; there are, however, also significant Amerindian-speaking communities from Chile northwards to Mexico.

In linguistic terms, the Spanish of Latin America shares a number of features, especially at the phonological level, with the southern variant of peninsular Castilian, Andalusian, which, given the provenance of most of the early settlers, is hardly surprising. There has been much controversy about the influence of substrate languages on aspects of the Spanish of Latin America other than the obvious lexical input, but almost all the phonological and morphosyntactic characteristics examined have counterparts in areas where the proposed source language was not spoken. That is not to say, of course, that there is not mutual influence, especially in bilingual areas — but taken as a whole the linguistic developments of Spanish in America seem best attributed to the normal processes of change, with only marginal effects from the Amerindian substrate.

In comparison with French and Portuguese, Spanish forms the base of relatively few creoles. The most significant group is to be found in the Philippines, known collectively as Chabacano, and deriving to a significant extent at least from Ternateño, a creole which developed in Ternate in the Moluccas from interaction between Spanish and an already existing Portuguese-Malay creole, and which was taken to the Philippines in the middle of the sixteenth century. This seems to have been the principal source of four major Spanish creole dialects in the Philippines, none of which, however, are prospering in the face of constant pressure from English. There are small numbers of speakers of Spanish-based creoles in various parts of the north-west of South America, the best known of these being perhaps Palenquero, showing features derived also from Portuguese and spoken by some 2,000–3,000 people in northern Colombia; for Papiamentu see p. 12. One should also note briefly the existence of an Italianised form of Spanish used by Italian immigrants in the Buenos Aires area (Cocoliche) and a range of Spanish–English hybrids spoken in the American south-west; none of these, however, are sufficiently developed to qualify as true creoles, and appear to have a limited future.

Finally, one should briefly note Judeo-Spanish, the language of the Sephardic Jewish communities who fled from or were expelled from Spain during the fifteenth century. Judeo-Spanish, the collective name for a range of mutually comprehensible variants, shows both conservative and innovatory features, the former above all at the level of phonology and the latter particularly in the lexicon, as would be expected. The best known Judeo-Spanish communities, in the Balkans, suffered greatly during the Second World War, but significant numbers are to be found in Morocco (where the influence of contemporary Spanish is strong), in the United States, particularly New York City (where the pressure to linguistic conformity is also strong), and in Israel, where the Spanish-speaking com-

munity is large enough to support radio broadcasts and a journal *Aki Yerushalayim*. Recent estimates suggest a total of some 200,000 speakers in all.

Portuguese

Portuguese, as we have already seen, developed as a concomitant of the southward movement of speakers of Galician, with which as a result it still has the closest of affinities. The inhabitants of Portugal currently number some 10.5 million, and metropolitan Portuguese is generally said to have two principal dialect groups, northern and southern (broadly reflecting the different times at which Moslem occupation ended), with transitional varieties spoken in the provinces of Beira Alta and Beira Baixa. From Portugal, the language was taken to Brazil, the effective colonisation of which got under way from the middle of the sixteenth century, and has gradually become the native language of almost the whole of the population, currently estimated to be in excess of 150 million. As elsewhere in Latin America, there was in Brazil a prolonged period of interaction between speakers of Portuguese and speakers of indigenous languages, the most important of the latter being Tupi, one dialect of which, like Quechua and Guaraní noted earlier, came to be used for missionary and other purposes well outside its own original territory. This *lingua geral*, however, has not persisted, and was already losing ground to Portuguese by the end of the seventeenth century. Portuguese in Brazil, influenced by the diverse origins of both the immigrants and the administrators sent from Lisbon, rapidly developed norms of its own, particularly in the more popular registers. The overall position is that while the official and literary standards on both sides of the Atlantic do vary, not least because of the changes which took place in metropolitan but not Brazilian Portuguese from the seventeenth century onwards, apparently as part of a process of fairly conscious linguistic distancing from Castilian, ease of communication ensures that this variation is kept within limits; no such constraints affect common speech, however, in which divergences at all linguistic levels can readily be perceived. Again as elsewhere, there have been attempts to demonstrate that the divergences between Brazilian Portuguese and that of Portugal are due to the influence either of Tupi and/or of the Portuguese-based creole (see below) which developed subsequent to the importation of black slaves; but whereas as usual no influences other than on the lexicon have been established to general satisfaction in respect of the standard language, the widespread simplification of suffixed morphology in particular in spoken Brazilian Portuguese is strongly reminiscent of a typical result of the process of creolisation.

Brazilian Portuguese is, for the reasons we have indicated, relatively homogeneous, although there are differences between a northern and a southern group of dialects. This division apparently goes back to the

pattern of early settlement and to subsequent patterns of economic and cultural evolution.

After Brazil, the biggest concentration of speakers of varieties of Portuguese and of Portuguese-derived creoles outside Portugal is in Africa. We should, however, note first the two Atlantic archipelagoes of Madeira and the Azores, both colonised in the first half of the fifteenth century, with populations of some 300,000 and 350,000 respectively: their speech is generally grouped together as 'Insular Portuguese', and is fairly close to the European standard. The Portuguese of Madeira has a number of marked phonological characteristics and a clearly recognisable local intonation, while the Azores embrace a number of varieties of Portuguese, that of São Miguel being the most distinctive.

Within Africa itself, one thinks first of the former Portuguese colonies of Angola and Mozambique, with populations of some eight million and 12 million respectively. Recent estimates suggest that somewhere between a quarter and a half of these populations, particularly those in the major urban areas, speak at least some Portuguese, the norms of the more educated being essentially those of European Portuguese but with an admixture of features more reminiscent of the Portuguese of Brazil, whether through common divergence from the metropolitan norm or from the effects of (partial) creolisation or both. Creoles today appear to be of limited incidence, but one does find references, particularly in respect of Angola, to the existence, above all around the towns and cities, of varieties often known collectively as *pequeno português*, parallel to the *petit-nègre* or *petit français* of Francophone Africa mentioned below (p. 16).

One group of Portuguese-based creoles is to be found in the Cape Verde islands and on the nearby mainland, in Guiné-Bissau. The islands were colonised in the fifteenth century, and served as the centre where slaves were collected, auctioned and despatched; hence they necessarily played an important role in the formation and diffusion of Portuguese-derived creoles, particularly to Brazil. Among the present population of some 250,000 one finds a linguistic spectrum ranging from near-standard Portuguese to a fully fledged creole (though decreolisation is apparently well advanced). Of the two principal varieties of creole, one, Sotavento, is very similar to that of Guiné-Bissau, with over half a million inhabitants, most of whom use this *crioulo* as their language of everyday communication, as do some 60,000 persons in neighbouring Senegal. Further south in the Gulf of Guinea are the islands of São Tomé and Príncipe, with some 70,000 inhabitants, whose creole is described as 'increasingly lusitanised' by one observer, and Pagula (formerly Annobon), whose creole shows clear signs of Spanish influence.

Looking further east, there were significant Portuguese settlements in India and Ceylon. Leaving aside such former colonies as Goa, with some half a million inhabitants, where the local Portuguese is, or was, fairly close

to the European standard, one finds remnants, or historical records, of Portuguese-derived creoles along much of the Indian coastline, and again much further east, in Timor and Java (and indeed many other islands in this region), in Malaysia, where there are still in Malacca some 3,000 speakers of a creole known as Malaqueiro (or similar names) and a community of speakers of a related variety in Singapore, and in Macao, where standard Portuguese appears largely to have ousted Macaísta, although this survives as a popular medium of communication among some 2,000 descendants of creole speakers who moved from Macao to Hong Kong.

The 'oriental' Portuguese-based creoles, formerly very much more widely diffused than at present, display a sufficient number of points of similarity with those of Cape Verde and Guiné-Bissau, noted above, for many to argue that they share a common 'pidgin' or 'proto-creole' ancestor, which has undergone relexification to a greater or lesser extent in different actual contact situations, to yield the varieties at present in existence. (A similar argument, the monogenesis theory, has been advanced more generally, cf. p. 425.) It can certainly be demonstrated that there were significant trading and other links between the Portuguese-creole-speaking areas, and while the observed similarities may with some plausibility be attributed to the directly comparable historical and sociolinguistic circumstances in which pidgins arise and are then creolised, there was certainly a Portuguese-based *lingua franca* in common use throughout the relevant period.

One should note also two other Portuguese-based creoles found in the Americas: one, Papiamentu, spoken by upwards of 200,000 speakers on Curaçao and another, Saramacano, spoken on the nearby mainland in upper Surinam. These creoles are both in some ways atypical, the former because of complex influences on its lexicon (above all from Spanish, but also from Dutch and to a lesser extent English) and the latter because it manifests an unusually high proportion of words of African origin. In addition, there have been reports of Portuguese-derived creoles in Brazil itself among rural speakers of African origin, for instance Fronteiriço spoken on the Brazil–Uruguay border, though these are apparently on the point of extinction if not already totally lost. Finally, one should note the existence of small Portuguese-speaking settlements in both the USA and in Canada, in all of which language-shift is apparently well advanced.

