

**LATE LATIN
AND
EARLY ROMANCE**
in Spain and Carolingian France

ROGER WRIGHT

X FRANCIS CAIRNS

ARCA Classical and Medieval Texts, Papers and Monographs 8
(ISSN 0309-5541)
General Editors: Francis Cairns · Robin Seager

PUBLICATION OF THIS VOLUME WAS ASSISTED BY A
GENEROUS GRANT FROM THE BRITISH ACADEMY

Published by
Francis Cairns
The University, P.O. Box 147, Liverpool L69 3BX
Great Britain

First published 1982

Copyright © Roger Wright, 1982

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

British Library Cataloguing in Publication Data

Wright, Roger

Late Latin and Early Romance in Spain and

Carolingian France. — (ARCA classical and medieval

texts, papers and monographs, ISSN 0309-5341: 8)

1. French language — to 1500

2. Spanish language — to 1500

3. Latin language, Medieval and modern

I. Title II. Series

447.01 PC2814

ISBN 0-905205-1-2-X

CONTENTS

Preface

Introduction

1. LATE LATIN, EARLY ROMANCE AND HISTORICAL LINGUISTICS

The Traditional View	1
Learned Vocabulary	1
The Identification of the "Learned" Word	4
Geographical Regularity	5
Statistical Regularity	6
Types of Change	8
Lexical Diffusion	14
Strengthening and Phonemes	16
Homonymy	18
a) Clash	20
b) The Formation of Doubtless	20
Subsequently Obscured Regularities	23
Postdating Latinisms	30
Morphological Archaism	33
Conclusion	38

2. PRONUNCIATION IN PRE-CAROLINGIAN ROMANCE COMMUNITIES:

THE TEXTUAL EVIDENCE	45
Unity and Evolution	46
Non-differentiation	47
Proto-Romance Reconstruction	48
Phonetic Script and Education	50
"Vulgar" Latin	52
The Evidence of Grammarians before 500	54
a) The Second Century	55
b) The Fifth Century	56

Contents	
5· LATIN AND ROMANCE IN SPAIN, 1050-1250	208
The Council of Burgos (1080)	208
Toledo	213
Santiago	220
Twelfth-Century Castille	229
Script	233
Vernacular Writing	237
a) Navarre	237
b) Castille	238
Literature	245
The Spread of the Reforms in the 1220s	254
CONCLUSION	261

Appendix	263
Bibliography	271
Indexes	293

Contents	
The Evidence of Legal Documents	61
The Evidence of "Rhythmic" Poetry	66
The Evidence of the Visigothic Liturgy	73
The Evidence of Grammarians, 500-770	78
Cassiodorus	78
Isidore of Seville	82
a) <i>The Differentiae</i>	83
b) <i>Origines I</i>	83
c) <i>De Lectoribus</i>	87
d) Isidore's comments on linguistic variety	89
e) <i>Vulgus and Rustici</i>	93
Julian of Toledo	96
Bede and the Insular Tradition	98

3· CAROLINGIAN FRANCE: THE INVENTION OF MEDIEVAL LATIN	104
Standard Delivery of the Liturgy	104
Standard Education	112
Sermons	118
The Invention of "Romance" Writing	122
a) Oaths	122
b) The Sequence	127
The Survival of Latin Speech and Romance Writing past 1000 A.D.	135

4· SPAIN (711-1050)	145
Catalonia	146
Moslem Spain	151
Toledo	162
Asturias	163
"Leonese Vulgar Latin"	165
The Cheeses	173
Verse	175
a) Rhythmic	175
b) Popular Verse?	177
Historiography	187
The Rioja	188
Glosses	195
Conclusion	207

PREFACE

A very large number of colleagues have helped me in the preparation of this book. In particular I wish to thank Glyn Burgess, Klaus-Dietrich Fischer, Jonathan Foster, Margaret Gibson, the late Harold Hall, Richard Hitchcock, Derek Lomax, Nigel Vincent, Max Wheeler, the patient staff of the Inter-Library Loan Office of Liverpool University Library, and the many members of tolerant audiences that have commented on one (or more) of the nineteen conference papers delivered since 1974 to test the ideas elaborated in this book. I am also very grateful to Sandie Murphy for typing the manuscript.

Liverpool, February 1982

INTRODUCTION

This book examines the implications of a single hypothesis: that "Latin", as we have known it for the last thousand years, is an invention of the Carolingian Renaissance.

At first sight, this idea may seem eccentric. Latin is, after all, by definition, the speech used in the Roman Empire, and written Latin, at least, seems to have been essentially the same ever since. The normal modern view of the medieval period has been that, broadly speaking, the educated spoke Latin and everybody else spoke whatever vernacular was normal in their community; that Latin remained the same, and vernacular existed alongside it.

When applied to communities in which the vernacular was some variety of Old Romance, this established theory does not appeal to the historical linguist. Old Romance is a chronological development from the Latin of the Roman Empire, and it is not *a priori* likely that any variety of language can stand still for a millennium in one part of the community while the rest of that community gradually evolves the vernacular in the normal manner. All languages change; that seems to be an empirical fact. What the historical linguist prefers to envisage in Romance communities is merely that over the years the spoken language developed in different ways in different places, without any pockets of total resistance, until eventually the different areas of the Old Romance world evolved their varieties of speech so much that orthographical reforms had to be undertaken.

The problem with this view is that it apparently conflicts with documentary evidence. Written Latin from every century survives on manuscript. In and after the twelfth-century Renaissance, there can be no doubt that Latin was taught, learnt, read, written, and at times spoken, as a conceptually distinct language from medieval Catalan, Portuguese, Sicilian, etc. At that point there are two languages in any one literate Romance community, Latin and the local vernacular. A thousand years earlier, within the Roman Empire, there had been one

might be spelt VIRGEN, VIRGINEM, VIRGINE or VIRGINI, in the same way as Modern French [ʒiːt] can be spelt *chanté*, *chantes* or *chantent*. The old written standard had to be followed by the few who wished to write, with the result that the written texts of the Early Middle Ages look *prima facie* as if their authors spoke more or less the same language as Plautus or Pliny. The old vocabulary, grammar and spelling could be picked up from books. Old pronunciation, however, cannot be picked up directly from books.

Pronunciation is thus the key issue in this case. Only when a consistently archaizing "Latin" pronunciation norm starts being consciously used in a Romance community is the ground prepared for a complete conceptual distinction between two separate languages, rather than the preexisting collection of varying styles that is normally found within one language.

If we look to see whether the hypothesis of the introduction of a new, but archaic, pronunciation norm during the Carolingian reforms is plausible, we discover that it is. The court scholars had the task of raising the level of Church education, including the introduction of a standardized manner of performing the Roman liturgy. To this end was prepared an authorized text of the Bible; Alcuin's *De Orthographia* (dealing with both spelling and pronunciation) was compiled at about the same time, 796-800 A.D. It seems probable that a standardized method of reading texts aloud was decreed, which involved the oral production of a specified sound for every written symbol in the text. This new pronunciation caused VIRGINEM to be read (in theory) as [virginem] or (in practice) as [virdginem]; this system of giving every letter a sound has been that used in reading Latin ever since. All scholars accept that this has been the case since c.800; the novelty of the view presented here lies in the suggestion that this had not happened before in any Romance-speaking community; that the result of the reform was suddenly to create a distinction between two coexisting pronunciation norms. Were such a reform applied in Modern English, *Wright* would be read as [wraɪt] (or [wɔɪt]) rather than [raɪt], etc.

At first the reform applied only to reading aloud. One of its consequences was the invention of a system of writing Old French elaborated for the purpose of specifying intelligible vernacular reproduction of texts such as oaths, sequences or sermons.

The implications of this theory are considerable. It requires reexamination of topics central to the study of Medieval Latin language and literature, of all the Medieval Romance languages and literatures, of

language (Latin). This one language had the sociolinguistic and geographical variation normal in any widely-spread speech community. So when did the distinction begin to be felt between "Latin" and "Romance"? The usual answer to that question suggests that the separation into coexisting, distinct norms had already happened by the end of the Western Roman Empire in the fifth century. This assumption is often vague rather than explicit. Philological evidence shows incontrovertibly what would be expected anyway, that the "Proto"-Romance speech of the end of the Empire had evolved considerably from the speech of the first century A.D.; the simultaneous existence in the fifth century of an old-fashioned Latin is not so incontrovertible but is usually taken for granted.

Such coexistence of educated, archaizing speech with normal regional vernaculars as is usually postulated for Early Medieval Romance communities is attested elsewhere. The varieties of Modern Greek and of Arabic are often cited as parallels to the Latin-Romance distinction characteristic of at least the later Middle Ages. Yet the preservation of Classical Arabic and of Katharevousa depends on a conscious, detailed, continuous and consistent process of education and training such as is hard to postulate for the whole of Romance Europe in the Early Middle Ages. If this is a sound analogy, we shall have to seek the start of the Latin-Romance distinction at some point when a conscious decision to change the nature of linguistic education in this way can plausibly be postulated in a historical context. The immediate candidate for such a turning-point is the period of reform around the year 800 A.D. in the realms of the Emperor Charlemagne; the suggestion offered in this book is that an integral part of the educational reforms of that time included the introduction of "Medieval Latin", originally as a method of reading texts aloud distinct from ordinary vernacular Romance.

This theory implies that before the advent of these reforms into a community there was no contemporary distinction between "Latin" and "Romance". Everyone used styles of their local vernacular, however well or badly educated they were. It is normal in any area and in any language for educated people such as lawyers, priests, poets and linguists to have a larger vocabulary and use complicated syntax more often than their neighbours; these are ordinary features of style variation and sociolinguistic patterning that occur in all communities and need cause no surprise here. Learning to write involved the techniques for reproducing on parchment outdated inflexional morphology (such as *-m*, *-us*, *-abit*) and a spelling system not closely suited to the evolved Romance. Seventh-century French [vjɛdʒə] or [vɛdʒə], for example,

Romance Philology, of historical linguistics, and of medieval European history. This book attempts to cover all these aspects insofar as they relate to the central theme, but no individual section claims to be a complete study of its subject. The focus is maintained on the relationship between Latin and Romance. For the sake of space, not all the Romance area will be investigated in detail in this volume. Spain will take the centre of the stage for the latter part of this book; it is there that the theory requires the greatest reevaluation of accepted views. The Roman liturgy only came to Spain (outside Catalonia) in 1080 A.D., some three centuries after its promulgation in France, and came to be used generally only in the thirteenth century. If Medieval Latin was indeed invented for use initially only with the reformed Roman liturgy, it should not have existed in Spain before that date. Various areas of early medieval Spain (Visigothic Spain, Moslem Spain, León, the Rioja) produced extensive writing in what is at first glance "Latin"; it is proposed here that the texts are compatible with the hypothesis that their originators spoke Old Spanish and that the texts are the written modes of that Old Spanish. It turns out that established facts (as opposed to accepted wisdom) concerning Spanish culture from c.600 to c.1250 are not incompatible with the perspective proposed here, and in many cases are explained, more satisfactorily than they have been hitherto, by the theory that Medieval Latin was in Spain a revolutionary innovation after 1080 rather than a miraculous survival from the very distant past.

1 LATE LATIN, EARLY ROMANCE AND HISTORICAL LINGUISTICS

The Traditional View

A widely accepted theory concerning Late Latin and Early Romance is that both coexisted as spoken languages during the centuries between the end of the Roman Empire and the twelfth-century Renaissance¹. The evidence for this is supposed to be both textual and philological. The textual evidence, of the surviving written works of this period, will be considered in the next chapter. The philological and linguistic evidence will be considered in this chapter.

This "two-norm" theory is phrased in different ways. The common factor is that scholars consider, broadly speaking, that throughout these years there is a significant distinction that can be made between the speech of the educated, who spoke some kind of "Latin", and the speech of the uneducated, who spoke their local evolved vernacular. The following representative quotations, here presented in chronological order, show how entrenched this belief has become.

Rice (1909: 7), discussing postulated loanwords from educated Latin into the vernacular, stated simply that:

The forms of early Old French loanwords prove beyond all doubt that a conservative form of language, based on the spelling of the time and on oral tradition, must have existed in both the seventh and eighth centuries.

Menéndez Pidal (1926: para.109) declared flatly that:

Latin was in the early centuries ... the language of elevated communication between all educated people; it was thus the only, or at least the main, norm of good speech. Romance was merely colloquial; only the uneducated had Romance as their sole means of expression.
(My translation)

1. This chapter draws on previous published work of mine; in particular Wright 1976a, 1976b, and 1980.

2 Late Latin, Early Romance and Historical Linguistics

Alfred Ewert (1933: para.500):

... the clerks carried into their pronunciation of Latin many of the features of the vulgar tongue ...

Mildred Pope's influential handbook similarly assumed that Latin and Romance were distinct when she stated that (1934: 229):

In Gallo-Roman, particularly in Early Gallo-Roman, the clerks carried over into their pronunciation of Latin the tendencies that were most strongly affecting the speech of the time, and the words borrowed at this epoch ordinarily reflect these developments.

E. Löfstedt (1959: 3):

There could be no sharper contrast than that between the conservative character of officialdom's language and the tendencies of popular speech.

Norberg has studied Medieval Latin with greater sympathy than most other scholars, but he too has been sure that pre-Carolingian Latin and vernacular are conceptually distinguishable (1968: 29):

Merovingian Latin has above all been profoundly influenced by the spoken language.

Lüdtke had the idea of rephrasing this supposed bilingualism in pre-Carolingian times into the sociolinguistic concept of diglossia (1968: para.5.1.3; for "diglossia" see Ferguson 1959):

... a gap is formed between the vernacular language, naturally learnt, continually evolving, and the language of culture ... the gap between the two languages must have been getting deeper and wider and the diglossia ever more obvious.

The most influential of all the North American Romance philologists, R.A. Hall, described the relationship between Latin and Romance in the following way (1974: 106) – suitable for the Late Middle Ages, but here projected back to the immediate post-imperial centuries:

LATIN AS A SECOND LANGUAGE. The antinomy between Romance and Latin did not imply, as has often been thought, that the latter ceased to be spoken, and survived only as a "dead" language used exclusively for writing and reading ... Latin was the regular language of instruction, discussion, and debate, in elementary and advanced learning, in the church, in law-courts, and in the offices of lawyers and notaries. It also served as a lingua franca for learned persons, not only in their written correspondences, but in their personal contacts with each other at home or on their travels.

Bustos Tovar was even prepared to envisage a three-fold distinction in seventh-century Spain (1974: 67):

While in very educated circles they would have continued to speak a Latin learnt at school ... similar to that written by St. Julian or

St. Isidore, on another level, without such specialized study, they would have spoken a Romance-affected Latin similar to what later became the popular Leonese Latin. On a lower level they would have used Old Romance.

Macpherson had no qualms about the distinction (1975: 94-5):

Until well into the medieval period Latin continued to be the common language of cleric, lawyer, administrator, natural philosopher and man of letters; Church Latin and written Latin exerted a continuous influence on the spoken language ... The popular phonetic development of a word may be arrested or modified at some stage in its development by the pronunciation of the more educated classes.

Bynon was prepared to assert (1977: 249-50):

There can be no doubt that both the dignitaries of the Church and educated people in general used Latin as the normal means of communication throughout the area from a very early period.