Catalan

Catalan, with some six million speakers, has not experienced the great overseas expansion of its two sister languages within the peninsula. We have already noted that Catalan, from its homeland in the medieval Counties, including almost all the present-day French department of *Pyrénées Orientales*, was carried as far south as Alicante during the Reconquest (albeit with little headway on its western flanks, despite a lengthy

period of political union with Aragon from 1137), and that the inhabitants of the Balearics (conquered between 1229 and 1235) and a declining number of older inhabitants in the Sardinian town of Alghero (occupied in 1355) are Catalan-speaking to this day. It has to be said, however, that Castilian has made substantial headway in the *País Valencià*, particularly in Alicante and the other southern coastal resort towns. Recent surveys suggest a figure of some 40 per cent of Catalan speakers in the *País Valencià*, and 50 per cent in Catalonia proper, where there has been very high immigration of Castilian speakers from the south of Spain into industrial towns around Barcelona over the last thirty years or so. In the province of Tarragona, south of Barcelona, on the other hand, some 70 per cent are speakers of Catalan, while to the north the number is as high as 80 per cent.

The language is fairly homogeneous, although two principal dialect groups are generally distinguished, western (which includes Valencian and the eastern fringe of Aragon) and eastern, which includes the Catalan of Roussillon, the Balearics and Alghero but also, much more significantly, that of the great city of Barcelona. It is not exaggerating to say that it is above all the fidelity of the majority of the inhabitants of this city, of all social groups, to their native tongue which has ensured that its fate has been so unlike that of Occitan across the border in France; and indeed it is difficult to point to any language in Europe which has not become the official language of a nation-state which is as strongly placed as Catalan today.

The Romance of Gaul

As will be apparent from what has already been said, a major division developed within Transalpine Gaul between the French dialects of the north and centre (and part of modern Belgium), known collectively in medieval times as the *langue d'oïl*, and the Occitan dialects of the south, the *langue d'oc* (*oïl* (> *oui*) and *oc* being the markers of affirmation in their respective areas). The fate of Occitan is discussed below (pp. 16–17), as is that of Franco-Provençal (p. 17), the collective name for the dialects of a smaller intermediate area in the south centre of eastern France together with the varieties originally spoken in parts of Switzerland and the Val d'Aosta in Italy. The northern group of dialects is one of the most innovative branches of Romance, partly because of the intensity of the Germanic superstrate influence referred to earlier and partly because of radical changes within French itself in the post-medieval period.

French

Within the three major areas just noted, linguistic fragmentation continued, and a wide variety of dialects emerged, the principal ones being shown in Map V. One northern dialect was Norman, which has had such a profound

influence on the development of English, and from this source developed those varieties of French spoken in Jersey, Guernsey and Sark, estimated to have 15–20,000 speakers between them in the early nineteen-eighties; few if any of these, however, are under 40 years of age, and the loss of these forms of French spoken within the British Isles seems certain in due course. Another French dialect to emerge was Francien, the dialect of the Ile de France, and it is from this dialect that, once circumstances arose which favoured the growth of a national language, modern standard French has developed. The establishment of a fixed royal court in Paris, the development of an educational and of a legal system centred on that same city, and the fact that the abbey at Saint-Denis, close by, was in effect the spiritual centre of the kingdom were all factors which tended to favour the dialect of Paris and the surrounding area for the status of national language. Since the twelfth and thirteenth centuries, when Francien (a modern name) gradually came to be accepted as a norm to aim towards, at least in writing and in cultivated speech in northern and central France, its advance has been slow but steady, although, as we shall observe later, it was not until the nineteenth and twentieth centuries, particularly in the south, that French came to be so wholly dominant within the boundaries of France, at first among the bourgeoisie and in the cities, and later also in the remoter rural areas. Indeed, French's long period of predominance as the major international language of culture and diplomacy long antedates its general use as a spoken language within France: by the end of the seventeenth century, French had in effect replaced Latin in the former role, to the point that the Berlin Academy was able to ask, as a matter of fact, in 1782, at a time when Francien was the native tongue of perhaps a quarter of the population of France, 'Qu'est-ce qui a rendu la langue françoise universelle?' ('What has made the French language universal?'). This enhanced role for French persisted until the First World War and even beyond.

Within Europe, French is now spoken by some 51 million people within France (and Monaco), and by some four million Walloons in Belgium, principally in the four francophone districts of the south, Hainaut, Namur, Liège and Luxembourg, and in the bilingual district of Brussels the capital. Somewhat less than half a million people live in the Grand-Duchy of Luxembourg, where the native language of most speakers is a Germanic dialect but where French is the language of education and administration, whereas in Switzerland the most recent figures suggest that some 18 per cent of a total population of some 6.4 million, mostly living in the Suisse romande, are French speakers. In northern Italy, the Val d'Aosta has a population of around 100,000, some two-thirds of whom use French and/or a local variety of Franco-Provençal according to the register.

Outside Europe, indigenous French speakers are to be found in almost every continent. In Canada, there are some six million Francophone descendants of the original colonists, three-quarters of these living in the

province of Quebec, where they form some 80 per cent of the total population. Strenuous efforts are made to preserve and strengthen French, particularly in Quebec, within what has been since 1867 officially a bilingual country. Descendants of another group of French colonists in Acadia (the easternmost provinces of Canada), driven out in the mid-eighteenth century, carried their language southwards down the eastern seaboard of the United States, a few travelling as far as Louisiana, which had earlier been claimed for France by explorers coming southwards down the Mississippi. As a result, although there are relatively few French speakers in Acadia today except in New Brunswick (some 200,000), there are significant numbers — approaching one million — in New England, where there is a major admixture also directly from Quebec. Further, in Louisiana, a French possession until 1803, where as we have seen the immigrants were primarily from Acadia, and are indeed called 'Cajuns' (< *(a)cadien*), *français acadien* is in regular use by perhaps a further one million people, together with both a small elite speaking more or less standard French and also a French-based creole, sometimes known as Gombo, spoken by a declining number of people in eastern Louisiana and a small part of eastern Texas and earlier also in a few communities in Mississippi.

Elsewhere, French is generally in competition not with another European language but with indigenous non-European languages and/or with French-based creoles in former French (or Belgian) colonies. In the West Indies, French is found for instance in Haiti, where it is the official language of approaching five million people but where the great majority actually speak creole, and in various islands such as Martinique and Guadeloupe, where also French-based creoles have been documented and described; similarly, in Guyane, there are upwards of 50,000 creole speakers. Important also are the countries of the Maghreb (Algeria, Morocco and Tunisia), where French appears to be holding its own since independence: in Algeria, for example, it is estimated that some 20 per cent of the population can read and write French as a first or second language, with a much higher proportion able to speak it, above all in the cities. In black Africa, there are sixteen independent Francophone states comprising a great swathe across the west and the centre of the continent from Senegal to Zaire, together with the Malagasy Republic. There is a further group of French and French-creole-speaking islands in the Indian Ocean, for example Mauritius (approaching one million speakers), Seychelles (c. 40,000 speakers) and Reunion (450,000 speakers). In most of these countries, the future of French as a second language, used for a variety of official, technical or international purposes in place of one or more indigenous languages, seems secure. In Syria and Lebanon, however, the use of French as a second language has declined greatly, while in Vietnam, Laos and Cambodia, French and French-derived creoles appear to have been almost entirely lost.

Like all languages with any significant degree of diffusion, French is not a single homogeneous entity. Just as in France itself there is within most regions a spectrum of variation from 'pure' patois (the original local dialect, now often moribund) through *français régional* (largely the standard grammar, with a more or less regionally marked phonology and a greater or lesser number of non-standard lexical items) to the standard language (which itself has a wide range of styles and registers), so too one finds a similar spectrum in most if not all of the areas discussed above, often with the added dimension of a French-based creole. In Quebec, for example, one finds 'educated Quebec French' shading imperceptibly through to the fully popular variant known as *joual* (from the local pronunciation of *cheval* 'horse') associated primarily with Montreal. French-based creoles are spoken not only in Louisiana (alongside Cajun, discussed earlier), Haiti and various islands mentioned earlier, but arguably also in parts of black Africa, in the form of such variants as *petit-nègre* or *petit français*, though there the precise boundary between a pidgin and a creole is not always clear in practice. As in the case of *français régional*, there is very frequently a standard–creole continuum, with more educated speakers tending perhaps increasingly towards the metropolitan norms.

Occitan

The other major branch of Romance found within present-day France is Occitan, the generic name for all those varieties other than Franco-Provençal and Catalan spoken south of the major east–west line in Maps V and VI, forming a great swathe from Provençal through Lengadocian, Auvernhat and Lemosin to the very distinctive Gascon south of the Garonne. (This group of dialects is still at times referred to by English speakers as Provençal, this being a widely used name in the medieval period for the *koine* which was at that time a major literary language, with significant output, particularly in the twelfth and thirteenth centuries, of lyric poetry, narrative verse and prose: today, however, Provençal refers properly only to the local dialect of Provence.) While Old Occitan and Old French were certainly more similar than their modern descendants, not least because many of the radical changes which characterise French post-date the medieval period, there are in fact far closer parallels, both synchronic and diachronic, between Occitan and Catalan than between Occitan and French; as one observer put it, whereas a 'Proto-Occitan-Catalan' is a quite plausible concept, a 'Proto-Occitan-French' (excluding Catalan) certainly is not.

Forms of Occitan remained in general use in the southern part of the country until the end of the fifteenth century and beyond; indeed, only the edict of Villers-Cotterets in 1539 really ousted it (and indeed Latin) as an official written language, though by this time French was widely seen as having greater prestige, with all the consequences which follow from such

an attitude. Occitan remained however as virtually the sole everyday spoken language in its home territory until the new social and political climate which followed the Revolution, after which French made rapid headway, above all in the cities and among the upwardly socially mobile. Various attempts have been made to re-establish some form of Occitan as a literary language, most notably by the ninteenth-century Felibrige movement associated with Mistral, a movement hindered, however, by the essentially conservative and folklorist attitudes of its adherents. Attempts to form a standard written Occitan are also fraught with difficulty, given the diversity of the varieties of Occitan in use today. More recently, effort has been directed rather towards the restoration of local pride in spoken variants of Occitan, both in their own right and as vehicles for the maintenance and transmission of local culture, perhaps the one strategy with any hope of success. The fact remains, however, that there are today probably only some two to three million people still happy to converse in their own form of Occitan, none of these being monolingual. Various relaxations in the absolute hegemony of French within France have been tolerated since the last war, above all by virtue of the Loi Deixonne (1951) with subsequent amendments which permitted the teaching of local forms of language at all educational levels, as a result of which some 14,000 secondary school children were following courses in Occitan during the school year 1983–4. However, given that it is now common for parents to speak to their children in French rather than Occitan, there are very few true native speakers below the age of 40, a fact which must call into question the long-term future of all varieties of Occitan.

Franco-Provençal

Franco-Provençal is the name given to the group of dialects spoken in south-east central France, roughly in a triangle bounded by Grenoble, Geneva and Lyon, in Suisse romande and in the Val d'Aosta in Italy, thence shading fairly sharply into the Gallo-Italian dialects of the far north of that country. Sharing certain features of French, Occitan and indeed of Italo-Romance, these dialects, having lost their hold on the cities mentioned earlier, are now reduced to the status of patois, the formerly relatively unified language based on the usage of Lyon having fragmented. It is worth noting that the separate treatment accorded to Franco-Provençal is due at least as much to its geographical diffusion over three countries and its characterisation as a zone of linguistic transition as to its linguistic distinctiveness; Gascon, for instance, is notably more different from the other forms of Occitan, of which it is nevertheless traditionally treated simply as one variety.

Italian, Sardinian, Rhaeto-Romance

Crossing Italy to the north of Florence, one finds a major phonetic isogloss relating to the voicing or otherwise of Latin intervocalic voiceless plosives. Forms of Romance found north and west of this line, traditionally referred to as the La Spezia–Rimini line, generally show voicing in this environment while those to the south and east do not. (Consider for example the derivatives of Lat. SALUTARE 'to greet', which gives Fr. *saluer* (with the consonant lost entirely), Sp., Cat., Occ. *saludar* (with [ð] in the first two and variably [d] or [ð] in Occitan) but It. *salutare* and Rum. *a sărută* (the modern meaning in the latter case being 'to kiss').) Within Italy, one finds to the north of this line the Gallo-Italian dialects (Piedmontese, Ligurian, Lombard and Emilian), which have already been mentioned as shading into Franco-Provençal and thence into the dialects of present-day France, and also Venetian, which is linguistically, although not geographically, transitional between the Gallo-Italian dialects and Tuscan. To the south of the line, where voicing does not typically occur, one must distinguish on the one hand Tuscan and on the other hand a range of central and southern dialects from Umbria through the Abruzzi to Campania, Calabria and Sicily, with Sardinian so distinct as to warrant the separate treatment accorded to it in this volume (Chapter 9). Apart from the La Spezia–Rimini line itself, there are no abrupt divisions between these dialects, but rather a spectrum the ends of which are markedly distinct from one another. Partly within Italian territory and partly in Switzerland, one finds those forms of Romance traditionally labelled Rhaeto-Romance, similar in many ways to the dialects of northern Italy but sufficiently distinct from them to deserve description in their own right (Chapter 10).

Italian

The position of the dialects in Italy today is much more solid than elsewhere in Romance-speaking Europe. While one can truthfully say that from the time of Dante, Petrarch and Boccaccio — that is, from the early part of the fourteenth century — Tuscan, and in particular the Tuscan of Florence, came to be firmly established as the literary language of Italy, this had little or no effect on the speech even of educated people elsewhere in the peninsula, except in the capital city, where the usage of the Papal court appears to have influenced the local educated norm as early as the fifteenth century.

With this one major exception, there was no historical process comparable to the Reconquest in Spain and Portugal and no socially cohesive pressure such as that experienced in post-Revolutionary France to lead to the diffusion of Tuscan outside its home territory other than as the literary language of a very small minority. While the *questione della lingua* — the debate about the form of Italian most appropriate for literary usage — continued for centuries, it was not until after the political unification of Italy in

1861 that the question of a national language for educational and administrative purposes was seriously tackled. Despite the controversial recommendation of a commission headed by Alessandro Manzoni that the basis of this new truly national language should be contemporary Florentine, it was in fact, perhaps inevitably, literary Tuscan which was disseminated through the school system, a tendency reinforced by a strong prescriptivist tradition which has only begun to recede within the last couple of decades. It should at all times be borne in mind that in Italy, more perhaps than anywhere else in Europe today, there is a gradation from the indigenous dialects of isolated rural communities through regional dialects to regional standard Italian, with the national standard language superimposed and with the vast majority of native speakers able to practise code-switching over at least part of this range. Certainly, a number of regional standards are still in everyday use by even the most educated of speakers, reflecting the deep regional loyalties and long-standing cultural traditions of the different parts of what is now Italy. (The survival of Catalan, it will be recalled, has been attributed to similar factors.) As against this, however, the normal pressures exerted by a national education system, by military service — a great linguistic leveller — and by increased geographical and social mobility together, in the twentieth century, with the all-pervasive influence of radio and television, have contributed to the fact that standard Italian is now understood virtually everywhere in a country where it is only very recently that, for the first time, more than half the population claim 'Italian' rather than a regional variety as their native language.

It has recently been suggested that Roman may be tending to supplant Florentine as the basis of the standard language. Such a view, however, would be rather too dramatic. We noted above that educated Roman speech was to a very considerable extent Tuscanised well before the process was felt elsewhere in the country, the indigenous centre-south dialect, *romanesco*, being downgraded socially and increasingly limited to rural areas surrounding the city. When one talks, therefore, of 'Roman' in this context, one is speaking in effect of a form of speech not so different from the standard, but with certain originally local features of phonology and lexis, which is certainly now widely diffused over the national radio and TV networks.

Italian today, in the very broad sense outlined here, that is, including all the dialects, is spoken by some 60 million Italians within Italy and San Marino, only some three-quarters of a million now having some other language as their mother tongue (see p. 14 for French, p. 21 for Ladin and Friulan, p. 13 for Catalan). Non-Romance languages spoken within Italy include German in the Alto Adige (South Tyrol), Greek in Puglia and Calabria, Albanian, Slovenian and Serbo-Croat (see Map VII). Italian is spoken also by some 10 per cent of the population of Switzerland, including around 250,000 people in one Italophone canton, Ticino, by a number of

people in Corsica (see next section) — where, however, the official language has been French since 1769 — and in not insignificant Italian-speaking communities of Venetian origin in Istria and Dalmatia, now within the borders of Yugoslavia. In Malta, the popular spoken language has always been a local form of Arabic, but Italian was an official language until 1934, although it had long been losing ground to English, a process accelerated by the Second World War. In recent years, the availability of Italian television in the island has tended to restore at least a passive knowledge of the language to many Maltese. Elsewhere there is a large Italian-speaking minority of some four million in the USA, second only to the even larger group of Hispanophones noted earlier, and sharing with them a renewed interest in their linguistic and cultural heritage. Italian is spoken also by a declining linguistic minority in Eritrea, where it also forms the basis of probably the only surviving Italian-based creole. We have already mentioned the Italianised Spanish of Buenos Aires, Cocoliche, and there is apparently a rudimentary Italian-based creole known as Fazendeiro which is, or was, spoken in São Paulo in Brazil. These are both reflexes of significant Italian immigration into various South American countries, where recent figures suggest one and a half million italophones in Argentina and half a million in Brazil; there are also at least half a million Italian speakers in each of Canada and Australia. Italian does not, however, have anything like the same degree of diffusion across the world as Spanish, Portuguese or French.

Sardinian

Sardinian, the most conservative of all the Romance languages in a number of respects, is spoken by some one million people, all (apart from emigrants) within the island of Sardinia. The inhabitants of this island were largely divorced from the historical and cultural development of the former Roman Empire from the end of the fifth century, and the language which developed locally was used for almost all purposes until the end of the fourteenth century. Since that time it has been rivalled by various forms of Italian, by Catalan, by Spanish and latterly by standard Italian for all purposes other than everyday speech and here too, except perhaps in the remotest rural areas, it is now losing ground, especially among younger speakers. The net result has been on the one hand the implantation of non-Sardinian forms of Romance in certain areas and on the other hand the failure of any one dialect to emerge as standard Sardinian.