The idea continues to flourish in the 1980s. Lapesa's widely-used and now much-revised *History of the Spanish Language* still declares that ecclesiastical pronunciation was different from the ordinary vernacular (1980: 110):

Some words have existed uninterruptedly in speech, free from the memory of their literary form and abandoned to the course of phonetic evolution ... others have not undergone a phonetic development unencumbered with learned reminiscences ... VIRGINE or ANGELUS ... in preaching and in religious ceremonies were pronounced in a way that may not have been exactly pure Latin but was essentially respectful of the Latin; people's ears became accustomed to the ecclesiastical pronunciation, whose influence prevented the usual phonetic tendencies from running their course: VIRGINE became *virgen*, not *verzen*, and ANGELUS *ángel*, not *añlo* or *amlo*.

There is, then, a consensus concerning the plausibility of the distinction between some kind of spoken "Latin" and the vernacular. In future chapters, however, the evidence in support of this dichotomy will be examined and found not to be substantial; it needs to be made clear here in advance that the view being argued against is indeed very widely held and believed, rather than being a mere straw man. Almost everyone who has studied the linguistic state of Romance communities in the centuries from 400 to 800 seems to accept that Latin and Romance speech both existed at the time, and some believe that they coexisted in earlier times as well (e.g. Leonard 1980, Pulgram 1975; cp. Janson 1981, reviewing Pulgram).

The argument in favour of this view seems to be (i) an instinctive feeling that "Latin" must have survived alongside Romance in the Early

Middle Ages, if only because there it is, incontrovertibly, in the Later Middle Ages, allied to (ii) the discovery, well-argued and by now indisputable, that many of the linguistic developments from Imperial Latin to Romance began long before the end of the Roman Empire. A few scholars have briefly rebelled against this idea; e.g. Janson (1979: 13), "Proto-Romance is Latin and Latin is Proto-Romance", but the implications of this identification have yet to be explored.

Learned Vocabulary

Around the turn of this century, philological attention was turning to words borrowed from Latin into French; Berger's study (1899) began a trend taken up by e.g. Paris (1909). Words taken into Old French from Latin in the Middle Ages are often identifiable because they have not evolved phonetically in the way that a comparable "popular" form would have done; a standard example is Fr. *fragile*, borrowed from Latin, with the *g* representing a consonant, as opposed to O.Fr. *freile* (Mod. *frêle*), a popularly evolved form with the *-gr-* group regularly simplified to *-r-* without a consonant (cp. MACGISTER > *maître*). The earliest French texts written in an orthography that attempts to correspond to the evolved phonetic forms are set down in the ninth century, at which point the coexistence in France of two ways of writing and pronouncing a word can be regarded as established knowledge. This distinction between Latin and Romance was at that time assumed not to be something new in the ninth century but to have been in existence for many centuries already. Yet we know anyway that the pronunciation of Latin was reformed in c.800 A.D.; the systematic learned-uneducated distinction, in phonetics as in orthography, can only be regarded as surely attested with the Strasbourg Oaths (842). The discovery of "learned" vocabulary was a real discovery but the conclusion that these unevolved "Latinate" forms had existed in such a form in anyone's speech before the ninth century was logically unwarranted. When Le Coultre made an acute analysis of Alcuin's *De Orthographia* in 1805, he claimed that Alcuin was describing Latin as traditionally spoken, and never considered the idea that this official "Latin" pronunciation might have been something new that had not existed in France previously.

Despite the uncertain basis of the argument, the supposedly proven existence of unevolved archaizing learned pronunciation, in the years between the Roman Empire and the emergence of written Romance in new spelling systems, came to be treated in Romance philology as established fact. This assumption had the consequence that

any word which seemed not to have evolved enough, according to the discoverable regularities of sound change, became a candidate for being described, on this evidence alone, as a "learned" word. Whereupon the existence of such an educated unevolved layer of phonetics in pre-Carolingian Romance came to be seen as apparent confirmation of the two-norm theory. The argument is thus circular: the theory of learned vocabulary depends on the Latinate layer, and the theory of a Latinate layer largely depends on the learned forms.

Yet the historical and textual evidence (to be examined in subsequent chapters) is compatible with a simpler hypothesis: that this Latinate pronunciation norm, which gave each written letter a separate sound, was established in a Romance community for the first time by the Carolingian Reforms of c.800 A.D. It is worth looking to see if the philological data that are supposed to support the two-norm theory can be explained in other ways with the procedures of historical linguistics. If so, the existence of borrowings from Latin can be accepted as plausible only after the advent of the new "Latinate" pronunciation into any community. The traditional theory can be thus supported for the years after 800 in France and 1080 in Spain, but words containing "Latinate" phonetic features cannot be regarded as evidence of the "Latin" social dialect in preceding centuries. The circularity of the traditional argument can thus be broken, and the hypothesis advanced that the only speech styles that existed in any Romance community before these reforms belonged to the evolving Old Romance vernacular of that community, and nothing else.

The Identification of the "Learned" Word

The words traditionally identified as "learned" are those that appear to be exceptions to the general regularities of phonetic change; they have not evolved as much as phonetically analogous words usually did, and the view is generally held that this is specifically due to the predominant use of these words by the educated classes in society, who used "Latin" rather than Romance. For example, the original consonant cluster *-ns-* usually loses the *[n]*: *MENSAM* > *Sp. mesa*; *MENSEM* > *Sp. mes*. The *[n]*, however, is still there in the descendants of *PENSO* in its sense of 'to think': *Sp. pienso*, *Ital. penso*, *Fr. pense*, etc. Similarly, philological orthodoxy states that words beginning originally with *FL-* palatalize in Spanish into *[ʎ]*: *FLAMMAM* > *llama*. The *[ʎ]* is still there, however, in *flor*, from *FLOREM*. Spanish *pienso* and *flor*, therefore, regularly feature on the lists of learned vocabulary such as that in

Bustos Tovar (1974: 482-83, 607-08), as words whose supposed phonetic retardation is due to their use by the educated.

The identification of those words that are supposed to have been used predominantly by the Latinate thus depends on the prior establishment of which developments are to be regarded as "regular". The regularities of sound change are sometimes called "laws", but they are not generally applicable in the same sense that the rules of grammar are. We can tell, for example, that in English there is a phonetic rule that adds [s] to the singular form of a noun to form a plural, or to a verb to form the third person singular present, when that noun or verb form ends in an unvoiced consonant (other than a sibilant). Neologisms such as the transatlantic verb *deescalate* (*Nixon deescalates U.S. involvement*) or the noun *raincheck* (*they took rainchecks*) form the third person singular present and plural form respectively in this manner. This suggests that native speakers have some kind of a "rule" in their heads that [-s] is added to words in such circumstances whether or not they have heard or used the words before. In such cases, where it is plausible to suggest that there are rules which speakers follow, it is reasonable to claim that any word which does not conform to this pattern is some kind of "exception". *Children*, for example, is a survivor from an age in which the rule of English was different, and it is thus legitimately characterized as an "exception" in the modern language. Sound changes, however, cannot be postulated as existing within the speakers' heads. The simultaneous availability of two variants might be known to a speaker, but the evidence does not support the idea that "Latin [-ns-] > Spanish [-s-]" had the same kind of psychological reality as "[prt] + plural = [brts]". Sound laws are statistical truths discovered afterwards by scholars; at the time of their operation they appear as cases of variation or as competition of forms, and unless the native speakers are etymologists they do not base their own practice on any knowledge of which form is the older and which is the newer.

Geographical Regularity

It is worth considering, then, why any particular sound change comes to be thought of as "normal" or "regular" at all. There are many factors in play, but the basic ones are geographical and statistical.

It is regarded as normal and regular for CA- at the start of original Latin words to develop to *ch(a)-* in France ([ʃa]): for example, CANTANT > *chantent*, CATTUM > *chat*; or, in a free syllable, to *che*: CARUM > *cher*. It is also regarded as normal and regular for CA- at the

start of these words not to develop at all in Castille; CANTANT > Sp. *cantan*, CARUM > *caro*. The presence of neither of these developments has caused the other to be considered by philologists to be exceptional or irregular, because the regularities concerned can be neatly allocated to different sides of linguistic boundaries, and hence the divergence is considered of no significance, as if it were believed that the political divisions of nineteenth-century Europe had been instinctively followed by the sixth-century inhabitants of Romania. In this manner, two different resolutions of the same original etymon have been considered to be both normal or regular. On the other hand, scholars have not usually thought such variants to be similarly regular if both occur in the same area. The inevitability of the sound law, remorselessly altering every word whose phonetic configuration brought it into the scope of that law, as the nineteenth-century grammarians are popularly believed to have envisaged it, only applied within geographical units. The same development could be regular in one valley but irregular in another. From the point of view of the early scholars of the subject, most of whom were specialists in the Romance of Northern France, this geographical proviso was an essential assumption. Without it, the logical application of the theory that the educated were supposed to use archaic forms could have been most unfortunate: Sardinia, in general, is the area whose speech seems to have evolved least from Latin, and Northern France the area whose speech, in general, seems to have evolved most, but they would hardly have wanted to conclude that early medieval Sardinia was the most educated and literate part of Romania.

This avoidance of the logical conclusion of accepting the two-norm theory was half-noticed, and sometimes rationalized away with another bold assumption, that increased intellectual activity leads to language change. Entwistle (1962: 72) expressed this in the following characteristically neat manner:

It is the languages which have the most active intellectual background that show the most rapid changes.

There is little serious evidence to support this view either. Even in Early Romance it is untenable, since although Merovingian France seems to have been more active intellectually than contemporary Sardinia, it was considerably less active intellectually than Visigothic Spain, whose language also, in general, seems to have evolved less than Northern French. But even if it were tenable as an *a priori* linguistic principle it would conflict logically with the two-norm theory; the educated élite, supposedly responsible for lack of evolution and Latinate speech habits,

and the active intellectuals, supposedly responsible for rapid evolution and vernacular speech habits, are the same people.

This confusion is heightened by yet a third belief, that remote illiterate dialects are usually more archaic in their speech than the educated norm. Some scholars seem to have quite cheerfully believed this at the same time as the two-norm theory, which is based on the contrary assumption, that the educated élite are usually more archaic in their speech than the illiterate rustics. There are, of course, archaic features preserved in some rustic dialects. There are, of course, also evolved features attested only in rustic dialects. Neither case seems to suggest to modern linguists that there is any link at all either way between comparative literacy and comparative archaism of speech in different geographical areas.

Scholars have avoided the problems caused by these mutually incompatible assumptions partly by vagueness, partly by ignoring them, but also by separating different geographical areas into different study areas. There is, therefore, no problem about having the evolution of CANTO > *chante* described as "regular" in one place and that of CANTO > *canto* described as "regular" in another; different places are said to have different languages, even in the sixth century. What has been definitely regarded as a problem, however, given this geographical slant on the question of regularity, is the coexistence in the same place of apparently different developments from the same kind of etymon. *Chante* is no problem to the philologist who is only considering CANTO > *canto* in Castile; but the different Castilian development of a different word in the same place, such as CATTUM > *gato*, is considered problematic, "sporadic", "irregular".

Statistical Regularity

Once the decision had been taken (subconsciously, in the main) that it was sensible to parcel out the field of Romance Philology into geographical units of study, the decision concerning which developments are to be regarded as "regular" was normally made on a statistical basis. Concerning the development in Spain of initial CA-, for example, 150 medieval entries in Corominas' brief etymological dictionary (1961) show no change – e.g. CAPRAM > *cadra*, CAPUT > *cadu*, CAPITALEM > *candai*, CASTELLUM > *castillo*, Greek *kará* > *cadá*, etc. – whereas 20 entries show the development of initial [ka] to [ga]: *gato* < CATTUM, *galápagos* < *CALAPACUM, *garnella* < CAMELLAM (cp. B. Löfstedt 1959: 52-3), *garnaza* < CAMOCUM, *gafoire* formed off *cafo*, *garrañhar* < *CARPINARE, *garraucha* < *carraucha*, *garúa* < *CALLIGINEM,

garuña < *CARULLAM, *garza* < Celtic *KARKIA, *garvilia* formed off CAVUM, *garniño* formed off CADMIA, *garzacho* perhaps formed off CASPA, *gazaza* formed off *cazar*, *gannara* formed off CAMLUM. (There are also several Arabisms: *gafo* < *qaf* a, *gasa* < *qazz*, *garbón* < *gabá*, *garraña* < *garāb*, *garnate* < **gannāt*. Arabisms do not always have this effect: e.g. *alcázar*, *alcáide*, *cafi*, *cafozo*, *candil*.) The latter development, of [ka] > [ga], is thus quite widespread.

The voicing of [ka] to [ga] is statistically more common if the [ka] is preceded by a vowel: e.g. VACARE > *vager*, PACANT > *pagan*, SECAT > *sega*, etc. These statistics are awkward to assess in view of the number of distinct morphological forms that might or might not be included, but it seems reasonable to suggest that taking all positions into account [ga] < [ka] might be as common as [ka] unchanged; however, the regularity is usually assessed (for good reasons) on the basis of word-position and phonetic environment, and there is no doubt that statistically in standard Castilian Spanish it is "regular" for word-initial [ka] not to change and for post-vocalic [ka] to voice to [ga]. In this way, for purely statistical reasons, *carran* and *sega* are both said to develop "regularly", even though they develop differently, whereas *gato*, etc. and e.g. *foca* < PHOCA, are "exceptions" to be explained in some further way, despite the same results being "regular" in *sega* and *carran* respectively. (*Foca* is a fifteenth-century borrowing from educated Latin.)

CANTO > *canto* is a statistically "regular" result in every respect. CA > ca is regular; NT- > -nt- is "regular", once we regard the normal development of -NTI + vowel > -ntz ([nt] > ntz > nt) as a separate case (e.g. CANTIONEM > *cantión*); -O > -o in the first person singular is "regular", once we frame a separate rule for monosyllables (SUA > *soy*, STO > *estoy*, VADO > *voy*, DO > *do*). *Canto* is thus "regular": vernacular, not ascribed to "Latinate" speakers, despite the fact that it is for practical purposes phonetically and phonetically the same as the Latin CANTO. Accordingly, mere similarity to the Latin is not the criterion for being thought a learned word whose retardation is attributable to educated use; the criterion is that of greater similarity to the Latin than to words evolved from phonetically analogous origins. *Pieno* < PENSO, for example, is said to be "culto", even though the stressed [e] has diphthongized in the vernacular manner; it is statistically more common for the -N- to disappear from -NS-, as in *mensa* < MENSAM, *mesa* < TENSAM, etc. with the result that *pieno*, although it has evolved a non-Latin diphthong, is officially a learned word, and *canto*, although it has hardly evolved at all, is allowed to have existed in the vernacular.