The first of these factors accounts for the small Catalan-speaking settlement at Alghero mentioned earlier, for two Genoese-speaking settlements, Carloforte and Calasetta, on islands off the southwestern corner of Sardinia, and for the fact that the two most northerly dialects, those of Gallura and Sassari, are so heavily influenced by Tuscan as to be best regarded as variants of Italian rather than Sardinian (p. 314). (The same applies to an even greater extent to the Italian of Corsica, formerly very

similar to Sardinian.) The second factor is responsible for the existence in the remainder of the island of three principal dialects, Campidanese (spoken over most of the south), Nuorese (centre and east) and Logudorese (north-west), none of which can really claim to be predominant (see Map IX). In recent years, in Sardinia as elsewhere, there has been a revival of interest in local languages and cultures; nevertheless, the long-term future of Sardinian as such looks far from secure (p. 349).

Rhaeto-Romance

The name Rhaeto-Romance is that conventionally given to a number of Romance speech forms found in the eastern part of Switzerland and northeastern Italy, characterised more by their differences from the major Romance language groups than by a set of shared features common only to themselves (p. 35). Indeed, it is certain that the areas in which these tongues are spoken have never formed a single administrative unit, and virtually certain that they have never been a homogeneous linguistic or cultural entity either. Equally, it is not possible to point to a common substrate or a common superstrate which would justify the tradition of treating these dialects together. It is perhaps simplest to think merely in terms of a set of dialects most closely related to those of the north of Italy, with each one showing a particular subset of a group of both conservative and innovative features said to characterise Rhaeto-Romance. Chapter 10 opens with an attempt to indicate these characteristics, but proceeds to talk on almost every page of the divergences between one form of Rhaeto-Romance and another.

There are three principal subtypes of Rhaeto-Romance (p. 351). Firstly, in the Swiss canton of the Grisons, one finds some 40,000 speakers of what is usually known as Romantsch, a number of related dialects of which the best known is probably Surselvan. Romantsch has been a 'national' (though not an 'official') language of Switzerland since 1938, and considerable work has been done on codifying the dialects and providing them with a standardised orthography; they are also taught in both primary and secondary schools. The fact remains, however, that the pressure of German, both standard and regional, is strong and incessant.

Secondly, there are, around Ampezzo and Bolzano in a number of valleys in the Dolomites and the Alto Adige in the eastern part of central northern Italy, something in excess of 10,000 speakers of Ladin. Much more significant, however, are some half a million speakers of Friulan, spoken in the region of Friuli-Venezia Giulia, around the cities of Udine and Gorizia. Neither of the forms of Rhaeto-Romance spoken within Italy enjoy the status or protection of Romantsch, and the influence of standard Italian and, in the case of Friulan, of Venetian also, is ever-present, and growing.

All in all, then, there are somewhat less than 600,000 speakers of

Rhaeto-Romance in a limited and non-contiguous geographical area, of whom probably none are now monolingual. The pressures of two high-status languages — German and Italian — on varieties which are essentially isolated from each other, coupled with the lack of any official support for what is numerically by far the most important variant probably bode ill for the long-term future of any member of this group of Romance languages, although some or all of them may well survive for a considerable time in informal spoken usage.

Balkan Romance

Dalmatian

Before passing on to look at the last major branch of the Romance family of languages, we should glance briefly at a group of dialects which has fared notably less well than Rhaeto-Romance, to the point of being totally lost at the end of the nineteenth century. On the coastal areas of what is now Yugoslavia and on the offshore islands, there existed a form of Romance generally known as Dalmatian, which was in a number of ways structurally intermediate between Italo- and Daco-Romance. This form of Romance, which should be confused neither with (Venetian) Italian introduced into Istria and Dalmatia from the Middle Ages onwards as a result of trade and settlement (see p. 20) nor with the tiny output of Istro-Rumanian found not far from Rijeka (see below), is best represented in records from Dubrovnik and the surrounding areas. Even the earliest records, however, show Venetian influence, and the indigenous Romance speech of that area is no longer attested after the end of the fifteenth century, ousted by the combined pressure of Venetian on the coast and Serbo-Croat inland. The last place where this form of Romance survived was apparently the island of Krk (Veglia), and the standard description of the local Dalmatian dialect, Vegliote, published in 1906, is based on personal interrogation of its last-known speaker — or rather of a son who claimed to remember it well!

Rumanian

The history of the easternmost branch of the Romance family of languages is rather more complex than that of the varieties we have discussed hitherto. Deriving from the Latin spoken in the Roman province of Dacia — and hence often known as Daco-Rumanian — the antecedent of modern Rumanian, although mentioned as early as the thirteenth century, is attested in texts only from the sixteenth century, with consequential uncertainty as to its precise history during the preceding thirteen centuries or so. In essence, there are two views about the persistence of a Romance tongue in what is today Rumania: one is that Latin was preserved without a break

north of the Danube, although this province was abandoned by Rome as early as AD 271; the other is that Latin was lost in this region but later re-introduced by Romance speakers from south of the Danube (an area much longer in Roman hands), as these migrated northwards under pressure from incoming speakers of Slavic languages. The second view is broadly speaking more plausible, given both the fact that the northern province was one of the last to be occupied and one of the first to be abandoned and also the dispersal of pockets of Romance speakers, as we shall see, well outside the boundaries of present-day Rumania, a distribution consistent with northward migration. This does not, of course, in any way rule out the possible persistence of a form of Romance north of the Danube, which would have reinforced, or been reinforced by, an influx of Latin speakers from the south.

Balkan Romance is generally divided into four principal types, of which one, Daco-Rumanian, is, as we have seen, the antecedent of modern Rumanian (see Map XI). The three other principal sub-Danubian variants are Istro-Rumanian, spoken by fewer than two thousand speakers around Ucka Gora in the eastern part of the Istrian peninsula not far from Rijeka, Megleno-Rumanian, also spoken by a few thousand speakers north-west of Salonika in Greece, and Arumanian (Macedo-Rumanian), spoken by far more — some 350,000 — in northern Greece, parts of Albania and south-western Yugoslavia. This group is not helped by the fact that it is spread over three countries, one of these being virtually cut off from the other two for political reasons, and that it has no single focus or national base, or official support in any of the countries in which it is spoken.

The northern branch of Balkan Romance has prospered considerably, Rumanian today having some 21 million speakers. There are two principal dialects: Moldavian, spoken in the northern part of the country and indeed right up to the Dniester river within the two Soviet republics of Moldavia and Ukraine, and Muntenian (Wallachian), spoken in the south of the country and underlying the literary language developed during the latter part of the eighteenth and the first part of the nineteenth century and based on the language of Bucharest. There are those who argue also for a third dialect, spoken in Transylvania, but most analysts prefer to see this rather as a transitional area between the two principal dialects.

Standard Rumanian shows no readily identifiable substrate influences but does of course bear very considerable marks of its long period of interaction with Slavic. (For lexical borrowings from other sources, see pp. 414–16). Rumanian also has more than one morphosyntactic feature characteristic of the Balkan *Sprachbund*. During the nineteenth century, serious attempts were made as a result of a (relatively brief) period of 'Romance nationalism' to reduce the Slavic element in the lexicon, but changes were largely limited to specialist registers, with little effect on everyday vocabulary; one countervailing pressure favouring Slavic has always been that of

religious language, Old Church Slavonic being the liturgical language of the orthodox church in the Balkans. One lasting effect of the pro-Romance movement, however, was the change within Rumania from the Cyrillic to the Latin alphabet, begun early in the nineteenth century and completed by the time of the union of the principalities of Moldavia and Walachia in the early 1860s.

Rumanian, then, is spoken by some 21 million people, in Rumania (18 million), in adjacent parts of the USSR (where it is often claimed to be a separate language and the Cyrillic alphabet is used) and of Yugoslavia (the Banat), and in a number of villages just across the borders of Bulgaria and Hungary. In addition, there are substantial communities of Rumanian speakers abroad, in particular in Australia and the USA, but Rumanian was not involved in the great period of colonial expansion, so there are no areas speaking Rumanian or Rumanian-derived creoles in the New World.

Conclusion

We have briefly surveyed in this chapter some of the historical, social and cultural factors underlying both the present patterning and the current distribution of the Romance languages as we see this today. Some branches of Romance — for instance in Britain or north Africa — died out before the attested emergence of any local vernacular; one other, Dalmatian, has died out almost within living memory, and the long-term future for several others must be at best doubtful. On the other hand, while none of the Romance languages can rival the claims of English to be the international language of the second part of the twentieth century, it should not be forgotten that among the dozen or so languages in the world with the greatest number of speakers, one finds no fewer than three Romance languages, with a continuum of variants from the metropolitan norm through regional standards to more or less creolised local forms. What the Roman Empire did for Latin, colonial expansion and contemporary ease of travel and communication have done for Romance. It is a more detailed examination of these languages severally to which we shall now turn.

Bibliography

Agard, F.B. (1984) *A Course in Romance Linguistics* 2 vols. Georgetown University Press, Washington DC.

Anderson, J.M. and J.A. Creore (eds.) (1972) *Readings in Romance Linguistics.* Mouton, The Hague.

Bal, W. and J. Germain (1982) *Guide bibliographique de linguistique romane.* Cabay, Louvain.

Bourciez, E. (1967) *Elèments de linguistique romane* 5th edn. Klincksieck, Paris.

Dittmar, N. and B. Schlieben-Lange (eds.) (1982) *Die Soziolinguistik in romanischsprachigen Ländern.* Narr, Tübingen.