The case of *flor* is instructive. It is arguable whether even the statistics support the common description of *flor* as a "cultismo". There seem to be six original Latin words with initial FL- that survived in the early medieval vernacular. Of those six, four have preserved the initial cluster: *flaco* < FLACCUM, *fleco* < *fluco* < FLOCCUM, *flojo* < FLUXUM, *flor* < FLOREM². One has palatalized the cluster to [ʃ]: *llana* < FLAMMAM. The other has reduced the cluster to a simple [l]: *lacio* < FLACCIDUM. Comparing FL- with CA-, it might have been expected that FL- > *fl-* would be considered the "regular" shape of vernacular forms, with *llana* and *lacio* being considered "sporadic" and in need of further *ad hoc* rationalization. This has indeed been the treatment given to *lacio*. Malkiel perceived a problem here (1963:64: 159); aware that according to prevailing orthodoxy FLACCIDUM "ought" to have become [katsjo], with an initial palatal [ʃ], he accepted the view that the form once existed as the standard, with the subsequent [ʃ] > [l] explained as a desire to dissimilate palatals in the light of the following [j]. (*Lacio* is attested as a variant.) This attempt to explain a development of [ʃ] > [l] through a debatably general palatalization followed by depalatalization is possible, but it may not be necessary, even though it has convinced Macpherson (1975:136) and Dworkin (1978:609). F- > *ø* is in fact the statistically normal resolution of F- before a vowel (FACT > *hace*, FILAM > *hija*, etc.). *Llana* might seem to be similarly sporadic, but in fact it has been called "regular", because FL- was grouped for analytical purposes with other initial clusters of unvoiced consonant plus [l]. Initial PL- and CL- offer us sufficient examples of palatalization, such as *llorar* < FLORARE, *lleno* < PLENUM, *llave* < CLAVEM, for it to be statistically true that, for all initial unvoiced consonant + [l-] clusters, [ʃ] is the regular evolved form.

Choosing a different brief and a different field of operations thus produces a different pattern of regularity; within FL- words, *flor*, *flaco*, *fleco* and *flojo* are "regular"; within F- words, *lacio* is "regular"; within unvoiced C + [l]- words, *llana* is "regular". The choice of the grammarians has fallen on the latter; *llana* is patted on the head for

meek obedience, and the others have seemed to require special explanation. In this case and others, it might seem to have been quite hazardous which form was said to be regular, although in practice the possibility of assigning unchanged forms to "culto" speech may have made the choice of the evolved form as the "regular" one more likely. Once the assignment is made it is seen as indisputable. Corominas took it as firm, for example, and was moved to speculate on the apparent fact that *flor*, *fleco*, *flaco* and *flojo* seemed to have been confined to the educated. Under *flojo* (1954:57: II,541b) he declared that "the conservation of the FL- group is due to the moral aspect of the word, which accounts for the triumph of the pronunciation of the educated classes, as happened in *flor*, *flaco* and other similar words." This attempt at an explanation is unconvincing. Were it not for the FL- phenomenon, would we consider "flower" (*flor*), "thin" (*flaco*), "fringe" (*fleco*), "feeble" (*flojo*) to be essentially more moral concepts than "flames" (*llamas*)? The flames of hell suggest not. Even if they are, what evidence is there that uneducated groups never discuss morality? Morality is surely a feature of the conversation of all, even if not of the behaviour, and evaluative terminology is a linguistic universal in any group. In cases such as this, Corominas seems to expect his readers to visualize the monks of Early Medieval Spain as being so determined to resist the postulated general Castilian change of initial FL- > [ʃ] that they refuse to call flowers by anything other than the Latinate *flor*; while having no particular reservations about losing the final [-e] of the spoken form, nor, indeed, about changing the gender from the Latin masculine to the Castilian feminine. (At the same time, the Italian monks successfully preserved the original gender, but were happy to palatalize into *il fiore*, so morality too has geographically distinct linguistic consequences, it seems.) How, subsequently, were these monks so successful at the task of encouraging the other 99% of the population to ignore the development that is supposedly taking place unanimously in their speech, and say *flor* rather than their natural **flor*? Why do the same monks say *llana* without a qualm? Monasteries had fires, the outside world had flowers, and the scenario is not plausible.

Corominas' brave attempt to explain the absurd consequences of his predecessors' views is surely untenable. The postulate, however, that FL- words need some kind of special explanation, survives. Malkiel explained the supposed "cultismo" of *flor* as arising "since generic terms are frequently *cultismos*" (1963:64: 29-31); this idea comes from Lapasa (1980: para.2.5.3), but is hardly to be taken seriously. "Flower"

2. The case of *fleco* is confused by the existence of the rare word *lleco*. If both come from FLOCCUM, as Corominas suggests, then *fleco/lleco* will be a pair of doublets; doubters are discussed below. But the existence of Medieval Latin *fleco/flocus/frocas*, with the same meaning as *lleco* ("uncultivated land"), suggests that this word, probably of Germanic origin, could be the source. Since the etymology is unclear, and the first sure attestation of *lleco* is in Covarrubias' dictionary (1611), I have felt justified in omitting it from this section.

is no more a generic concept than is "man" or "fish"; *hombre* < HOMINEM and *pez* < PISCEM are hardly Latinate. *Fleco* is also developed to a greater extent than the vernacular *regular* would foresee: *PLUCCUM* > *flueco* is fair enough, but *flueco* > *fleco*, phonotactically explicable through the awkwardness of [fwe], is over-evolved. Were this really a word confined to the supposedly archaizing literate, **fluco* would seem more consistent. Malkiel, however, "granted in principle the existence of a significant learned transmission of these clusters" (1963-64: 157), so even when the attribution of learnedness is as statistically dubious as it is with *fl-* Malkiel felt unable to dispute it.

One voice has been raised against it. Badia Margarit (1972) has diffidently declared his belief that the [A] was a late evolution, and that such obviously non-learned words as *claro* < CLARUM were either words that had not succumbed by the time of standardization in the late thirteenth century or words borrowed from a non-Castilian dialect (since PL, CL > [A] is exclusively Castilian). Badia's suggestion has been ignored, even by the scholar in whose honour it was offered. There are many such cases in which a reconsideration of the statistical basis of the attribution of "regularity" shows that a form might as plausibly be considered "regular" as "irregular", and in which any postulated "rules" have many exceptions.

It is not the intention here to explain all such variations, but in the case of *flor* it can be simply done along traditional lines. Even if we accept PL > [A] as the "rule", we can make *flor* "regular" by adding a rider that the palatalization never operated in such a way as to produce monosyllables. As a statement of fact, it is true that in Castilian there are no monosyllables beginning with [A] and there probably never have been. Latin ILLOS, for example, palatalized to *ellos* [eʎos] in tonic positions, but reduced to *los* rather than **los* in atonic positions; this suggests there was a distaste for such a form, even perhaps a morpheme structure rule. Being monosyllabic is acceptable as a respectable conditioning factor for change; for example, the retention of final nasal consonants is confined to monosyllables (QUEM > *quien*, CUM > *con*, but NUMQUAM > *nunca*). Had Menéndez Pidal included this rider, the idea that only the educated were moral or generic enough to talk about flowers would not have needed to rise. It is not my purpose here to claim to have explained *flor* by this rider, for there seems to me to be nothing to explain, but the possibility of such an explanation is an example of the way in which, if we wish them to, irregular forms can be made "regular" by a stroke of the legislative pen. The assignment of regularity or irregularity is not always a simple process or even a meaningful one.

The supposedly anomalous forms, such as *pienso* and *flor*, have thus only been identified as such on the basis of the prior establishment of the "rules". In Spain, the rules were engraved on tablets of stone under the title of *Manual de Gramática Histórica Española* (Menéndez Pidal 1904). Any recalcitrance to those rules became a sufficient reason for the word in question to be placed among the ranks of "cultismos" or "semicultismos", learned words confined to or influenced by the "latinate" élite. This attribution could be taken to extremes. The case of *engendrār*, for example, takes the theory of "cultismo" into the realms of mild absurdity. Latin INGENERARE has undergone at least eight identifiably "regular" sound changes in its development to Modern Spanish *engendrar*, viz: the change of the initial "short" [i] to [e]; the loss of the internal atonic [e]; the subsequent insertion of a [d], which is legitimately regardable as "regular", for the [nr] cluster was unstable in Old Spanish (it sometimes survived (*honor*), sometimes reversed (GENERUM > *yerno*), but often acquired the [d], as in the future forms of *poner* and *rener*, *pondre*, *rendre* (which coexisted in O. Sp. with *ponre*, *renre*)); the loss of the final [e] in approximately the twelfth century; the change of Latin [g] to Old Spanish [dʒ] before a front vowel (as in Modern Italian *genero*); the deaffrication of the [dʒ] to [ʒ] in the late Middle Ages; the unvoicing of the [ʒ] to [ʃ] in the Renaissance period; the subsequent backing of this voiceless fricative to [x]. It might thus be thought to be not obviously an archaizing Latinate form. But the theory was held that in Old Spanish the affricate [dʒ] ought to have been reduced to a palatal [j] before a front vowel, as indeed happened in e.g. GELUM > *hielo*, GERMANUM > *germano* > *hermano*, GENERUM > *yerno*; nor do the forms with stress on the stem diphthongize (*engendra*, not **engendrā*). As a result, *engendrār* is stigmatized as being retarded and learned (cp. Bustos Tovar, 1974: 444). This may not be justified. There are no cases at all that I can see of Latin -NGE- becoming [nje] in a pretonic syllable; *hielo* and *yerno* have tonic vowels that diphthongized anyway, and the instability of the pretonic [j] in the transient [germano] — which was never the standard form, soon becoming Sp. *hermano* (and Port. *irmão*) — is another symptom of the medieval Castilian tendency I have discussed elsewhere to avoid diphthongs in preclitic position (Wright 1976c). It is hard to be certain, but there seem to have been only three cases of -NGE- > [nje] even in a Latin tonic syllable; *quientos* < QUINTENTOS, O.Sp. *puniente* < PUNCENTEM, *cientos* < CINGENTES (disregarding affixes such as in *tañendo* < TANGENDO; Menéndez Pidal 1926: para. 49). In addition, there is INGENERARE the morpheme boundary after IN-

14 Late Latin, Early Romance and Historical Linguistics

which is able to inhibit other changes (e.g. *INSIGNARE* > *enseñar*, not **esefar*), and the whole situation with -NG- plus front vowel is fluid; we have, for example, *INGERE* > O.Sp. *unzír* (O.Sp. [dz]), *GINGIVAM* > *encía* ([s]), *TANGIT* > *tañe* ([n]), and other resolutions to show that "regularity" in this case is hardly discoverable at all (see Malkiel 1975a). Such confusion could well inhibit diphthongization; criticism of the form for not undergoing the "regular" change may be misplaced.³

Engendrar is an extreme case. Many of the supposed "cullo" forms have little firm basis for such an attribution, however, and the label has merely hindered their rational examination. Even so, there are forms for which surprise at their archaism is legitimate; the status of exceptions in historical linguistic theory needs to be examined in connection with them.

Types of Change

Some sound changes are "isolative" ones that apply in all phonetic circumstances. The Proto-Romance merger of originally long and short [a] is one such. Completely isolative changes probably do not occur in Early Romance; Latin stressed or counter-tonic [u] > [y] in French, as in *tu, nul, durer*, has been said to be isolative, but the presence of stress is a conditioning factor even here. For practical purposes, Early Romance developments can be regarded as "conditioned" changes, ones that applied in some phonetic circumstances but not in others.

This, for example, explains how the loss of [n] in *mesa* < *MENSAM* can be described as "normal". In general, Latin [n] does not normally disappear, neither initially – e.g. *NATAM* > *nada* – nor intervocally – e.g. *TENERE* > *tener* – nor in a cluster – e.g. *DENTEM* > *diente*. Its loss is, however, statistically regular under the phonetic condition of having a following [s], so it is reasonable to draw up the regularities [n] > ø / - [s] (and [k] > [y] / V - V).

The relationship between isolative and conditioned change is not clear, although they may well be connected. Wang (1969), for example, wondered whether even long-established conditioned changes may be isolative changes that have not yet spread as far as they will.

3. *GENERUM* > *yeruo* ("son-in-law") may have chosen [j] and [m] at the same time as *engendrar* chose to avoid [j] and adopt [ado]; the two words being given opposing choices in the current free variations [tɛʒ-] and [j-] / [m-] and [pdr-] in order to avoid confusion of "son-in-law" and the central morpheme of *engendrar*. Cp. the variants *gente* ~ *yente*; and Malkiel (1979) on *harta* and *hacia*.

Maybe all examples of Latin [ka] will be Spanish [ga] in a millennium or two, as all positions of Latin /b/ seem to have become eventually Spanish /β/ (Penny 1976). Krishnamurti (1978) has taken the cue and cheerfully described a change that has now been going for twenty centuries and could last for as many more.

There is one general point concerning conditioned change that is of particular relevance to the words under examination in this chapter; the fact that after the conditioned change has operated, the old sound will usually survive as a phonetic entity in those contexts unaffected by the conditions. The only exception to this could be the suffering by a phoneme of more than one conditioned change simultaneously in such a way that no contexts are left unaffected; this situation is possible but most unlikely. For example, Latin [t] became Spanish [d] and then [ð] intervocally, became [ts] and then [θ] before [j], and disappeared word-finally; even so, it survives initially, and even intervocally as the simplification of some consonant clusters in which it was originally the second part (e.g. *gota* < *GUTTAM*, *escrito* < *SCRIPTUM*), and the phoneme itself is as vital as ever even if relatively fewer words contain it now than two millennia ago. This means that, in general, if any form, for any reason, does survive unaltered a change which we might otherwise expect it to undergo, the form may stick out like a sore thumb to modern philologists, but speakers of the time, even synchronic phonologists of the time, are unlikely to think of it as peculiar at all. *Pensar*, for example, is immediately obvious to the modern philologist as being a remarkable exception to the [ns] > [s] "rule"; but the [n] and the [s] are in themselves very common sounds of Early Romance, and even the cluster [ns] has moral support from forms such as *cansar* (< *CAMPARE*). If we permit ourselves to ignore the morpheme boundaries that sometimes blocked the change, and involve the same sounds, the support from analogous forms is quite large (e.g. *enseñar* < *INSIGNARE*, *consolar* < *CONSOLARI*). Accordingly, there need have been no prolonged or tense struggle for survival on the part of such a sound as the [n] in *pensar*. Whatever the reasons for its survival, the form itself would not have seemed peculiar or undesirable at all within the superseding phonological system. So far as I know, no one thought *pensar* in any way to be a strange word until the philologists decided it was last century. If there is a sensible reason for a word's not undergoing a conditioned change, not changing is a simple goal for the word to achieve; "exceptions" within historical linguistics are not usually "exceptions" within the synchronic phonological structure of the language.

If then, for some reason, speakers prefer not to develop a word in the same way as most analogous words, there is rarely a synchronic reason for it not to remain unchanged. As the next section shows, there is reason to suppose that some words are less amenable to a change than are others, for reasons that are at times discernible but do not normally have anything at all to do with the relative literacy of different strata of society.

Lexical Diffusion

Wang's 1969 article arose from the problems that can confront a word that is being subjected to two changes at the same time; if the changes are incompatible — for example, if the application of each change removes the condition for the other — it is impossible, with the best will in the world, for a word to evolve in a regular way. A Romance example is Latin *limpidum* 'clear, attractive' > Spanish *limpio* 'clean', *lindo* 'pretty'. Unless Corominas (1954-57) and Dworkin (1978: 607) were right in thinking that Spanish *lindo* descends from *LEGITIMUM*⁴, — in Portuguese *lindo* the case is stronger — we can see that both result from 'regular' changes; between two unstressed vowels, [d] regularly disappears — cp. *ribis* < *TEPIDUM*, *recio* < *RIGIDUM*; between two consonants it is regular for atonic [l] to disappear — cp. *brade* < *LDATTEM* — and if three consonants thereby cluster the middle one disappears and the others assimilate — cp. *once* < *UNDECEM*. *Limpio* is the result of the first change, which removes the second consonant, with the consequence that the conditions for the second change do not apply; *lindo* is the result of the second change, which removes the vowel, with the consequence that the conditions for the first change do not apply. The two words share the semantic content of *limpidum* in a sensible manner, and the incompatibility is resolved.