Elcock, W.D. (1975) *The Romance Languages.* Faber and Faber, London.
Hall, R.A. Jr (1974) *External History of the Romance Languages.* Elsevier, New York.
Hope, T.E. (1971) *Lexical Borrowing in the Romance Languages* 2 vols. Blackwell, Oxford.
Iordan, I., J. Orr, and R. Posner (1970) *An Introduction to Romance Linguistics.* Blackwell, Oxford.
Kontzi, R. (ed.) (1982) *Substrate und Superstrate in den romanischen Sprachen.* Wissenschaftliche Buchgesellschaft, Darmstadt.
Lüdtke, H. (1956) *Die strukturelle Entwicklung des romanischen Vokalismus.* Romanisches Seminar an der Universität Bonn, Bonn.
Malkiel, Y. (1978) 'The Classification of Romance Languages'. *Romance Philology* 31: 467–500.
—— (1978) 'Factors in the Unity of Romania', *Romance Notes* 18: 263–71.
Manoliu-Manea, M. (1985) *Tipologia e historia: elementos de sintaxis comparada románica.* Gredos, Madrid.
Posner, R. and J.N. Green (eds.) (1980–2) *Trends in Romance Linguistics and Philology* 4 vols. Mouton, The Hague.
Rohlfs, G. (1971) *Romanische Sprachgeographie.* Beck, Munich.
Sampson, R. (1980) *Early Romance Texts: An Anthology.* Cambridge University Press, Cambridge.
Tagliavini, C. (1972) *Le origini delle lingue romanze* 6th edn. Patron, Bologna.
Vincent, N. and M. Harris (eds.) (1982) *Studies in the Romance Verb.* Croom Helm, London.
Weinrich, H. (1958) *Phonologische Studien zur romanischen Sprachgeschichte.* Aschendorffsche Verlag, Münster.

2 Latin

Nigel Vincent

1 Introduction

Latin differs most obviously from the other languages that are dealt with in this volume in that it is a dead language, for which direct recourse to the phonetic output and intuitive judgements of native speakers is by definition not available. Nor, of course, is it possible to conduct surveys into the correlation of linguistic and social variables, language attitudes and the like. Rather, our knowledge of Latin grammar, lexis and phonology has to come from the evidence of surviving texts, admittedly available in generous quantity and over a long period, supplemented by the occasional explicit comments of native grammarians and authors. To this may be added the data retrievable by the backward projection, through the techniques of reconstruction, of the evidence of the daughter languages, these in turn offering us a combination of living data and a variably rich textual tradition. The integration of the two sorts of evidence, attested and reconstructed, takes place in the further context of those general principles and constraints that have emerged, and are still emerging, from the various subdisciplines of general linguistics. For our understanding of the social, cultural and ethnic factors governing the use of Latin, we must turn to the evidence of history, where once again the surviving records have to be matched against the current state of affairs and interpreted in the light of our general knowledge of social, political and economic processes.

A further caveat is necessitated by the fact that the term Latin is often employed as a convenient designation for a broad range of related but distinct varieties. Thus, in origin, Latin refers to the speech of the city-state of Rome as it emerged towards the middle of the first millennium BC (p. 1); yet it may still be appropriately used a thousand years later as the name of the dominant administrative language of the Roman Empire at the time of its collapse, conventionally taken to date from the deposition of Romulus Augustulus by the Huns in AD 476. Add to this the social and geographical variation also discussed earlier (pp. 2–5), and it will be clear that the potential for terminological confusion is considerable. The traditional means of resolving the ambiguities of the term Latin has been to

use the expression Classical Latin to refer to the language of the educated classes of Rome in the period of Cicero and Caesar (first century BC) and a variety of other labels — Late, Popular, African, etc. — to distinguish different eras, regions and social levels as necessary. Of these alternative epithets, the most abused by far has been 'Vulgar', which, although it might seem on etymological grounds to delimit a social variety, has come to be used by scholars in such a wide range of senses as to be virtually useless. In this book, therefore, we shall follow both the more ancient and the more modern practice and employ the undifferentiated term Latin. Only in cases where ambiguity or confusion seems likely to arise will more precise (and transparent!) designations be adopted.

Our aim, then, in the present chapter is to narrate the structural transformations that Latin underwent in the course of the passage towards the modern Romance languages. Unfortunately, this task is not made any easier by the fact that the chronology of the texts, from which a large part of our evidence is of course drawn, does not necessarily reflect what must be presumed, on internal criteria, to be the natural sequence of linguistic developments. Thus, the evidence of an early but innovative writer like Plautus (254–184 BC) will often tell us more about the sorts of direction linguistic evolution must have taken by the fifth century AD than we could learn from a contemporary but conservative writer such as Boethius (AD 480–524). It will be important to remember, therefore, in what follows that there are two distinct stages involved in describing the linguistic history of Latin. First, the evidence, culled from such sources as are available, has to be integrated into a linguistically plausible diachronic sequence; only at a second stage can that sequence be projected back onto a sociolinguistic model of Rome and the Empire at the various stages of its history. Of course, much of this latter task remains to be done, and it should be clear that the present chapter is concerned primarily with the former exercise.

As an example of the separability of internal and external factors, consider the development of the tonic vowel system. Structurally, we must explain as far as possible how a system based on a phonemic opposition of vowel quantity with predictable assignment of word stress came to be replaced by a number of systems where the crucial factors were vowel quality and free, or partially free, stress. Such an account can, in part at least, stand or fall independently of our ability to date precisely the changes involved, or to say when and where there was a historical overlap of systems, or to be able to identify sociolinguistic parameters of stratification, even though these are all important questions, to which answers must ultimately be sought. Similarly, in the realm of grammar, a strictly linguistic history needs to explicate and, if possible, explain how the classical 'accusative and infinitive' with verbs of saying and thinking (p. 67) came to be replaced by the emergent pattern of finite complementation with QUID/QUOD + embedded sentence. An increased use of the former over the

latter in the prose of a Boethius or a Claudian will simply be taken as an index of their conservative usage, not as indicating the need to revise our relative chronology of the two constructions.

Two dangers, apart from that of incompleteness, threaten too strict a separation of internal and external history. The first is that we might take as a starting point or intermediate stage of a diachronic linguistic trajectory a usage which never existed outside the confines of the written page or the rhetorician's manual. Consider, for example, the freedom of adjective position, and in particular the distance between noun and associated adjective, as derivable from an analysis of classical poetry. On the one hand, it might be argued that this is no more than a literary artifice; certainly, the position of the adjective seems to be a good deal less free in prose. On the other hand, the position of the adjective is considerably more constrained in all of the modern Romance languages than in even the most colloquial of prose. It must also be said, moreover, that a number of the world's languages seem to exhibit positively Virgilian degrees of freedom even in the absence of a written tradition, so we must be on our guard against ruling out all poetic usage simply on the grounds that it is poetic usage.

The second danger alluded to above is that in some circumstances a structural argument will require the contemporaneity of two or more linguistic patterns. There is a grave risk of circularity if the only evidence for such contemporaneity is the necessity of completing the argument, and in such cases — say as regards the relative order of the pairs auxiliary + verb and verb + object — specific attention will have to be paid to establishing as accurately as possible the absolute as well as the relative chronology.

With these considerations and caveats in mind, we turn now to a treatment of individual aspects of Latin phonology, morphology, syntax and lexis. The general format of each section will be a synchronic description of the relevant Latin structures, followed by a diachronic account of how those structures were transformed into the patterns, call them Late Latin or Early Romance, which underlie the development of the modern Romance vernaculars, whose individual properties will form the contents of the succeeding chapters.

2 Phonology

The phoneme inventory of Latin is set out in Table 2.1. Latin orthography provides a one-to-one representation for the phonemes of the language, except for its failure to indicate the systematically contrastive length of vowels. It has also, of course, provided the basis for most West European orthographies and thence for the International Phonetic Alphabet (IPA), so that given the sound–letter equivalences *c* = /k/ and *x* = /ks/ and making due allowance for the special cases to be discussed below, it is a

Table 2.1: The Phoneme Inventory of Latin

Consonants

p	b	t	d	k	g	k^w	g^w	
f		s						h
m		n						
		l						
		r						

Vowels								Diphthongs	
ī	ĭ					ŭ	ū	ai	(written *ae*)
	ē	ĕ			ŏ	ō		au	
			ā	ă				oi	(written *oe*)

straightforward matter to read off the phonemic structure from the spelling. Among the special cases, note first that we have listed /k^w, g^w/ (orthographically *qu, gu*) as separate phonemes rather than as clusters of /k g/ plus /w/. This decision is based partly on morphophonemic evidence (/k^w/ alternates with /k/ as in RELĪNQUO 'I leave' vs RELĬCTUS 'left', and /g^w/ alternates with /k/ and /w/ as in NĪNGUIT 'it snows' vs NĪX 'snow (nom. sg.)' and NĪVEM 'snow (acc. sg.)' and partly on phonotactic evidence (initial clusters of the form /Cw/ are otherwise absent in the language apart from a handful of words in /sw/ such as SUĀDEO 'I urge', and SUĀVIS 'sweet'). Next, there is the problem of the status, phonemic or otherwise, of [ŋ], generally agreed to be the phonetic value of the *g* in the orthographic sequence *gn* in DĬGNUS 'worthy', RĒGNUM 'kingdom', etc. (and perhaps also in *gm*, as in TĔGMEN 'cover', etc.). Minimal pairs such as AGNUS 'lamb' [aŋnus] vs ĂNNUS 'year' [annus] on the one hand and AGGERE 'earthwork (abl.)' ['aggere] vs ANGERE 'to choke' ['aŋgere] on the other are sometimes adduced to show that [ŋ] cannot be an allophone of either /n/ or /g/ and must therefore be accorded phonemic status in its own right. It seems preferable, however, to follow the lead of the orthography and say that we have here a case of partial overlapping with [ŋ] as the allophone of /n/ before velars and of /g/ before nasals. The orthography also seems to be right in treating as variants of /i u/ the values [j w] which must be attributed to *i* and *u* (the systematic use of a separate lower case *v* is a Renaissance innovation) in words like IAM 'now', IŌCOR 'I jest', VŎLO 'I wish', VĀDO 'I go', etc. The conditioning factor here is the syllable structure: [i u] when the item is in nuclear position and [j w] in onset or (more rarely) coda position.