This general idea (with Chinese evidence, in the main) led Wang to consider whether some morphemes are more keen to change than

4. Corominas decided that the stressed [i] of *limpidum* 'ought' to have opened to [e], despite the indisputable provenance of *limpio* from *limpidum*, the evidence of the exact opposite in *libro* < *TEPIDUM*, and the form *lindo* from *limitem*. *LEGITIMUM* > *lindo* involves several unattested and debatable steps, and we would expect to find many Medieval variants (cp. *METIPIDUM* > *meismo*, *meimo*, *meimo*, etc.), which we find in Portuguese but not in Spanish. Dworkin remarked (wrongly), that 'most Hispanists reject the equation *limpidum* > *lindo*', phonetically this etymon is still surely the most likely, but the Spanish semantics may well have been affected by *LEGITIMUM* or its Portuguese descendants, and the form of *Pg. lindu* may have been affected by *Sp. lindo*.

others. The traditional theory of sound change assumed that the progress of a change operated the same way at the same time in every relevant word, but this was never more than an assumption; a simple look at changes in progress has shown without any doubt that often, perhaps always, some relevant morphemes undergo a change before others⁵. If this is true, as it seems to be, then the problem of accounting for 'archaic' unevolved forms that survive in speech could be redefinable as the problem of deciding why certain words should find themselves at the beginning, and others at the end, of the queue for the application of that change. If the change fails to reach every item, those at the end of the queue will be left unaltered, and as we saw in the last section there are unlikely to be dreadful phonological consequences of this.

The idea of a queue has subsequently been elevated into a theoretical school, that of 'Lexical Diffusion'. Chen attacked those who use 'dialect borrowing' as a panacea for any problem; he pointed out (1972: 462) 'the inadequacy, both methodological and factual, of a facile recourse to interdialectal borrowing as an "explanation" of irregular sound changes.' This recourse is at least feasible if the dialects concerned are indisputably known to exist; inventing hypothetical dialects from which to borrow otherwise inexplicable forms, such as inventing a spoken Latin distinct from Early Romance in Early Medieval Europe, would seem even less adequate as a way of snatching cases of ostensibly insufficient evolution. Chen and Wang then joined forces to attack the traditional preference for ignoring exceptions (1975: 256):

The neogrammarians were tightly perceptive of the significance of the over-all regularity of phonological processes, since there was no a-priori reason why sound laws should operate with such remarkable uniformity... the neogrammarians often blamed the exceptions on such culprits as dialect mixture and analogy. Cross-dialectal interferences are certainly present in most real linguistic situations; but more often than not, linguists have used dialect mixture as an excuse for not producing evidence of a substantive nature.

This theory now has an established place within historical linguistics; most scholars, including many specialists in dialectology, accept that sound change in practice often affects some words later than others. 'For dialectologists, the theory of Lexical Diffusion is credible in a way that the structuralist and generativist hypotheses are

5. In 1980 BBC English, the [ɔ] > [æ] change that seems to be slowly occurring seems to have affected the morpheme *trans* ([træns]) but not *plane* ([plæn]); hence *transplant* often has two different vowels ([trænsplænt]).

gemination to *tutta* in Italy and preservation as an unvoiced consonant in Catalonia (*totz*) and Northern France (*toute*). Posner suggests that in France, at least, TOTA saw the change coming in words already affected (e.g. *VITA*, *NATA*, in which [-t-] voiced and eventually disappeared, > *vie*, *née*) and shied away in horror; "one must assume that intervocalic surds have been reinforced, so that there is no longer an input to the voicing rule" (1974: 108).

We can hardly claim that every regional dialect which fails to undergo a change suffered in the dialect of a cognate neighbour has positively resisted such a change; but the sporadic strengthening of some words in the face of a change, when they seem to have deliberately disqualified themselves from the conditions of that change, can only have occurred if the theory of lexical diffusion holds, or else by the time TOTA saw what had happened to *NATA* it would be too late. The theory does not explain why it happened, but it does explain how it was possible at all. French *toute* is in fact a more interesting case than Italian *tutta*, since so few Tuscan forms did eventually voice, and some Italian words generated — under different stress conditions — without the threat of such a change (e.g. *FEMINA* > *femmina*). (Wanner and Cravens 1980 also reject the usual "facile recourse to dialect mixture" to explain the voiced forms in Tuscan.)

One non-phonetic reason which might lead a word to prefer not to change, and hide at the back, is what Samuels called "phonaesthetics" (1972: paras 3.10, 5.4 and 7.2) and others have called "sound symbolism". One example discussed by Samuels concerns the English change of *er* to *ar* as has happened in such words as *clerk*. An exception to this was the word *swerve*, which remained as it was; the possibility of phonaesthetics arises here because the central sequence may have some kind of psychological connection with meanings of circular movement, as in *twirl*, *whirl* or *whirr*. A variant of this can concern onomatopoeic words, whose motivation might run the risk of being lost after the change. Dworkin (1975: 468) has suggested that the mildly surprising decision to choose variants with the inter-vocalic consonant *g* ([ʒ], now [x]) in Spanish *rugir* (to roar) < *RUGIRE* and *mugir* (to moo) < *MUGIRE*, compared with e.g. *ruído* < *RUGITUM*, could be attributable to the desire to preserve suitably loud consonants in such words (although the existence of English *moo* shows that this is not an overwhelming tendency). This might also apply to *maullar*, "miaow", since the possible [aw] > [o] development would have lessened its felinity. This word first appears in Spanish orthography in the *Glosario del Escorial* (2484, *CATELO por maullar como gato*) of c.1400, but it

not ... (Chambers and Trudgill 1980: 176). Not all scholars consider this to be particularly important, and until we can gather why words change in the order they do it is not particularly illuminating. The theory does not imply, as some of its critics apparently think it does (e.g. Harlow 1975-76), that the order is random. It has hardly been applied to Romance at all yet; Wang (1977) edited a collection of articles on various phenomena in different languages, none of which are Romance. Wang himself is centrally concerned with Chinese, and, despite using occasional data supplied by Malkiel, has no interest in Latin or Romance; but the comments above on "dialect mixture" apply well to Early Romance, in which the traditionally postulated surviving "Latin" has been one such dialect to be wheeled on when the "rules" have too many recalcitrants.

"Lexical Diffusion", as a simple statement of the fact that in many sound changes words do not all change at once, is probably the dominant view by now. Chen even suggested (1972: 473) that it is so obvious that the burden of proof rests on its opponents; one of these would have been Saussure, who regarded change as affecting sounds irrespective of the words they happen to be used in, but the evidence is not obviously on his side. The lexicon is apparently destined to assume a much greater importance in linguistic theory of the 1980s than in that of the 1960s (e.g. Bresnan et al. 1978), and the theory of lexical diffusion is merely one aspect of a wider realignment. Meanwhile, with regard to our problem of phonetically retarded "learned" words that seem to have existed in speech before the Carolingian reforms, lexical diffusion has permitted us to suggest that these words might merely be words that were at the end of the queue, which the change never reached. Why words should go to the end of the queue is the question that needs to be answered next.

Strengthening and Phonaesthetics

The idea that some words seem deliberately not to want to undergo a change that is otherwise general in the language is not in itself new. For example, as a general point, Zonneveld (1978: 261) suggested that "it may become plausible to at least entertain the idea of a sound change being competed against by its own converse", as hypercorrection; and within Romance Posner (1974: 106-09) has implied such a reaction in her comments on the strange phenomenon of the occasional "strengthening" or doubling of some intervocalic Latin consonants that refrain from voicing, comparing for example the voicing of the [-t-] in Latin *TOIAM* to *toda* in Spain, predictably, with its

undoubtedly existed earlier; for example, it appears (spelt MEOLARE) in a delightful list of animal noises compiled in tenth- or twelfth-century Moslem Toledo (Castro 1936; Díaz y Díaz 1976b: 154). Such strictly onomatopoeic words may seem to have a greater motive for recalcitrance than the psychological phonesthesia of *swerre*; neither would have sufficient strength to resist the steamroller effect of the neo-grammarians' sound laws, but both can be seen as plausible reasons for the words concerned being further to the back of a diffusion queue than other words unaffected by such considerations.

Archaism and reluctance to change is not the only consequence that has been envisaged for phonasthetic considerations; Malkiel (1980b) has even suggested, tongue apparently not in cheek, that the excessive assimilation in the opening of the initial vowel of Sp. *marzavilla* < MIRABILLA, where the most closed Latin vowel, [i], became the most open Spanish vowel, [a], is due to what he calls "phonosymbolism": marvels leave their witnesses open-mouthed. Any explanation is perhaps better than none! Sound symbolism is taken by Samuels further than it might be thought plausible to take it, but if all we are looking for are the reasons for the order within a queue we already have reasons to believe exists, even the most minor considerations might turn out to be relevant.

Homonymy

a) Clash

Romance philology made one of its rare contributions to linguistic theory when the dialect geographers proposed that some strange phenomena could be attributed to a desire to avoid homonymic clash. When two words are destined to become indistinguishable if the ordinary evolutions were to occur, it often happens that one or both of the words either disappears or evolves in an unpredictable manner. The standard and original example concerns Gascon farmyards, where CATTUS 'cat' and GALLUS 'cock' were both destined to lead to Gascon *gat*; but that form is now confined to 'cat', with the 'cock' being referred to with a variety of different words entirely (cp. Ullmann 1967: chap. 7). Such cases of the disappearance of one of the lexical units are sufficiently well attested for homonymy to be plausibly offered as the reason. Within Spanish, this is visible in such cases as ALT > o, UBI > o > ø; ET > y, BI > y > ø; FENCULUM > *hinojo*, GENCULUM > *hinojo* > ø (except in the unambiguous phrase *de hinojos*, 'on one's knees'). UBI has been replaced by *donde* < DEUNDE, BI by

allí < AD ILIC, and GENICULUM by *rodilla* < ROTAM + the diminutive (cp. ROTAM > *rueda* 'wheel').

It has been occasionally argued that prospective homonymy is a reason for lack of evolution, on the assumption that this prospect could be sufficient to keep the word back in the first place. But it has in the event proved difficult to find incontrovertible examples in which a homonymic clash has prevented a change from happening at all. Speakers cannot foresee a clash before it arrives, nor hear it until it has been uttered. The above Spanish examples in fact confirm that although homonymic clash can lead to subsequent readjustments, they do not regularly prevent the development that caused it. There was, for instance, no particular reason in the first place for the loss of the intervocalic [b] of UBI or BI, the development that precipitated the clashes with ALT and ET; [b-] survives more often than not (e.g. *subir* < SUBIRE, *habe* < HABERE, ET > y is a similarly unusual change, since [j] > [e] is commoner than the reverse. The loss of the aspiration in FENCULUM > *hinojo* > *hinojo* was isolative, for [h] does not exist in Standard Modern Castilian and the word could not therefore survive as such, but strengthening to [f] might have been a way of escape. Vincent has argued, in a brief but brilliant study, that since in such cases the sound changes concerned were not stopped in time to halt the creation of the pathological state, all the traditional examples of homonymic clash are 'counter-examples to the view that sound change is itself directly teleological' (1978: 416). The clash is on us before it is foreseen. In another influential study, Andersen concluded that teleology is acceptable as a 'teleology of function' rather than a 'teleology of purpose' (1973: 789), but the point cannot be hidden: sound changes do not occur because the speakers have a positive desire to arrive at their destination, even if they sometimes occur because speakers have a negative desire to leave their starting-point. If the desire to avoid clash were as strong as it has sometimes been presented, FLAMMAS, for instance, would have followed its phonetic analogues and survived as *flamas* in Spanish, since *flamas* < CLAMMAS has become an exact homonym for *flamas* < FLAMMAS.

This is unfortunate for our desire to explain why some words go to the back of the queue for a sound change; an initially promising and traditionally trumpeted cause célèbre turns out to be little more than a damp squib. However, one of the diffusion theorists has considered this question. Although he agrees with Vincent, that 'homophony-inducing sound-changes will cause compensatory mechanisms to develop...' subsequently, Lyovin (1977: 132) also suggests from his own evidence

that until such mechanisms are developed the standardization of the form of both the relevant items might be slightly delayed. Given this, the general pattern might be for homonymic clashes that arise, but are thought intolerable, to lead to irregularities of various kinds; and one of these irregularities might, if the old form is still synchronically acceptable, be a subsequent reversal of the change: i.e. not the prevention of a hypothetical change $[x] > [b]$, but its being followed by another change that happens to take the form $[b] > [x]$. This is observable in the Modern English pronunciation of *guerrilla*; aware of unfortunate consequences of failing to distinguish *guerrilla* from *gorilla*, given that $[ga]$ is the naturally evolved initial syllable of both, the BBC has self-consciously adopted $[ge-]$ for *guerrilla*. As a result this appears to be comparable with a case of strengthening, for surely the $[ge]$ cannot have been delivered so emphatically before this therapeutic reaction. Within the next century, unless terrorism disappears, the BBC is likely to have a similar problem with *gunmen* and *government* $[ganmen]$. I have caught myself doing the same when lecturing on the additions to the 1502 editions of *Celestina*, distinguishing *addition* from *edición* with an emphatic $[a-]$. Blaylock (1973: 53) rephrased this phenomenon as a decision to choose unevolved variants: "If a tendential loss is to be thwarted, it can only be through the eventual favoring of variants in which the loss never occurred." This recognition of the coexistence, at some time, of both old and new forms in the community as a whole, is a key idea developed in the next section.

One minor group of words might have the traditional theory of homonymic clash available still to explain retardation; those in which the homonym avoided is taboo. The case of the avoidance of $[ʃɪt]$ in the unchanged English forms *shut* and *shuttle* is mentioned by Samuels (1972: 143). Exactly the same mechanism has been thought to explain the non-diphthongization of stressed $[e]$ and the retention of internal atonic $[i]$ in the Spanish town of *Mérida* < EMERTAM, a provincial capital in Roman times; the desire not to be associated with *mierda* < MERDAM. Dialectal explanations are less implausible than usual here, however, judging from what we know of Western Mozárabe. On the other hand there is the delightful case of the citizens of Coyaanza complaining to the King about the embarrassing name of their town in c.1208⁶, this problem was solved not by retarding or reevolving the

6. I am grateful to Derek Lomax for drawing this to my attention. The petition to Alfonso IX of León is mentioned by one Jimeno de Daroca in Sancho's y Sivera (1921, II: 374). The King changed the name and imposed a fine of 100 *solidi* on anyone who spoke the old name. Coyaanza, on the Esia just south of

name to distance it further from *cojones* (< COLEONES, "testicles"), but by changing it entirely to Valencia de Campos, and eventually to Valencia de Don Juan.

b) The Formation of Doublets

The avoidance of the production of homonyms is probably a redish herring; but the application of sound changes to already existing homonyms and polysemic forms can have the effect of appealing to retard phonetically one of a pair of homonyms or part of the meaning of a polysemic form. In the case of two distinct homonymic words, it is reasonable to expect under diffusion theory that one will start to change before the other. In the case of a polysemic word – a word with more than one meaning – the mechanism for such differentiation is less straightforward but no less real. In each case, the effective result is the production of what are often known as "doublets".

Sometimes two Romance descendants from the same one etymon coexist in the same modern language. For example, Latin *PENSO* gives both Spanish *pienso* ("I think") and *peso* ("I weigh"); *MUNDUM* gives both *mundo* ("world") and *mondo* ("clean"); *RATIONEM* gives *razón* ("reason") and *ración* ("ration"); *CATHEDRAM* gives *cadéra* ("hip") and *cátedra* ("University Chair"); *LITIGARE* gives *lidiar* ("fight") – now confined to the bullfighting and *litigar* ("litigate"); *RADIUM* gives *rayo* ("ray"), *radio* ("radius") and *radium* ("radium"); etc. In some of these cases it is clear that the less evolved variant has been consciously borrowed from the old Latin form at some time later than the arrival of the twelfth-century Renaissance into Spain. *Radium*, for example, was recently borrowed, as in English ("radium"). *Radio* was borrowed ("radius") at an earlier time when the endings of Latinisms were normally adjusted to the norms of Spanish morpheme structure. *Cátedra* was a borrowing from educated Latin, probably in the thirteenth century. *Litigar* was a literary transference made by e.g. Juan de Mená in the fifteenth century. In these cases the origin of the less evolved form is obvious and indisputable; *rayo*, *cadéra* and *lidiar*, the more evolved forms, were in the vernacular all along, but *radio*, *radium*, *cátedra* and *litigar* were comparatively recent borrowings.

Unfortunately it has become traditional to use these words as the model for cases in which both forms are already in current use at

León, was an ecclesiastical centre, variously spelt with *i*, *y*, or *gi*, probably representing a $[j]$. For the Council of Coyanza in 1055, see Chap. V. COLEONES in Old Leonese would have given both $[kojones]$ and $[kojones]$, so the connection would indeed be perceptible. For Leonese, see Zamora Vicente (1970: 146-49); for COLEO see Celia (1969).

the time of the earliest texts in reformed Spanish orthography (c. 1200). In the same way as *litigar* is demonstrably a borrowing from Latin rather than a word existing in that form in the vernacular throughout, it has been assumed that surviving pairs of descendants from the same etymon are ascribable to comparable causes; that the more evolved form was used in the vernacular and the less evolved form was used by the educated in the "Latin" they were thought to speak. Thus it is that *pensar*, *mundo* and *razón* are said to be "culto" forms used by the educated, and *pesar*, *mundo* and *razón* to be "popular" vernacular words. In this way, doublets were seen as clinching evidence for the two-norm theory.

There is, however, a reasonable linguistic explanation for such doublets that does not depend on the postulation of a hypothetical Latinizing pronunciation; not only that, it has a recognized parallel in the development of English, wherein the same phenomenon is dealt with by philologists in a different manner.

It is generally realized that language change is neither uniform nor sudden. A development takes time to become established, within both the community and the individual. This means that while a change is in progress, both the old and the new form of a word will be encountered, and either form will be functionally sufficient. For example, during the progress of the diffusion of the [ʃ] pronunciation, with nasalized vowel, instead of [ens], *MENSA* would sometimes have been uttered with and sometimes without a consonantal [n], and either form would have been intelligible to all speakers, regardless of which form they preferred to use themselves. Normally this state of variation tends eventually to resolve itself into the survival of one form (e.g. > Sp. *mesa*); the surviving form will usually be the newer form, since whatever reason it was that led to its original emergence is likely to continue to be valid at the time of the decision to standardize one form at the expense of the other. There are, however, cases in which the retained form is the original form, and we appear to be confronted with a change that reversed direction in mid-flight; a much discussed case is that of Latin *NOCTEM* > O Sp. *noche* > *noch* > *noche*; *noch* was apparently a common twelfth-century variant which eventually failed to become the standard form (e.g. Lapesa 1975; Hooper 1974: 126-30).

If, however, for any reason, both the coexisting forms survive, with different meanings, then we have an institutionalized pair of doublets. This will only happen in practice should there be a sensible reason for speakers to prefer the survival of two forms rather than one. In cases of homonymy — two words with the same pronunciation — or

of polysemy — one word with more than one meaning — there often is a clear advantage in preserving two forms, in that the previous potential ambiguity can thereby be resolved. This has happened in English, for example, to the word *person*. The change of *-er-* to *-or-* led to the co-existence in functionally free variation of at least two pronunciations for the affected forms. Most English words eventually chose the evolved form (*clerk*, *clergy*), although some chose to remain with the older form (*swere*, *clergy*), but *person* and *parson* both survived. This survival is almost certainly due to the usefulness of being able to distinguish people in general from vicars in particular, and we need not posit the existence of systematically distinct dialect systems to explain it. (In addition, the fact that *parson* happens to be the more evolved form is enough to demonstrate the falsity of the idea that ecclesiastical terms are *ipso facto* more archaic in phonology than the general vocabulary.) In practice no one has thought of ascribing the older forms of such doublets to a special retarded norm of educated pronunciation in English philology, and if anyone did they would be laughed out of court (Waldron 1967: chap. 3).

The idea that doublets are evidence of two systematically distinct pronunciation norms ought to imply that each word of such pairs meant the same thing. Otherwise we will have to postulate that educated and uneducated people talk about systematically distinct subjects; for example, that in early Medieval Spain only the uneducated talked about weighing or sorrow (*pesar*) and only the educated talked about thinking or feeding horses with *pensar* (see below). If we concede that even illiterate people are likely to have used this common word to talk about thinking or feeding horses, we must also concede that in Medieval Spain they used a "literate" form when they did so, since even the earliest attestations regularly include the [n] or the manuscript *ñde*; and we can hardly avoid conceding that the literate used the *pesar* lexical item, since there it is, abundantly attested in early medieval Spanish literature. This systematic separation of subject matter according to the relative level of education in different groups is unconvincing; the two-norm explanation of doublets raises more awkward questions than it solves, despite its apparently charismatic initial attractions.

Homonymy is common in many languages, and is not usually in practice a cause of great ambiguity or awkwardness. If the chance arises to avoid it, however, it is easy to see why it might be thought sensible to differentiate the items in some way, within a lexical diffusion image of linguistic evolution, such a differentiation can be easily achieved by

having one of the items near the front of the queue and the other at the back. Since the meanings are semantically separate in homonymy, no problem arises there. Chen has documented some startling evidence of extensive and pervasive "homonym splits that cannot be accounted for either as phonologically conditioned divergences or as a result of inter-dialect borrowing" in Chinese (1972: 492); if this is an accurate diagnosis, we need not be too chary of proffering comparable evidence in Romance.

For example, Latin *MUNDUS* was homonymic, with the two meanings of "clean" and of "world". Originally short Latin [u] regularly develops to Spanish [o], as in the second syllable of these forms; this has happened in Spanish *mundo*, "clean" (and its cognate verb *mundar*, "to clean"), but the surviving form for "world" is *mundo*. According to the rules engraved in the philological handbooks, therefore, the [u] of *mundo* is under-evolved, and *mundo* is thus categorized as "culto". Thereby the impression has been given that the uneducated in Spain rarely or never used this form when referring to the world. But such a conclusion can be avoided if we suggest that during the inevitable period of free variation a gradual decision was taken to factorize out one meaning per form, and thereby permit each form to survive with its own no longer confusable meaning. The phoneme /u/ of *mundo* is not at all strange phonologically in the superseding phonology, being the regular development of the originally long /ū/, and the form has no reason to be thought undesirable. There are conditions under which Latin [ū] becomes Spanish [u]; e.g. before laterals (e.g. *CŪTELLUM* > *cuchillo*, *CŪLMINE* > *cumbre*, *SŪLPUR* > *azufre*) and before a yod (e.g. *FŪGIO* > *huyo* etc., *CŪNEAM* > *cuña*, *ŪNGULAM* > *uña*). Other examples include *cruz*, *culpa*, *duda*, *lucha*, *trucha*, *mucho*, *una*. *nunca*. Ū > u is not in itself avoided or considered unacceptable.

A similar case to *mundo* and *mundo* in Modern English concerns *blooming*; at present, given the [u] / [u:] variation in Britain, it usually has the long [u:] when it is used to refer to flourishing vegetation, and the short [u] as an adjectival expression of mild annoyance. Consider, for example, the likely pronunciations represented by:

- 1) The ivy is blooming in my garden.
 - 2) Get that blooming ivy out of my garden!
- When the orthographic revolution comes, the doublets will for the first time be attested in writing. The semantic separation will by then be ancient, but any postulated accompanying separation between educated and illiterate usage will be chimeric. Both the literate and the illiterate can and do get annoyed; both can and do discuss flourishing vegetation.

Cases of polysemy are not likely to resolve themselves like this so simply. Polysemy is normal in language, and not in itself pathological. The fact that many words have more than one meaning, or vagueness of potential reference, is often a pragmatic advantage, since if it was always essential to be precise in speech people would hardly ever talk at all. Even so, polysemy can lead to a word becoming unacceptably vague, ambiguous, or even misleading, and should a sound change begin to affect the form at a time when speakers are subconsciously becoming aware of the existence of an excessive range of semantic possibilities connected with the form, the same separation into two forms with distinguishable meanings can eventually be achieved.

The Spanish adventures of *PENSARE* are, in fact, of particular interest in this connection. Latin *PENSARE* meant originally to "weigh"; the Romance languages have words for "weigh" that presumably derive from a Proto-Romance form without the nasal (Sp. *pesar*, Cat. *pesar*, Ptg. *pesar*, It. *pesare*, Fr. *peser*, etc.). A metaphorical extension of this, meaning to "depress, grieve, weigh down with sorrow", seems to have been particularly widespread in Spain; in the nominalized forms, a further distinction arose between *pesar* "sorrow" and *peso* "weight". The verb remained potentially ambiguous between "weigh" and "grieve", although in practice a problem would very rarely arise since "grieve" requires an animate object and "weigh" usually took an inanimate one. The adjective *pesado*, however, is still occasionally ambiguous between "heavy" and "wearisome". As well as this extension of the semantic area covered by *PENSO*, there existed a particular use confined to the weighing out of animal foodstuffs. *PENSO* in this meaning developed a separate use in which the object of the verb was not any more the fodder (Sp. *pienso* < *PENSUM*) but the animal itself. It is possible that a potential semantic confusion between weighing a cow and feeding a cow could have been felt worth avoiding by the eventual decision to specialize the "fodder" meaning into the variants containing the consonantal nasal. In this way, Spanish *pensar* *las vacas* can only mean to "feed the cows". But this was not all. A further metaphorical extension from "weigh" arose to become, in some areas at least, apparently the statistically commonest use of *PENSO*; the sense of "mentally weigh up", "consider". This meaning has survived in the Romance languages with a nasal consonant (Sp. *pensar*, Cat. *pensar*, Ptg. *pensar*, It. *pensare*, Fr. *penser*, etc.), which should suggest that the Proto-Romance form had the nasal consonant. (This choice does not lead to Spanish confusion with "to feed", since "to think, intend" does not take a direct object.) This conclusion concerning Proto-Romance

has not been the usual analysis, however. The traditional analysis of *pensar*, etc., was that at a later date every Romance area decided to borrow this form, exclusive to "Latinate" speakers, into their vernacular language; by coincidence every Romance area seems to have done this at approximately the same time, and by further coincidence they all chose to give it more or less the same meaning; and by a coincidence that verges on the miraculous, they all decided to give it a meaning that it did not regularly have in Imperial Latin. This analysis overlooked the awkward fact that the Spanish forms that preserved the nasal, and were thus declared to belong to the Latinate, were exactly the same forms as the ones that diphthongized the [e] in the vernacular manner; whereas the Spanish forms that lost the nasal, and were thus declared to be vernacular, did not diphthongize.

This latter failure to diphthongize in *peso*, but success in *pienso*, is legitimately surprising. The sequence *-ens* in Old Latin was regularly "long", even in words whose immediate cognates were short — *CP. INGENS, INGĒNTIS; DENS, DĒNTIS, DĒNTIO* —, which means that technically *PENSO* > *peso*, *MENSAM* > *mesa* are, from the point of view of the vowel, "regular", and *PENSO* > *pienso*, *TENSUM* > *tieso* ("stiff") "irregular". Whatever the reason, the Romance evidence shows that after the loss of phonemic length the vowel in *PENSO* was open enough to qualify for diphthongization in Spanish. The diphthongization of [e] began early (as can be seen from the comments of Pompeius quoted in Chapter 2), and may have been early enough to coincide with the [os] ~ [s] variation. The fact that in this possible simultaneity of free variations both affecting *PENSO* some forms chose the evolved [je] and the old [n], and the other forms chose the old [e] and omitted the [n], suggests that the choice of variants was not made on the criterion of which variant was the older and which was the newer. It is unlikely that many speakers knew whether the [n] forms were older or newer, or whether the [j] forms were older or newer, or considered the information relevant to their own choice of usage if they did. The average Briton neither knows nor cares which is the older of [kap] and [kʌp], etc., and if he has a reason for choosing one rather than the other it is usually connected with neither his relative literacy nor his relative conservatism.

There are so many difficulties arising from the traditional explanation of *pensar* and *pesar* that it is bewildering to find that this word has become the standard oft-repeated example of a "cultismo" in contemporary textbooks. Hall (1974: 107), for example, used it as the paradigm case and even theoreticians of historical linguistics cheerfully

repeat this awkward case as the standard example. Anderson (1973: 185), for instance, holds it aloft, despite knowing Spanish reasonably well.

Other doublets can be explained in the same way. The desire to resolve homonymy can be invoked to explain, for example, the medieval pair of *obra* and *huedra* (< *OPERA*). *Razón* and *racón* (< *RATONEM*) is another fairly clear case, although the waters are muddied by the apparently deep-rooted idea that *-ción* and *-encia* are in some way necessarily "learned" forms rather than being, as they are, much the commonest medieval Spanish result of *-ATONEM* and *-ANTIAM*. Sometimes they occur in free variation with the *-ción* and *-anza* forms that were given the accolade of "regularity"; Patison's evidence (1975: 93) concerning forms with verbal roots has suggested, however, that the *-ción* ending is so much commoner than *-ción* that it can plausibly be regarded as the statistically normal vernacular form. Of 149 derivatives he listed, only 2 are attested with the *-ción* form alone, *cañazón* and *háygon*, and these are rare words, which were used in literary circles anyway (Berceo and the *Primera Crónica General*). In Medieval Spanish, *razón* was a very common word with an unboundedly generic meaning — "thing", "words", "affair" — which it has since lost (Lancheiras 1900: 626). Lapasa still apparently thinks that "cultismos" are likely to be generic; here is a case of the supposedly non-"culto" form gaining a bafflingly "generic" use while the supposedly "culto" form has a comparatively precise use. The four attestations of *racón* in the *Poema de Mio Cid*, for example, have recognizably the meaning of "part", "share" (lines 2329, 2467, 2772, 3388).