In the absence of any orthographic marking of vowel length, we have assembled some minimal contrasts in the chart of vowel length contrasts.

Vowel Length Contrast

(a) Lexical

mălum	'evil'	mālum	'apple (acc. sg.)'
lătus	'side'	lātus	'broad'
ĕsse	'to be'	ēsse	'to eat'
lĕgo	'I read'	lēgo	'I send as ambassador'
pŏpulus	'people'	pōpulus	'poplar'
ŏs	'bone'	ōs	'mouth'

(b) Morphological

rosă	'rose (nom. sg.)'	rosā	'rose (abl. sg.)'
(and similarly for all 1st decl. nouns)			
lĕgit	'he reads'	lēgit	'he read'
(and similarly for a number of verbs)			
gradŭs	'step (nom. sg.)'	gradūs	'step (gen. sg.)'
(and similarly for all 4th decl. m./f. nouns)			

(c) Lachmann's Law

dŭctus	'past part. of *duco*'	lūctus	'past part. of *lugeo*'
frĭctus	'past part. of *frico*'	frīctus	'past part. of *frigo*'
lĕctus	'past part. of *-licio*'	lēctus	'past part. of *lego*'
făctus	'past part. of *facio*'	āctus	'past part. of *ago*'

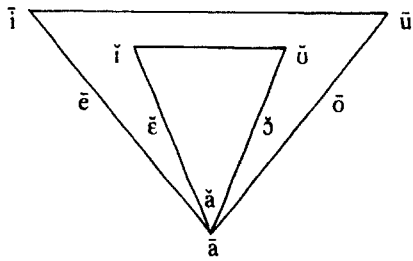
The examples given here show that length oppositions were functional both in maintaining lexical contrasts and in the working out of nominal and verbal morphology. Of particular interest are the length alternations introduced by so-called Lachmann's Law, according to which the stem vowel of the past participle is long if the root ends in a voiced consonant and short otherwise, even though the consonant voicing opposition is subsequently neutralised before the past participle suffix -TUS. Note, too, that vowel length in closed syllables, as in the examples of Lachmann's Law and elsewhere (cf. *nōsco* 'I discover' vs *pŏsco* 'I demand'), can only be discovered through the testimony of grammarians and through subsequent historical development, since both orthographically and metrically such quantities are 'hidden'.

Latin also had the three diphthongs of Table 2.1., which always counted as equivalent to long vowels. Of these, /oi/ was already recessive in the classical language, most instances having shifted to /ū/ (e.g. OLat. OINOS > ŪNUS), just as the Old Latin diphthong /ei/ had given /ī/ (OLat. DEICO > DĪCO). The cases of /oi/ that did survive, e.g. POENA 'punishment' (cf. PŪNIO 'I punish'), later shift to /ē/ and are thenceforth indistinguishable from primary /ē/ — thus Fr. *peine* 'suffering' (< POENA) beside *veine* 'vein' (< VĒNA). On the later fate of /ai, au/ see below.

Of crucial importance in integrating the consonant and vowel phonology of Latin were two suprasegmental factors: stress and syllable structure. The position of stress in Latin was predictable in terms of the segmental composition of the word. If we allow that a short vowel contributes one unit of

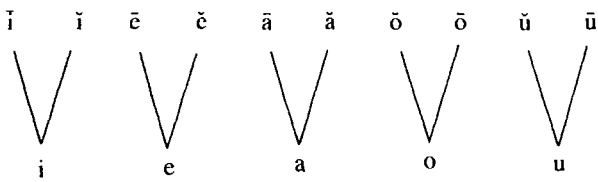
phonological 'weight', a post-vocalic tautosyllabic consonant likewise one unit, and a long vowel two units, and if we agree that a syllable with two or more units of 'weight' is heavy and a syllable with one unit is light, then we can easily see the logic behind the traditional rule that stress falls on the penultimate syllable if that is heavy, and otherwise on the antepenultimate syllable. Note, however, that such a formulation is not inconsistent with the existence of stressed light syllables — e.g. ĂNĬMA — since the antepenultimate must be stressed regardless, as must the penultimate of disyllables, e.g. CĂNIS 'dog', FĬDES 'faith', nor with the existence of so-called hypercharacterised syllables, having a long vowel plus a post-vocalic consonant (three units of weight), e.g. SCRĪPSI 'I wrote', MĪLLE 'thousand'. Nonetheless, there is evidence within the development of Latin of a tendency to eliminate such situations, either by shortening the vowel, e.g. ĂMĀNTEM > ĂMĂNTEM, or by reducing a geminate consonant, e.g. CAUSSA > CAUSA (a diphthong in Latin is always counted as two units and therefore long), CĀSSUS > CĀSUS, *DĪVĪDTOS > DĪVĪSSUS > DĪVĪSUS. Similarly, the process of iambic shortening, whereby a sequence ∪ – becomes ∪ ∪ is well attested, e.g. BĔNĒ > BĔNĔ, ĒGŌ > ĒGŎ, DŬŌ > DŬŎ, and serves to eliminate the anomaly of a long vowel after a stressed light syllable. Such tendencies eventually dictate a new norm in spoken Latin, according to which stressed syllables are never more nor less than two units in weight. The consequences are twofold: stressed short vowels in open syllables lengthen, e.g. CĂNEM > /ka:ne(m)/, and long vowels in closed syllables shorten, e.g. ĀCTUM > /aktu(m)/. We have deliberately cited words with [a] in this connection, since with this vowel no difference in phonetic quality accompanies a difference in quantity. This is easily seen in the fact that, for example, in Italian the pairs *cane* 'dog' and *pane* 'bread', *fatto* 'fact' and *atto* 'act' rhyme even though their Latin etyma differ in vowel length (CĂNEM/PĀNEM; FĂCTUM/ĀCTUM). In the case of the mid and high vowels, however, differences in length were accompanied by differences in quality, according to the general phonetic principle that short vowels tend to be laxer (more open) than their long congeners. Thus, the phonetics of the Latin vowel system is best represented as in Figure 2.1.

Figure 2.1: Phonetic Values of Latin Vowels



Each pair of non-low vowels is doubly distinguished — by quality and by quantity. Suppose now that, as a result of the changes described above, the differences of quantity become tied to the structure of the syllable. Two possibilities ensue. The first is that the loss of the quantity opposition entails the loss of the quality opposition since the two are linked. If that happens, the two triangles in Figure 2.1 will, as it were, be superposed, and where previously there were ten distinctive vowels, there will now only be five. Exactly this situation obtains in the Sardinian vowel system, also attested in a small area of Calabria (see below). Hence:

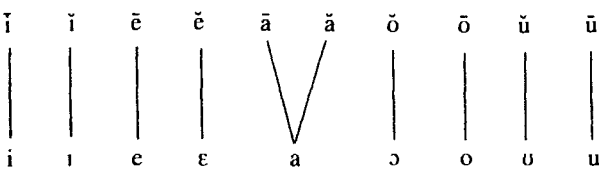
The Sardinian Vowel System



(For examples, see p. 317.)

The second alternative is that the quality contrasts will outlive the loss of the length distinction (except of course for the low pair /ă ā/), and we will get a nine-vowel system thus:

'Transitional' Late Latin Vowel System



Although this system does not survive as such, it is a necessary transitional stage in the development of the other Romance vowel systems. Evidence for early variation between the two possibilities can be found in examples such as CAUDAM > It. *c*[o]*da*, (contrast AURUM > It. [ɔ]*ro*); LŪRĬDUM > LŪRDU(M) > It. *l*[o]*rdo*, Fr. *lourd* (contrast NŪLLUM > It. *nullo*, Fr. *nul*); FRĪGĬDUM > FRĪGDU(M) > It. *fr*[e]*ddo*, Fr. *froid* (contrast MĪLLE > It. *mille*, Fr. *mille*). These developments suggest that after a change /au/ > [ɔ̄], the subsequent change may go with quantity, viz. [ɔ̄] > [o], or with quality, viz. [ɔ̄] > [ɔ]. Similarly, the long high vowels /u i/ may shorten (but not centralise) to give [ŭ ĭ] (not [ʊ ɪ]), wh ch subsequently develop to either [u i] or [ʊ ɪ], as the examples above testify.