This concept, of doublets exploiting free variation for semantic convenience, applies equally well to doublets that are not of Latin origin in the first place. The reconstructable Germanic root *RAUP*, for example, "pillage", provided a noun (> Sp. *ropa*) and a verb (> Sp. *robar*). It is probable that the word was borrowed at a time when [p] and [b] were in variation in several parts of Iberia; the preservation of the unvoiced [p] in *ropa* commits no solecism against the phonemic system — cp. the ALP > *op* postulated in SAPUŪ > **SALPIT* > *sopo* > *sapo*; or *copa*, from the geminate CUPPAM — and in view of the semantic divergence of *ropa* ("booty") > "clothes" and *roba* ("pillage") > "rob" there is no semantic value in maintaining cognate transparency. Here we apparently have a phonetically retarded word which unfortunately happens not to be Latin, so calling it a "cultismo" is impossible.

7. If *uncia* were "culto", how could it combine with a Germanic root in *garrañca*?

The differentiating mechanism envisaged here can even be exploited consciously (like the [ge] in *guerrilla*, above); the current Spanish spelling of *barón* "baron" and *varón* "male", identically pronounced, exploits a free variation in spelling that does not correspond to any phonetic distinction to separate the polysemy that has arisen from the originally Germanic *BARO*, "noble male". The spelling of English *flour* and *flower* achieves the same result. One of a pair can also change gender for the same purpose; e.g. Sp. *el mañana* ("tomorrow"), *la mañana* ("morning"); German *der Schild* ("shield"), *das Schild* ("sign").

The doublets once seemed to be self-evidently convincing evidence in favour of the two-norm theory. The existence of comparable phenomena in English, and of doublets of Germanic origin in Romance, show that they are not; the fact that the doublet forms can be explained with recourse to established principles of historical linguistics, and without recourse to hypothetical Latinizing speech systems, shows that the two-norm theory cannot use the doublets as a crutch. If the two-norm theory is to survive, it must be for other reasons than that.

Subsequently Obscured Regularities

It is likely that in some cases a statistically regular conditioned sound-change has had its statistical regularity obscured from the eyes of later scholars because of the accretion of new vocabulary that arrived too late to be affected. For example, it now seems certain from the surviving graffiti that several of the well-known developments from original Latin to Old Romance were already well under way by the time of the eruption at Pompeii (79 A.D.); so it is reasonable to propose that some of the new Christian vocabulary borrowed from Greek in the fourth century and later might in fact have turned up at a time when the conditions for phonetic changes already suffered by analogously spelt words had ceased to have evolutionary consequences. Lapesa was quoted above as exemplifying "cultismo" with *ángel* < *ANGELUS* < *áygelos*; the unexpected loss of final [-o] makes this attribution to Latinate speakers implausible anyway, but the reason for the attribution of "cultismo" is the retention of the atonic interconsonantal vowel, which is often lost in comparable words; cp. *SINGULOS* > *señalos*, *QUINDECIM* > *quince*, etc. This syncope, however, is likely to have been well under way by the time of the wide Christianization of the Empire, and the [e] of the new Graecisms could well have been phonologically distinguishable from the short [ə] of the forms with which *ángel* is

compared at the time [ángelo] (or [ándzelo]) was first uttered by monolingual Romance Christians. These Church words are as often over-evolved as under-evolved, in fact; the [-o] > ø of *ángel* is glossed over by the "culto" analysis as neatly as is the strange [e] > [o] of *obispo* < *EPISCOPUM*, which has in fact lost the syncopeated interconsonantal vowel, but is called "culto" because the [i] did not become [e]. Given that the time of borrowing is unclear and the stage reached by any phonetic change at any specified time is also unclear, few such words are likely to be clarified surely in the light of this possibility; on the other hand, we cannot necessarily expect new vocabulary to fall under the scope of conditioned changes that have by the time of the borrowing lost their vigour. It seems that some did and some did not.

I have examined elsewhere (Wright 1976c) a more recent example in which the existence of a conditioned development can be glimpsed in Medieval Spanish, but that development has not normally applied to an untypically large number of comparable words borrowed subsequently; the conclusions alone are presented here. (Malkiel 1980a comes to a contrary conclusion concerning counterexamples.) It concerns the fate of the pretonic diphthong [wa], which in Old Castilian is only found after initial velars (e.g. *cuaderno*, *guardar*). In Latin the [w] was part of the /kʷ/ phoneme; /gʷ-/ did not originally exist, but arose in order to accommodate Germanic vocabulary with initial /w/. Pretonic [wa] was the only pretonic rising labial diphthong; before all other vowels the [w] had dropped long before (e.g. *QUEM* > *quien* [xjen]). Partly as the final fling of that development, partly perhaps in the wake of a contemporary decision to confine [je] to tonic position (e.g. *GERMANUM* > *hermano* > *hermano*), a change began which lost the [w] in pretonic positions (e.g. *catorce*, *galardón*). The philologists stated that [wa] > [a] was the rule, and thus surviving pretonic [wa] was said to be a sign of "culto" vocabulary confined to educated usage. Hence words such as *cuaderno* and *cuarenta* were said, on this evidence alone, to be "learned" forms. *Cuarenta* has developed considerably from *QUADRAGINTA*, but that was not thought relevant; *cuarenta*, like *guardar*, was exceptional.

One word, which had a Latin *QUA-*, is in Spanish both an undeniable Latin borrowing and a word which has lost the [w]. *Casi* "almost", borrowed from *QUASI*, apparently c.1400, is a "cultismo" without a [w]. This ought to cast doubt on the wisdom of using the retention of the [w] as a sure symptom of Latinate usage. Two other words make the picture even less certain; the strange case of *guadamascí*,

borrowed from an Arabic word in which there was no [w], and the stranger case still of the adjective *cuadril*, formed off the noun *cadere* (< CATHERAM: there was never a [w] in this root), in which a pre-tonic [w] has actually been inserted.

What seems to have been happening is that for several centuries the change of [wa] to [a] was gradually working through the list of morphemes. The words at the front of the queue for this change included all cases in which there was the condition of a following [s] or [z]: literally all, whether tonic or not (*cascar* < QUASSICARE, *gascar* < VASTARE, *garcones* < VASCONES); and (with specifiable exceptions mentioned below) all cases in which the syllable was not originally immediately pretonic (ante-pretonic: not necessarily counter-tonic): QUALITATEM > *calidad*, QUANTITATEM > *cantidad*, QUATTUORDECIM > *cuore*, Germanic *WITRALAUN > *galardon*, Germanic WRANON > *garçon*. *Casi* is a pre-clitic word in which the initial syllable was regularly ante-pretonic, so when borrowed, even though new, it could plausibly have shot straight to the front of the diffusion queue. Words to the back of the queue include all words in which the [a] is followed by a [ð]: *cuaderno*, the Germanisms *guadaflores*, *guadalupe*, *guadanyar* (Aragonese), *guadapero*, and the many Arabic-based toponyms such as *Guadalupe*, *Guadalupe*, *Guadalupe*, in some of which the initial syllable might once have had full stress anyway; all words in which [a] is tonic: *guante*, *guay*, *cuadro* and cognates, *cuál*, *cuadro*, *cuanto*, *cuarto* and cognates, *cuatro* and cognates; and all words with a following [r] (except for ante-pretonic *garraón* mentioned above): *cuarenta*, *cuarenta*, *guarlar* and cognates, *guarir* and cognates, *guarir* and cognates. This is not particularly neat, but it is a possible form for a diffusion queue to have taken. Given that the evidence of indigenous forms shows that both a following [s] and ante-pretonic position encouraged priority for the change, it is understandable that the newly borrowed *casi* should go to the front; given that the evidence of indigenous forms shows a strong preference for [wa] rather than [a] before a [ð], it is at least understandable, even if unpredictable, to find the newly borrowed *guadameci* acquiring an etymologically inexplicable [w], and the newly formed *cuadril* doing the same. Even now there is only one word with initial [gað-] in the Academy Dictionary: the rare (and Latinate) *gañano*, 'citizen of Cádiz'.

If the vocabulary of Castilian in c.1500 had been approximately the same as that of c.1500, this conditioned regularly would probably have been noticed by Romance philologists. During the decades following 1500, however, Castilian took on board a large number of

words from various American Indian languages, including a remarkably high number of words with initial [ewa]. There were at least 25 words, both with initial stress and initially atonic, and several of them included a following [s]: e.g. *guano*, *guacamayo*, *guacho*, *guagua*, *guayaba*, *guasaca*, *guasaca*, *guaso*. The number of words with [g] or [k] + [wa] had never been high, and the original ones were outnumbered by the newcomers. The regularity that a fifteenth-century Castilian might have been able to intuit dimly as a morpheme structure rule had gone by 1600, and the diffusion of the change stopped where it was. Those at the latter end of the queue never arrived, and almost certainly never will, but those at the front had changed and have never looked back. Modern philologists, lacking the time to hack away the branches and perceive that a kind of regularity did once exist, peered through the undergrowth, perceived chaos, and cut the knot with the well tried formula, that pretonic [wa] > [a] was the norm and therefore the [wa] forms were exceptional (Menéndez Pidal 1904: paras 39, 68; 1944: 176-77).

There might well have been a number of such cases in which a change is methodically working through the diffusion queue and is then suddenly stopped in its tracks and rendered unforeseeably impotent by the action of such extrinsic forces as the sudden accretion of extra vocabulary. Vocabulary borrowing does not occur at a uniform rate. New vocabulary can either succumb to changes under way, integrating as neatly as *casi* or the [o] > Ø of EPISCOPUM, or it can discontinue the progress of that change, as in *guaso*, or the [e] (> [o]) and the survival of the [i] of EPISCOPUM. The latter is likely to be the result of the arrival of sudden large chunks of new vocabulary such as the Graecisms in the fourth-century Church, or the Indianisms in sixteenth-century Spain: the former is likely to be the result of the occasional assimilable individual borrowings that occur all the time. In the light of the effects of large-scale vocabulary influences, the attribution of 'Latinate' archaizing speech levels either to the new vocabulary, or to old vocabulary undergoing a development whose progress is affected by that new vocabulary, is likely to be unjustified if made simply on the grounds of historical phonetic peculiarity.

Postdating Latinisms

There can be no doubt at all of the coexistence of two pronunciation norms, at least for those with some knowledge of official liturgical practice, after the operation of the Carolingian Reforms (in France after c.800; in Spain, outside Catalonia, after 1080): the only

DULCIS had been in a dilemma considerably worse than that faced by the FL- words discussed above, or even than LAMPIDUM. It was not at all clear which sound changes it was supposed to be due for. The [k] before the front vowel evolved to [ts] without any problem; that was achieved early. (The [ts] seems to have been a regular feature of Hispanic reformed Latin also, whatever the Carolingians had originally decreed about written 'c' leading to spoken [k]: see Chapter 3.) But the -UL- and the -E were not so easy to deal with. If the -UL- was to be regarded as comparable to the -ULT- in e.g. MULTUM > mucho, which would be reasonable given the following [ts] in dulce, that would lead to an awkward [dujtsɛ] (or even [dujtsɛ]) in need of further resolution: hence written dulce, presumably representing [dujtsɛ]. If the -U- was to be regarded as a candidate for the normal [u] > [o] change, then dulce would be the result. If the -LC- cluster was thought to have a casting vote, DULCEM might have been expected to follow the pattern of FALCEM > *foze > hoz, the [l] thereby disappearing and a form doze resulting. The [ts] of [lts] would not normally have voiced in the presence of the preceding consonant — cp. LANCEAM > lança — but in the versions that lost the [l] the problem would arise of being thought to qualify for the queue for the voicing change [ts] > [dz], as happened in FACERE > O Sp. faser ([hadzɛ]). After which, about the eleventh century, there arose the development of the loss of final [ɛ]; this seems to have been dropped quite simply after single alveolar or dental consonants (e.g. RETEM > rede > red, etc.), but in other words the change seems to have reached a stage of free variation (e.g. corte/cort, noche/noch, dende/dend) which eventually resolved itself in favour of the original form. Final [ts] and [dz] remained without a subsequent [ɛ] (e.g. paz, juez, etc.), which is the model followed by duz, dulz, doz, dolz; but speakers using the forms in final [-lts] would not have felt easy, since in most words ending in consonantal clusters the [-ɛ] in fact remained; in addition all analogous words seem to have been verbs (e.g. the subjunctive of alzar [alɛzɛ]). In all finite verbs the [-ɛ] was reinstituted whatever the preceding consonant (e.g. fize 'I did', viene 'he comes'). Even if speakers who normally pronounced the word without an [l] then lost the [-ɛ] as well, that was not the final problem, because it looks as if there might have been some vacillation over whether [dz] and [ts] at the end of a word — a position which they had never occupied before the loss of [-ɛ] — ought to neutralize their opposition or not, as final [-t] and [-d] had. In the event, the modern view is that the spelling -z often came to represent a [ts] sound in the thirteenth century, long before the unvoicing of [dz]

problematic issue is whether or not the distinction between Latin and Romance has any validity before these reforms. Accordingly, any apparently Latinate form whose first use in Romance can be plausibly dated to a time subsequent to the reforms might remain attributable to "Latin"-speakers, as a borrowing from the new spelling-pronunciation institutionalized as the ecclesiastical standard. Such a form is not the property of archaizing conservatives but of innovating progressives, yet meaningfully "Latinate" even so. This can only apply to words that undergo none of the changes that have run their course before 1080 in Spain and 800 in France; the loss of the final syllable of ángeł, the backing of the first vowel in obispo, the reduction of -AGI- to [ɛ] in cuarenta, the diphthongization of pienso, etc., are sufficient to prevent many of the traditionally "culto" words from retaining their "culto" status by escaping through this loophole. Some of the evidence adduced by Rice for France, however, which led him to his categorical affirmation of the existence of Latinate speech in the seventh century (quoted above), is quite compatible with the idea that the forms concerned did not exist in Romance before the ninth century.

Malkiel (1975b) has offered a very interesting case of a word which had a large number of competing forms in Old Spanish, which were eventually all abandoned for a Latinizing pronunciation instead. Malkiel showed that dulce, Latin DULCIS, was in fact an innovation in twelfth-century Spain, a new form introduced to resolve the anarchy, confusion and embarrassment caused by the existence of at least ten attested competing variant pre-existing forms. This evidence, that a twelfth-century "cultismo" can be an innovation rather than an archaic survival, as later borrowings from Latin are, is of greater significance than Malkiel seemed to realize, and the case is worth analysing in greater detail.

Castilian dulce is clearly "retarded" for some reason. The preservation of the [l] and of the [ɛ], the retention of [u] rather than opening to [o], might be individually explicable, but collectively they suggest Latin influence of some kind. This influence, however, cannot be certainly ascribed to before the twelfth century. In the twelfth century the traditional orthography taught to scribes was still DULCIS; but there seems to have been a multiplicity of spoken forms, none of them sufficiently authoritative or widespread to be the obviously correct Castilian. Malkiel listed ten different attested written forms: e.g. dulz, doz, duz, duce. (Hartman has since argued (1979) that [dujtsɛ] could be one of the popular variants as well; in which case the arrival of Latin [dujtsɛ] had the effect of favouring an existing variant rather than introducing a new one.)

to [s] in other positions in sixteenth-century Castilian. The consequence of all this cumulative uncertainty was the simultaneous presence in twelfth-century Castille of a large number of variant pronunciations for the one word, all of them reasonable on the basis of analogous words, none of them authoritatively taken to be the ordinary spoken form. At this point the new Medieval Latin turns up, following the adoption of the Roman liturgy in 1080, offering yet another alternative, but in a reassuringly acceptable context. Speakers can occasionally become conscious of free variations and alternative pronunciations, and when they do, they often become awkward or embarrassed when of necessity they have to choose one form rather than another in speech; if there are a large number of available pronunciations, it is possible that the arrival of an apparently authoritative alternative could be used as a lifeline. In any event, [dulse] was taken into the vernacular and used increasingly in the thirteenth century. This was a case when "speakers favoured the culturally unimpaired 'learned' form on account of some collateral feature not immediately related to its erudite status" (Malkiel 1963:64, 162).

The suggestion that the arrival of Medieval Latin in the twelfth century had a deciding influence in cases of vernacular vacillation can be applied to phonology as well as to the lexicon. Penny's analysis (1972) of Medieval Spanish /f/ argues that [f], which was in the tenth and eleventh centuries either absent or an occasional allophone of /v/, become reestablished as /f/, a separate phoneme from /v/, in approximately the twelfth century. The labiodental [f] of French and other Romance languages may never have existed previously in Old Castille, for it has been suggested that a bilabial [p] corresponded there all along to the written Latin letter *f*; whether this is true or not, the Castilian phoneme could have been phonemically, according to Penny's evidence, some kind of fricative bilabial ([p], [hp]) or [h], such as survives in many rural areas still. This may have been an uneasy and awkward moment, when the dialect speakers were beginning to wish for some clearer form. In any event, the twelfth century is the period when French and French-taught clerics were dominant in educated circles; among other things they were doing their best to spread the pronunciation of official liturgical Medieval Latin to be used in conjunction with the Roman liturgy, newly adopted in North-West Spain from 1080. Penny suggested that in urban or educated Castilian of the twelfth century the [f] sound emerged "under the superstratum influence of French and Provençal speakers", to become a distinct phoneme again "on the introduction of large numbers of learned words"

(1972:482). The implication of Penny's comments is that the [f] is due to the pronunciation of French or Occitan. The influence of French or Occitan phonology in Spanish is not an initially attractive hypothesis, despite Lapesa's much-reiterated view (1975) that it was largely French or Occitan pronunciation that led to the [-e] > ø development. (Arabic /f/ was labiodental, and could have helped in such loans as *añérez*; [f] existed in Aragón, León, and probably *mozárabe* as well.) Many urban and educated Spaniards were self-consciously learning the new artificial Medieval Latin as it was taught by Frenchmen, but there is no reason to suppose that many learnt French; it seems more plausible to suggest that the /f/ arises not from French or Occitan itself, but from the European pronunciation of Medieval Latin. If this Latin is indeed new at this time, the (re)emergence of [f] is interpretable rationally; if there was a "Latin" [f]-preserving layer of speech in previous centuries, the problem remains.

The influence of French pronunciation of Latin might also have contributed to the decision to reverse the loss of [-e] in words ending in other than single alveolar or dental consonants and in verbs (RETEM > *rede* > *red*, COHORTEM > *corre* > *corr*, POTUI > *pude* > *pud* > *pude*, etc.); this reversal can be partly attributed to the phonetic conditions of syllable structure, and partly to morphological conditions, but (despite Lapesa 1975) not at all to the French, who apocopated throughout. The presence of a final syllable in the new Latin, however, could have been one factor to give the reversal in different forms at the end of the queue a push.

In twelfth-century Spain, according to the theory outlined later in this book, the introduction of Medieval Latin was something new. It was taught largely by Frenchmen. The European pronunciation of Latin can plausibly be seen to have been a serious factor in cases where the direction of the vernacular was in the balance at this time. In earlier centuries it is very doubtful whether a distinct Latin phonology existed, so it may not be justifiable to use it as a source of enlightenment. In later centuries, vernacular spelling could be looked to for a similar casting vote, but that was not in general use yet in the twelfth. For a brief period, in the twelfth century, it remains plausible to postulate the influence of Latin pronunciation on vernacular development in the way that it has traditionally been postulated for previous centuries. Latin-ATTO might, for example, have then helped decide vacillation in the vernacular between *-cción* and *-zón*. In later centuries, when at last Latin and the vernacular were coming to be thought of as fully separable languages each with its own speech and writing systems, the

influence of Latin on vernacular is noticeable in the growth of the vocabulary, but can only rarely be perceived in the phonology of individual forms. The coining of *catedra*, *litigar* and other Latinisms, for example, seems to have had no effect on the development of *cedera* and *lidar*. From the late thirteenth century, Latin would have had the same kind of influence on Spanish as any other non-native language.

Morphological Archaism

Morphological considerations can also retard or channel a phonetic change (e.g. in Malkiel 1968). The eventual decision mentioned above not to lose [-e] in Medieval Castilian finite verbs, for example, is perhaps attributable to the greater morphological usefulness of [-e] as a marker than -ø, given the syllabic nature of all other Spanish verb endings. A further case concerns the Spanish decision, which remained widely valid till recently, not to lose [-s], which was needed both to distinguish the familiar from the formal second person singular (e.g. *tú cantas* < *CANTAS*, *Vd. canta* < *CANTAT*) and to preserve the marking of nominal plurals. Once the Latin accusative form was all that survived in Spain, it became – presumably by accident – normal for [-s] to be the sole distinguishing feature of plural forms (e.g. *LUPUM* > *lobo*, *LUPOS* > *lobos*; *MENSAM* > *mesa*, *MENSAS* > *mesas*; *MONTEM* > *monte*, *MONTES* > *montes*). Whereupon some originally singular forms with an [-s] were reanalysed as plural and acquired a singular without an [-s] (e.g. *TEMPUS* > *tiempo*; *tiempo*). It is possible that [-s] was beginning to disappear in Spain (as it did in Italy); in Italy the nominative form of nouns was usually the one to survive, and second-person verb forms developed other distinctive marks (e.g. -i in *cantat*), so [-s] had no such status, but in Spain it became an essential morphological marker, and survived in verbs and nouns where it was needed. We may be entitled to envisage here a period of free variation in Spain between [-s] and either [-h] or ø, which ended in most cases in a victory for the [-s], which was the older form in most but not all relevant words. The [-s] > ø development, however, was merely postponed; in Andalusia and much of América the [-s] has since either aspirated to [-h] or gone entirely. (Felix's analysis (1979) of the nature of this development in the Canary Islands is illuminating on the interaction of the phonetic, morphological and lexical forces.)

We have here a case in which the second-person forms of all verbs and the plural forms of all nouns and adjectives have refrained from undergoing a phonetic change that otherwise, judging from Punic Romance evidence, might perhaps have been expected to occur (cp. the

adverb *fuera* < *FORAS*). Common sense has prevented scholars from suggesting that only the educated used plurals or second-person verbs; but in any event, morphological forms are not thought to fall under the "cultismos" category in the same way as other sounds.

The dissociation of morphology from phonetics is sensible in synchronic analysis; even so, it might seem reasonable to expect the analysis of morphological evolution to have some features analogous to the analysis of phonetic evolution. In practice, however, the neogrammarians never expected that regularities of morphological change would "operate without exception" in the same way as they seem to have instinctively expected sound laws to apply. Except for the fact that morphemes have meaning and sounds do not, there is no *a priori* reason why the two should be expected not to operate in comparable ways; in both phonology and inflectional morphology there is an inventory of bound forms, individually combined in different words, with the items on that inventory being considered as the units of change, rather than each separate form in which the item can occur. We say that [-k] > [-y], for example, rather than enumerating the catalogue of *LACUM* > *lago*, *PLICANT* > *llegan*, *AQUAM* > *agua*, etc.; we say that the Old Latin futures (-BO, -AM, etc.) were replaced with a compound of the infinitive plus *HABEO* (-é, -d etc.), rather than saying that *IBIS* was replaced with *irás*, *LEGET* by *leerá*, etc. And yet when cases occur of statistical regularity in morphological change, comparable to the statistically regular sound laws mentioned above, the forms which survive unchanged have never been ascribed to the influence of "Latin"-speakers.

For example, it is a regularity of the development of Latin to Spanish that many of the verbs that originally had strong perfect forms abandon those forms in favour of new analogous ones created in the usual manner by adding the regular past endings to the verbal stem; e.g. *CURRERE* > *correr*, but *CUCURRI* > ø, and the formation of *corri*; *CRESCERE* > *crecer*, *CREVI* > ø, *creci*; *RIDERE* > *reír*, *RISI* > ø, *reí*, etc. But several of these verbs have remained irregular in the past tense; e.g. *CONDUCTERE* > *conducir*, *CONDUXI* > *conduje*; *FACERE* > *hacer*, *FECI* > *hice*, etc.

These latter strong perfect forms belong to paradigms that are no longer productive, and yet have managed to avoid entirely the opprobrium that has been cast on such forms as *pensar*, *mundo* and *cuaderno*. However, there seems to have been no consistent phonetic or semantic rationale for the choice of which perfect forms to reorganize. For example, Latin *VDXI* (from *VIVERE*, "live") survived

into Old Spanish *viñque*, but from the fourteenth century a synchronically regular form *vivi* seems to have been commoner, and *vivi* is the only modern survivor. *DIXI* (from *DICERE*, "say"), on the other hand, developed with phonetic regularity to *dixe*. Both are very common ordinary words. I know of no scholar who has even considered calling *dixe* and the other strong perfects "cultismos", despite their being exceptions to the perceptible vernacular regularity. This reluctance is understandable, for nearly all the surviving strong perfects belong to common verbs that were clearly in vernacular all the time.

It is likely that there is a process of lexical diffusion here too, in which the timelag from head to tail of the queue is abnormally long. Spanish verbs are still in the process of reallocating their pasts; *ardar*, for example, is said in the grammars to have the past form *ardé*, but the form *ardé* is often the one used now. English has also been undergoing this development for several centuries; it looks, for instance, as though *dreamed* is currently overtaking *dreamt*. Vincent (1980) has discussed the fascinating (but probably unworkable) idea that morphological forms which resist changes may be just the ones that occur in those verbs whose irregular forms have to be individually specified in the lexicon rather than being derived by productive inflexional rules, with the consequence that they can remain stable even when the rule changes; in which case the change in e.g. *creci* is the lexical loss of *CREVI* (as part of the entry for *CRESCERE* — *crecer*) rather than a specifically morphological change. The recalcitrant forms could well be those which occur very often precisely because they are encountered early by the child as he learns to speak, and are thus stored lexically "before there is a sufficient range of data to permit the extracation of generalizations"; whether or not Vincent is right, there is no arguing away the fact that the exceptions to morphological developments often belong to words that cannot possibly be said to be exclusive to literate levels of speech. If, even so, we wished to attempt to explain the retention of the irregular *dixe* as the result of "Latin" influence, we would be faced with a further awkward fact: that Latin words indubitably borrowed from Latin into Spanish in later centuries when the distinction between the two is indisputable, and borrowed by speakers who are by definition Latinate, are taken without the strong perfect forms they had in the original Latin and are given analogous weak ones instead. Latin *PRÆSUMERE* had a past *PRÆSUMISTI*; when *presumir* was borrowed in the fifteenth century, it was allotted the "regular" past *presumí*. Latin *PERSUADERE* had a

past *PERSUASI*; the fifteenth-century borrowing *persuadir* has a past *persuadi*. Latin *CEDERE* had a past *CESSI*; the sixteenth-century borrowing *ceder* has a past *cedí*. And so on.

The fact that discernibly retarded processes of morphological evolution are analysable along lines quite different from the way in which equivalent processes of phonetic evolution have been traditionally analysed cannot be explained away by the generalization that morphological developments are by their very nature prone not to be all-inclusive. For the generalization is not true. There are no pockets at all in Romance, neither lexical nor geographical, of, for example, the retention of a morphologically distinct paradigm of neuter nouns, of synthetic passives, or of ablative cases. The loss of a complete gender, voice or case might be thought to be a sufficiently radical revision of a linguistic system to encourage recalcitrance among some of the affected vocabulary, but it has not. These laws have indeed "operated without exception", to a greater extent than many sound-laws, which were reputed to be more remorseless.

Morphological developments have not normally been treated with the same puritanical rigour as phonetic developments. The parallel cannot be wholly irrelevant, however. The fact that recourse to hypothetical Latinizing dialects in order to account for retarded or evaded change has, to judge from the evidence, little plausibility within the history of Romance morphology is enough to cast doubt on the wisdom of such a recourse within phonetics. Language is divisible into units for analytical purposes, but even so it would be reasonable to hope that generalities applicable to one variety of change might refrain from being incompatible with those applied to an allied variety.

The fact that, in early medieval Romance Europe the educated cannot easily be envisaged as having used archaic morphology in speech, while the uneducated evolved, lies oddly with the treatment of morphology in the grammatical treatises of the same period. The grammars used and compiled before the Carolingian Renaissance have little to say on pronunciation, as will be shown in the next chapter, and could not have been used as a serious teaching aid to reinforce Latinate pronunciation norms. On the other hand they do lay an inordinate amount of stress on morphological forms. The grammatical patterns of Latin verbs and nouns are reproduced in considerable detail; they take up, for example, almost the entire subject matter of Donatus's standard manual, the *Ars Minor* (Keil 1855-80: IV 355-66). There we find neuter nouns, nominal case systems, synthetic passives and synthetic futures, all lovingly elaborated in full technicolor. Accordingly, there

can be no doubt that the educated had indeed encountered old-fashioned word-endings. In order to write at all acceptably the aspirant scribe needed to master a whole array of such endings. It is thus initially rational to hypothesize that the old synthetic future and passive forms of verbs might have existed in educated speech, since the educated undoubtedly had to use them in writing. And yet there are no surviving synthetic futures at all in Spanish. All Spanish futures come from infinitive plus HABEO. There are no surviving synthetic passives in Spanish. All Spanish passive forms come from "to be" plus the past participle. Outside pronouns, there are no morphological inflexions at all in Spanish nouns other than the plural-marking [-s]. In short, the whole old-fashioned morphological apparatus that the educated encountered inevitably when learning to read and write seems to have had absolutely no effect whatsoever on the development of the vernacular. This makes it all the more surprising that the old-fashioned phonology, which the educated need not necessarily have ever encountered, should ever have been assumed to have had a serious impact on the development of the vernacular.

It is more plausible to propose that noone, however educated, actively used the old morphology in their speech. The morphological endings preserved on paper (e.g. -TUR, -ABIS, -IBUS) would have formed a subsection of the passive vocabulary of those who could read and write, as comparable English forms such as *etia* do now. Sabatini's analysis (1965: 994) of legal documents compiled in Early Romance Europe has shown that outside the formulaic sections of those documents the case system tends to break down, with the cases being used as they were in the local vernacular. That is, Italian scribes use the original nominative form, scribes in Spain use the original accusative form, scribes in France usually use only the nominative and their "oblique" case, whatever the syntax. For more literary genres, more "correct" morphology was required, and apparently taught. The practical question in writing the old morphology may well have involved such matters as when to insert a silent 'b' into plural nouns such as *MENSEBUS* (i.e. into Spanish vernacular *meses* < *MENSES*), comparable to the problem of whether or not to insert a silent 'n' into the same word. Occasional written hypercorrections including an extra 'n' before 's' (e.g. *occassionibus*, *Hercules, ariens*), or an extra 'b' into ablatives of nouns from other declensions (e.g. *annibus, dorsibus*), suggest that this was just one more problem encountered in trying to use the old writing system. The problem of what to write in a future tense, when vernacular used HABEO after the infinitive (e.g. *estará* <

STARE HABET) rather than the earlier analytic form (in this case, STABIT), was no less possible to resolve in practice; omitting writing the spoken [r] in [estará], STABIT, is no less learnable than the comparable routine needed to ignore the spoken [e] at the start and to insert a "silent" written 't' at the end⁸. It is child's-play compared with the modern French child's need to learn to write [swet] as *souvent*, or the modern English child's need to learn to write [wenzdɪ] as *Wednesday*. Morphological developments made spelling harder than it had been in earlier centuries, but it did not make it impossible, and those who could write need never necessarily have used the archaic morphology in their speech. Indeed, Politzer (1961) even argued that the more accurately suffixes were transcribed, the less likely they are to have existed in speech.