The possibility of variation is eliminated, however, once the inter-predictability of length and quality is disrupted by the merger of /ī/ and /ē/, attested in early inscriptions at Pompeii and elsewhere in the southern half of Italy, and plausibly attributable to Osco-Umbrian substrate influence. The effect is that the reflexes of Latin pairs such as PĪRA/VĒRA rhyme in all the Romance languages:

Rum.	pere	vere
It.	pera	vera
Fr.	poire	voire
Sp.	pera	vera
Port.	pera	vera

The exception of course is Sardinian (see the diagram above), where we find *pira* but *bera*. A parallel merger in the back series of /ŭ/ and /ō/ is later and does not affect the Latin/Romance of the Balkans (pp. 391ff) and of a small area of Lucania. Thus, whereas the reflexes of Latin NŬCEM/ VŌCEM rhyme in Fr. *noix/ voix*, It. *noce/ voce* and Port. *noz/ voz*, they remain distinct both in Sardinian, *nuge/ boge*, and in Rumanian, *nucă/ voce*. The effects of these various mergers are set out in Table 2.2, where the Rumanian (or Balkan) system can be clearly seen to parallel the Sardinian system in the back series and what we will henceforth call the Proto-Western-Romance (PWR) system in the front series. The Sardinian

Table 2.2: Romance Vowel Systems

Latin	ī̆	ī	ē	ĕ	ă̆	ŏ	ō	ŭ	ū
Sardinian	i		e		a	o		u	
Balkan	i		e	ɛ	a	o		u	
PWR	i		e	ɛ	a	ɔ	o	u	

system therefore represents the 'default' development of the classical vowel system consequent upon the loss of distinctive quantity, but this is interrupted by two waves of merger spreading outwards from Rome and/or Campania at different times. That all three systems were copresent in the Italian peninsula is strongly suggested by the survival of the 'Balkan' and 'Sardinian' systems in residual dialect areas of Southern Italy, as Map VIII indicates. In dealing with Romance vowel systems, manuals traditionally cite a fourth, so-called 'Sicilian' system, where Lat. /ī ĭ ē/ all merge to /i/ and in parallel fashion /ū ŭ ō/ yield /u/, Thus, we find Sicilian rhyming triplets such as:

vinu	'wine'	(< Lat. VĪNUM)	luci	'light'	(< Lat. LŪCEM)
sinu	'breast'	(< Lat. SĪNUM)	nuci	'nut'	(< Lat. NŬCEM)
kinu	'full'	(< Lat. PLĒNUM)	vuci	'voice'	(< Lat. VŌCEM)

This pattern, however, is probably best viewed not as an independent system evolving directly from Latin but as a subsequent development in Sicily (and parts of Calabria) of Proto-Western-Romance, involving the merger of the Proto-Western-Romance oppositions /i e/ and /u o/. Its explanation is plausibly to be sought in the strong substrate/adstrate effects in this area of the pentavocalic Greek pattern.

Within this pattern of tonic vowel developments the two Latin diphthongs /ai/ and /au/ had rather different roles. The former for the most part monophthongised to /ɛ/, and in so doing must have helped to blur the correlation between vowel quality and quantity, since /ɛ/ was normally the reflex of Latin /ĕ/, whereas any vowel deriving from /ai/ (or /au/ for that matter) would retain the diphthong's inherent length. Indeed some scholars have gone so far as to argue that it is the monophthongisation of /ai/ which finally brings about the collapse of the Classical Latin quantity-based system. Unfortunately, the inscriptional evidence militates against this view, which in any case imposes a rather heavy burden on a minor sound change affecting a small number of items, and which even then is not without exceptions. Thus, It. *seta*, Sp. *seda*, Fr. *soie*, etc. 'silk' presuppose /e/ not /ɛ/ from Lat. SAETA, and note It. *preda* 'booty' with /ɛ/ from PRAEDA, where the French cognate *proie* requires an earlier /e/.

The diphthong /au/ by contrast resisted monophthongisation in a large part of Romania, including Rumanian, southern Italian dialects, most varieties of Rhaeto-Romance, Occitan and Portuguese (where it has subsequently become /ou/). Elsewhere, by a development parallel to that of /ai/, /au/ yields a long open vowel, /ɔ̄/. The exception is Sardinian, where we find /a/. Thus, Lat. TAURUM 'bull' > Rum. *taur*, Cal. *tauru*, Occ. *taur*, Port. *touro* vs Sp. and It. *toro* and Sard. *trau*, this last by metathesis from **taru* (cf. Sard. *laru* < LAURUM 'laurel'). In a few cases, however, even /au/ monophthongised early. Hence Lat. FAUCES 'jaws' > It. *foci*, Sard. *foge*, Occ. *fotz*, Sp. *hoz*, Port. *foz*, which presuppose Lat. /o/ and not /ɔ/, as do forms like It. *c*[o]*da*, Fr. *queue*, etc. from Lat. CAUDAM.

A much debated issue in Romance historical phonology has concerned the nature and unity of Romance diphthongisation. The facts are relatively straightforward, even if their interpretation is not, and may be encapsulated in the following two generalisations:

(a) a number of languages and dialects, including French, Rhaeto-Romance and northern Italian, have developed so-called 'falling diphthongs' (i.e. having a vowel nucleus as the first element followed by an offglide) from Latin stressed monophthongs, principally /ē/ and /ō/: e.g.

Lat. PĪLUM 'hair' > Fr. *poil* (subsequently [pwal] (cf. p. 211)), R-R *peil*, Vegl. *pail*; Lat. TĒLAM > Fr. *toile*, R-R *teila*; Lat. GŪLAM 'throat' > Fr. *gueule*, R-R *gula* (both with subsequent remonophthongisation); Lat. -ŌSUM 'adj. suff.' > Fr. *-eux*, R-R *-us* (again with remonophthongisation). Vegl. *-aus*.

(b) Many languages and dialects have undergone a process of metaphony, whereby the stressed vowel, usually, but not necessarily, /ĕ/ and /ŏ/, has diphthongised (often with subsequent raising and remonophthongisation) before unstressed, usually final, /i/ and/or /u/, the end product being a rising diphthong (i.e. an on-glide preceding a stressed nucleus): e.g. VĔNTUM 'wind' > SIt. *vientu*, *DĔNTI 'teeth' > *dienti* vs DĔNTEM 'tooth' > *dente*, or PŎRCUM (m.) 'pig' > Port. *porco* vs PŎRCAM (f.) > *p*[ɔ]*rca*.

This second type of diphthong is more widespread in Romance, and is generally thought to be older, in all probability attributable to a period before the break-up of Latin into the individual vernaculars. It also seems clear that the two phenomena must have different explanations, since metaphony is essentially a harmonising or assimilatory effect between a stressed and an unstressed vowel, and therefore occurs regardless of syllable structure, while falling diphthongs are best seen as due to phonetic lengthening, a conclusion endorsed by the fact that they are limited to open syllables. To see the historical problem, consider now the patterns of standard Italian and of Spanish:

Latin	*Italian*	*Spanish*	
BŎNUM	buono	bueno	'good (m. sg.)'
BŎNAM	buona	buena	'good (f. sg.)'
MŎRT(U)UM	morto	muerto	'dead (m. sg.)'
MŎRT(U)AM	morta	muerta	'dead (f. sg.)'
CAECUM	cieco	ciego	'blind (m. sg.)'
CAECAM	cieca	ciega	'blind (f. sg.)'
DEFĔNDIT	difende	defiende	'defends (ind.)'
DEFĔNDAT	difenda	defienda	'defends (subj.)'

Both languages have the rising (putatively metaphonic) diphthong, but Italian has it only in open syllables, the classic falling (non-metaphonic) diphthong environment. Spanish, by contrast, has the metaphonic type of diphthong in open and closed syllables, but regardless of the final vowel. The Spanish pattern can be explained fairly straightforwardly as being due to the analogical extension of the originally metaphonic pattern to all stressed syllables. The Italian, more strictly Tuscan, pattern is more puzzling. It can be quite neatly explained, however, according to Schürr, if one sees the diphthongs as borrowings from surrounding metaphonic dialects, specifically those to the north-west, with the borrowing being

mediated by native phonotactics that forbid long vowels (to which diphthongs are equivalent) in closed syllables. Once borrowed the diphthongs would then generalise by analogy to the non-metaphonic environments. Alternatively, one might see metaphony as original in Tuscan, but with the effects having been obscured by the local developments in syllable structure which filtered out those diphthongs which did not conform phonotactically. Either way it seems clear, as most scholars would agree, that we should seek a unified solution for rising diphthongs in Romance, though whether or not in terms of the operation of metaphony is more controversial, while treating the development of falling diphthongs under a separate heading.

Whereas the tonic vowels, despite the kinds of shift discussed in the preceding sections, tend to preserve their segmental identity as the nucleus of the stressed syllable in a word, the vowels in the unstressed or secondarily stressed syllables tend to be subject to a wide range of elisions, assimilations, reductions, and so forth, which it would be impossible conveniently to summarise in a short chapter such as the present one. We will content ourselves, therefore, with an indication of the main types of process involved. Two in particular were important:

(a) *Syncope*: the loss of the immediately pre- or post-tonic vowel, as in ŎC(Ŭ)LUM 'eye' > Fr. *oeil*, It. *occhio*, Sp. *ojo*, Rum. *ochi*; BŎN(Ĭ)TĀTEM > Fr. *bonté*, Cat. *bontat*, R-R *bundet*. The evidence for this is already to be found in documents such as the third-century *Appendix Probi*. The pattern is attested in all the Romance languages but is less widespread in the east than the west — cf. for example, in comparison to the foregoing, Rum. *bunătate*, Sard. *bonidade*, or It. *dodici* beside Fr. *douze*, Port. *doze* < DŬŎDĔCĬM 'twelve'.

(b) *Glide formation*: this occurred when a Latin unstressed /i/, /e/, or /u/ was in hiatus with another vowel: e.g. FĪLĬUM 'son', CĀSĔUM 'cheese', BĂTTUO 'I beat'.

These changes are less important in themselves than because they trigger many other changes such as palatalisation, gemination and cluster reduction due to the new sequences that are thereby created. Other changes are significant because of the way they interact with marking of morphological categories, as in the general late Latin loss of the opposition between /u/ and /o/ in final syllables or the differential loss of final vowels in French (all except /a/), Spanish (all except /o/ and /a/ and some /e/), or the frequent tendency in southern Italy to reduce the final syllable inventory to /i u a/ (and even to just [ə]).

It is a commonplace of historical phonology that sounds change in different ways in different environments. We have seen that the principal determining environments in the case of vowels are the location of stress

and the nature of the syllable structure. Occasionally, too, the vocalic context may be relevant, as in the case of metaphony. Less commonly, vowels may also change under the influence of adjacent consonants. Thus, rounding due to the following or preceding labial consonant has occasioned the changes in the development of It. *dovere* < DĒBĒRE, *nuvolo* < NŪBĬLUM, *-evole* < -ĒBĬLEM, etc. Even here, mention must be made of stress, since only atonic vowels are affected — contrast *deve* < DĒBET. When we come to consonant changes, we find three principal types of influence: position in the word, position in the syllable, and the nature of the adjacent segment (vocalic or consonantal). To these we may add two further, minor and usually sporadic, effects: metathesis (It. *formaggio* vs Fr. *fromage* < FŌRMĂTĬCUM) and dissimilation (IR *irimă* < ĂNĪMAM, Sp. *cárcel* < CARCĔREM).

The influence of position in the word is perhaps most dramatically obvious in the almost total loss of word-final consonants. Already in Latin the inventory of consonants in this position was severely reduced from that set out in Table 2.1. No words end in /p g k^w g^w f h/. We can easily add /b/ to this list since the only possible instances would be ĂB, SUB and OB, which could only occur proclitically. Likewise /d/ is final in the prepositions ĂPŬD, ĂD and the conjunction SĔD, to which the same reasoning applies, and in a few pronominal forms (ĬLLŬD, QUĪD, etc.), all of which are neuter and thus destined to disappear anyway. Note too that inflectional -D in the Old Latin ablative (-ĀD, -ŌD, etc.) had already been lost. Final /k/ has a slightly wider range of occurrence, being found in the demonstrative HĪC (whose neuter HŎC gives *oc* as in *langue d'oc* (p. 13)) and in adverbial forms such as ĬLLĪNC, ĬSTŬC, etc. (in all these instances the final consonant is the residue of an Old Latin deictic particle -CE) and in a few imperative forms of verbs (FĂC, DĪC, etc.) and in the exceptional noun LĂC 'milk'. We have the explicit testimony of contemporary grammarians that /m/ in final position was at most realised as nasality on the preceding vowel and often completely absent, so that it is really only the dental series /t s n l r/ that show up regularly at the end of words. Of these /t/ is lost everywhere except in Sardinian (p. 326), although the final consonant, now protected by a paragogic vowel, in forms such as Cal. *vividi* < BĪBIT may be a relic and not a secondary development. The resonants /n l r/ likewise lose their word-final position by virtue of a paragogic /-e/: hence, CŎR > *CŎRE > It. *cuore*, etc.; MĔL > *MĔLE > It. *miele*, etc. with the classic diphthongal outcome in an open syllable. Compare the French reflexes of these etyma, *coeur*, *miel*, whose vocalism indicates a formerly open syllable, even though, in line with normal French developments, the paragogic vowel has now been lost, and the syllable thereby closed again (p. 213). This leaves only /-s/, whose survival marks one of the classic diagnostics for the division of the Romance languages, not simply on phonological grounds but because of the morphological consequences (cf. pp. 59, 63). Thus, to

take by way of example two representative languages, compare the following forms of the verb 'to sing':

	Spanish	*Italian*
2 p. sg. pres. ind.	cantas	canti
1 p. pl. pres. ind.	cantamos	cantiamo
2 p. pl. pres. ind.	cantáis	cantate
2 p. sg. past subj.	cantases	cantassi

Contrast, too, the pattern of plural formation in the two languages:

	Sg.	*Pl.*		*Sg.*	*Pl.*	
Spanish	casa	casas	Italian	casa	case	'house'
	libro	libros		libro	libri	'book'
	monte	montes		monte	monti	'mountain'

The survival of /-s/, again in paragogic contexts, in those same southern Italian dialects that are claimed to preserve /-t/ is probably an independent phenomenon: e.g. Luc. *kàntasi* < CANTAS, Cal. *yèrasi* < ERAS.

There are also a few word-final clusters in Latin, namely /-ns, -rs, -ks, -ps/ (the last two may be preceded by a sonorant). There are also /-nk/ only in adverbs, /-nt/ only as a verbal inflection, and /-ls, -ms, -rt, -st/ attested in scarcely more than one word each. None of these survives in Romance.

Word-final is generally taken as a weak position in the structure of a word, and one where changes may be expected in any language. It is, of course, equivalent to final position in the last syllable of the word, and it is perhaps no surprise, therefore, to find another series of changes taking place in the final position of internal syllables. We may distinguish three types of pattern:

(a) assimilation to following dentals, characteristic of central and southern Italy and Sardinia: e.g. It. *otto* 'eight' < ŎCTŌ, Sard. *mannu* 'big' < MĂGNUM;

(b) palatalisation of velars, typical of Western Romance in the case of syllable-final stops as witness the reflexes of ŎCTŌ like Fr. *huit*, Sp. *ocho* 'eight', etc., but also of Italian in the case of nasals, e.g. *le*[ɲɲ]*o* 'wood' < LĪGNUM;

(c) labialisation of velars before dentals found in southern Italian dialects and normal in Balkan Romance, whence Rum. *opt* 'eight', *lemn* 'wood' (cf. p. 397).

The series of consonant changes traditionally known as lenition relate to the segmental rather than the syllabic environment. They occur when original stops are in intervocalic position, or followed by a liquid (CĂPRA >

Sp. *cabra*, Fr. *chèvre*), and involve a set of shifts generally characterisable as:

	Voiceless stop		*Voiced stop*		*Voiced fricative*		*zero*
	p t k		b d g		β ð ɣ		∅
VĪTA 'life'	It. *vita*	>	Mil. *vida*	>	Sp. *vi*[ð]*a*	>	Fr. *vie*

As can be seen from the above example, the Romance languages can be set out according to the extent to which they undergo the change. Generally speaking Italian and Rumanian do not exhibit the change at all (although see p. 287 for a qualification on this point), while the Western Romance languages, including northern Italian dialects, all do so, but to differing degrees. The extent of the change also differs according to the place of articulation of the consonant, with labials being the most resistant to total loss and the velars least so: SĂPĔRE 'to know' > It. *sapere*., Sp.,Port.,Cat. *saber*, Fr. *savoir*, ĂMĪCAM 'friend' > It. *amica*, Sp., Port., Cat. *amiga*, Fr. *amie*. Note that voiced stops usually follow the same paths as their voiceless congeners, so that, say, Sp. [ð] from /-t-/ and from /-d-/ are not synchronically distinguishable, and Fr. *devoir* < DĔBĒRE 'to owe' shares the same final syllable as *savoir* < SĂPĔRE. Note, too, that those languages which exhibit lenition also show reduction of geminates to single consonants: CŬPPAM > Fr. *coupe*, Sp. *copa* but It. *coppa*; CATTUM > Fr. *chat*, Sp. *gato*, It. *gatto*; SĬCCUM > Fr. *sec*, Sp. *seco*, It. *secco*. This strongly suggests a chain-shift effect connecting the two developments.

The above changes serve in part to alter the distributional possibilities of existing phonemes and allophones and in part to introduce new elements — e.g. voiced fricatives — into the inventory of Romance sound types. However, by far the most radical effects on segmental inventories are those wrought by palatalisation. The segments affected, the environments and the eventual outcomes are subject to considerable regional variation (for details, see the individual chapters), but the pan-Romance nature of the phenomenon indicates that its beginnings must be sought in the pronunciation of Latin, and specifically in the tendency of certain sound types to change their place and manner of articulation when adjacent to front vowels and glides. Table 2.3 seeks to summarise the relevant parameters. Some comments on and clarifications of this schematic representation are necessary. First, we must specify the geographical bounds of the double reflexes of /ki ke/, namely /ts s/ in Portuguese, Spanish, Catalan, Occitan French and some northern Italian dialects versus /ʧ ʃ/ in Rhaeto-Romance and contiguous Italian dialects, standard Italian and Rumanian. The fricatives /s ʃ/, and also their voiced congeners /z ʒ/, simply represent a further stage of weakening, or more precisely de-affrication, of /ts ʧ dz ʤ/, and are not strictly part of the palatalisation process itself. In the voiced series, the distribution of /dz z/ is much more limited than /ts s/, being