Archaic morphology is accessible to the literate from written works. Archaic phonology is not similarly accessible; it cannot be taken from books; if it could, historical phonology would be simple. As regards Romance, if the hypothesis of a Latinate speech level were true, morphology might have been expected to play an important part in the influence of that speech level on subsequent vernacular; it plays no part at all. This casts considerable doubt on the existence of such a level, and since within historical linguistic theory there is no great problem in envisaging its absence, we can now be prepared to examine the textual evidence to see if it is compatible with the proposal that Old Romance was the only speech used in Old Romance communities.

Conclusion

Irregularity is regular in phonetic change; cp. Wang (1969: 21): A close scrutiny reveals that even the best phonetic laws are frequently riddled with irregularities.

Samuels (1972: 125) may have been led by Wang's article to remark:

In view of such possibilities, it would seem that completely regular sound-changes can be expected only as rarities.

Concerning Romance, Malkiel has often pointed this out (e.g. 1963-64: 148):

Instead of further laboring the infelicitously posed issue of the regularity of sound change one may candidly admit that, within each set of diachronic sound correspondences, some (whatever the reason for this peculiarity) allow a distinctly higher degree of predictability than do others.

8. Perhaps popular etymology considered -ABIT and the identically pronounced HABET to be the same in any case.

Those irregular forms thought to have proved the existence of a Latin pronunciation system distinct from Old Romance, before the implementation of the Carolingian reforms into any community, can be explained in a reasonable manner in other ways. Lexical diffusion, as the simple statement that some words change later than others, is widely accepted; reasons for some words being late in this list are not difficult to find; indeed, a common view of non-Romance specialists is that "items . . . of high incidence of use are among the most resistant to regular sound-change" (Collinge 1978: 71), rather than that such resistant items should be attributed to esoteric minorities.

The idea that any group of people could resist phonetic changes that were generally affecting the rest of the community, for a thousand years or more, is difficult to take seriously. It is a view that has never been suggested in any other context, and would probably strike any historical linguist from another field as absurd, even if he or she were well acquainted with the manner in which in Romance Philology interim hypotheses regularly fossilize into accepted wisdom. Nor would historical linguists see much to be said for the traditional Romance assumptions that people who spell well necessarily speak in an archaic way, or that their "pressure" can hold back a sound change. This is supposed to have happened in the Early Middle Ages, with no mass communications and a very small literate community.

Given the sociolinguistic and stylistic variation to be found in any community, "Proto"-Romance was the speech of all; it is unnecessary to postulate anything else. From the point of view of theoretical historical linguistics it is reasonable to suggest that "Latin" as distinct from vernacular did not exist in Romance communities before the Carolingian reforms. After these reforms, "Latin" influence can on occasion be justifiably adduced to explain vernacular developments, but previously existing phenomena can and should be analysed with the usual tools of historical linguistic theory. If the textual evidence analysed in Chapter 2, from grammarians, scholars, lawyers, poets and liturgists, can be seen to be compatible with this view, there can no longer be justification for believing in pre-Carolingian "Latin" speech. It was once thought that the sun revolves round the earth. Pliny asserted that when their tusks fall off, elephants bury them underground (*Natural History*: VIII 4). Western civilization has acquired the telescope, a knowledge of elephants, and a theoretical understanding of linguistic evolution; primitive misconceptions formed in the absence of such modern aids can be discarded.

2

PRONUNCIATION IN PRE-CAROLINGIAN ROMANCE COMMUNITIES: THE TEXTUAL EVIDENCE

Latin and the Romance language of any community are nowadays regarded as separate languages. As we saw in the first chapter, it has been normal for this division to be postulated as existing as far back as the final stage of the Roman Empire. One result of this has been that many specialists in the early Romance languages have felt instinctively that the study of "Late Latin" is outside their brief, and many specialists in Medieval Latin have felt that "Early Romance" is outside theirs; thus the assumption that they are separate fields of study has been reinforced. The existence of two distinguishable languages in the period between the fall of the Roman Empire and the twelfth-century Renaissance is usually taken for granted to the extent that scholars of Romance confidently talk of the influence of Latin on the vernacular, and scholars of Latin can talk of the influence of the vernacular on Latin, as if the existence of two names were a guarantee that there are two separate entities. But, as we have seen in the first chapter, the evidence available to the historical linguist is such that there is no longer any need to assume *a priori* that the two were separate or separable before the arrival of the Carolingian reforms into any community. The reforms themselves will be examined in the next chapter. In this chapter the textual evidence from earlier centuries will be examined to see if it is compatible with the theory that there is no general distinction between two co-existing languages at that time, and that the linguistic state of the Romance communities was approximately that prevailing in Modern France or England: i.e. that there may on occasions be visible stylistic variation in syntax and vocabulary between the usage of those who can write and the usage of those who cannot, but there is no general phonological variation correlated with the ability to write.

There exist further differences of attitude concerning the nature of "Latin" and "Romance" in the Early Middle Ages. Some philologists and most linguists appreciate that in any community vernacular speech evolves long before signs of these changes are reflected in the techniques of writing taught to professional scribes; in this case, that Latin vernacular was always evolving, like every other vernacular, and the majority of the orthographic changes visible in early Romance texts reflect phonetic changes undergone in the distant past. Most of them, however, accept at the same time the theory that "Latin" also existed in some circles for some purposes in the Early Middle Ages as we know it did in the Late Middle Ages, spoken, if only by a tiny minority, more or less as it had been in Classical times.

Latinists, on the other hand, and a few philologists also, have often not realized how early and how widespread the evolution was, and tend to operate as if there was very little change in speech from the first century to the seventh. Muller, for example, believed he had shown that the spoken language of the time was close to the written version. Many medieval historians follow the Latinists, understandably assuming that Latinists know best (e.g. McKitterick 1977: 186-87). These arguments need to be considered before the texts.

Unity and Evolution

The history of the sixth and seventh centuries shows that people from the separate Romance communities were able to communicate with each other. This has led to the idea that there must necessarily have been some kind of "lingua franca", of a nature at least partly archaic, as a standard available for such communication between members of the different early Romance areas to be possible (e.g. Burger 1943). The assumption that such a common language had necessarily to exist is possible to refute if we consider the state of Modern English. The English vernaculars of Somerset, Shropshire, Wiltshire, Devon, Cornwall, etc., pronounce words in recognizably different manners, even though schoolmasters in these communities still teach the same spelling norms for most of those words. Given practice, at least, a Scot and a Pakistani, etc., can communicate without difficulty; and even if they could not, their problems would be aggravated rather than alleviated if they attempted to use some pronunciation based on treating traditional orthography as a phonetic script, in the manner of post-Carolingian liturgical Latin. Even if these varieties of English are due to become mutually unintelligible in a millennium's time, as Early Romance has since become Portuguese, Tuscan, etc.,

such mutual unintelligibility is not the norm now despite the dissimilarity between the phonology of Somerset and Jamaican English. Mutual intelligibility does not of itself imply similarity, nor lack of evolution. St Leander of Seville and Latin speakers from Byzantium may have sounded odd to each other, but they seem to have succeeded in communicating.

Non-differentiation

A slightly different view has in the past been influential; the hypothesis of the "non-differentiation of the Western *kouff*". The proponents of this view agree with the consensus of many medieval Latinists that Western Latin hardly evolved at all until the seventh century, and the dialectal diversification in separate areas was only just beginning in the ninth. This theory arose from Muller's study of written texts of the time (Muller and Taylor 1932; Muller 1945) and came to be endorsed by E. Löfstedt (1959: 2-4) and others. The arguments were originally concerned with morphology and syntax rather than pronunciation, but some scholars have taken Muller to have demonstrated that the pronunciation was probably not evolved much until the eighth century. To a historical linguist such an idea seems absurd, so some have attempted to rephrase these views into something more acceptable. For example, Gaeng (1968: 296) has reinterpreted Muller as follows:

The question arises in our mind as to whether Muller indeed meant an *absolute* unity of speech throughout the Western Roman world, with practically non-existent dialectal variations, as his critics seem to think. Muller may have had in mind the kind of *relative* unity of speech that has been attributed to American English.

Whatever Muller may have "had in mind", this relative unity is a useful idea; if we accept that varying dialects can be mutually intelligible, we can simultaneously suggest that a Leonese might have understood the Strasbourg Oaths in French (842), particularly if he cottoned on to the apocope of final vowels, and that neither speaker would have needed to archaize for communication to be possible. Indeed, even in the twelfth and thirteenth centuries, the literature of the time suggests that many people found no great problem in understanding speakers of other Romance languages (e.g. in macaronic verse); the fact that at this later time reformed vernacular spellings were in use can hide from us phonetic similarities, just as much as eighth-century orthographical uniformity can conceal phonetic dissimilarities.

It was not immediately clear whether Muller believed in two coexisting separate pronunciation norms or not, although he does

contrast "Vulgar" with "Ecclesiastical" Latin. His main thesis was that until the late eighth century Latin had in general changed so little that "the living language of Christianization, of practical civilization, was probably similar enough everywhere to retard dialectalization" (Müller and Taylor 1932: 24). This undifferentiated *koine*, however, is similar to the kind of Latin envisaged within the two-norm theory discussed in the previous chapter (e.g. by Rice 1909 and Pope 1934); it seems that the one language that Müller was envisaging was not so much the Romance vernacular part of the vernacular/Latin pair as usually envisaged, but the "Latin" part. He refers to seventh- and eighth-century documents being written by people who "wrote the language they spoke and heard" for example (Müller and Taylor 1932: v). The one language that I argue for here is the Romance part of that dichotomy.

The view that "Romance" only evolves after the seventh century has been untenable since the discovery of the Pompeii graffiti (datable to A.D. 79), which attest many of the changes postulated as not existing before then. Analysis of the graffiti, and other recent analyses of informal writing such as Adams' (1977a) discussion of scribal habits in the letters of Claudius Terentianus (early second century), make it clear that several of the important differences between Romance and "Classical" Latin had at least begun by then. No linguist can seriously doubt any more that the vernacular was changing long before 600 A.D. The only debatable point concerns the postulated co-existence alongside that vernacular of an unchanged spoken "Latin".

Proto-Romance Reconstruction

The theory that dialectal divergence, as well as evolution, began very early, has in the past also been influential. The "substratum" theory held that Latin in each area of the Roman world was profoundly affected by the preexisting languages spoken in that area by the pre-Roman inhabitants, with the result that speech differentiation began with colonization (Mohl 1899; for a sophisticated modern version of the general theory, see Whinnom 1980). Within Romance this theory is no longer held in much regard, since no phonological development of any consequence has ever been shown to have undeniable origins in a substratum language, and most of the linguistic features that were originally candidates for such influence have since been found to occur outside the zone of the language in question. The early chronology remains respectable, however. There are those who have "reconstructed" Proto-Romance by arguing backwards from the evidence of later attested Romance forms. R.A. Hall Jr (1950, 1976), for example,

regards deductions based on the evidence of later vernaculars as more likely to reflect the speech of the early times than deductions made on the evidence of pre-Carolingian spelling, and has produced traces of diverging dialects that apparently date back to before Christ. This seems to imply that the speech of the Empire was already in several distinguishable vernacular divisions long before the period of what Müller and many others regard as a fairly uniform speech. In Hall's view Proto-Romance was only uniform for a short while, if at all, and contemporary with the related but largely artificial language called "Classical Latin". This view is sometimes interpreted as suggesting a coexistence of separate spoken systems (Classical and Vulgar Latin) from 200 B.C. onwards (e.g. Maurer 1962, Leonard 1980), but the evidence that is supposed to support this duality is easily compatible with the normal phenomena of social and geographical differentiation to be found in any one language used in a complex society.

Thus it is that in the Proto-Romance "reconstruction" theories the early chronology for dialectalization remains current; an early date for several sound changes is certain, and whether or not we believe distinguishable dialects from such an early period to be identifiable depends essentially on our definition of a dialect. As long as we do not imply any mutual unintelligibility, and we realize that any suggestion of strict geographical delimitations of such dialects would be an anachronism (conditioned by the modern tendency for isoglosses to coincide as a result of the politics of national education), we can accept much of Hall's theory. His views would probably be seen generally as extreme, however, partly because he seems to assume that language change is neat, linear and unidirectional, when it seems from modern evidence that it tends to be haphazard, non-teleological and multidirectional, and it is notoriously difficult to locate early medieval texts on the basis of their "vulgar" phenomenon, and partly because he proclaims that recourse to documentary evidence is not only pointless but "unscientific". What is unsatisfactory is to see documentary material as face-value evidence of speech; if we can accept the existence of old Romance vernaculars approximately as the reconstructionists have envisaged them, examination of documents can be a rational and indeed illuminating exercise. Even so, Hall's instinctive mistrust of written texts as evidence for contemporary pronunciation has a sound basis. This mistrust has also been well expressed by Pulgram (1950), who pointed out that nobody writes unless they have taken time and trouble to *learn* to write; that learning to write is a practical discipline, consciously taught in accordance with established norms; and that until the

Ask:
WHY DOES
HALL DISTURB
TEXTS?

norms are consciously reformed they remain the same. The scribe is likely to reflect linguistic evolution in semantics and syntax, particularly in word order, but is unlikely to reproduce changes in pronunciation. Whoever he is and wherever he is writing, he will not produce simple phonetic script:

He cannot help it, for he has learned to write in the literary language only, according to certain rules of spelling and grammar. What he produces may not always conform to those rules, but neither will it be a faithful rendering of what he would SAY. (459)

A comparison with Modern English will help to make this point clear. For example, virtually all written English documents of the twentieth century that include the words *orange* or *knight* have them spelt "correctly". A thirteenth-century Muller might deduce that speakers of the twentieth century must have said [ʔrændʒ], or [ʔrændʒə], and [knaɪt] or [knaɪt]; but if the evidence of linguistic reconstruction is going to suggest that we probably now say [ʔrɪndʒ] and [naɪt], specialists will be right to accept the reconstructed forms in preference to the apparent testimony of the texts. In the same way, later evidence (the written form *siglo*) suggests that SAECULUM in a written Spanish text of the ninth century is more likely to represent a vernacular [sjɛlo] than [sekulum]. Noone wrote *siglo* in the ninth century; noone writes *orinda* or *naɪt* nowadays; because noone is taught to write without also being taught how to spell. Phonetic (or phonemic) script does not come automatically to people in the way speech does. Indeed, there is every reason to suppose that this analogy underestimates the extent to which written pre-Carolingian "Latin" falls short of being a phonetic script; any manuscript thought worth preserving would probably have enough care expended on it to approximate to correct old-fashioned orthography, and many were in any event re-copied "correctly" in the Carolingian period (as happened, for example, to the Benedictine Rule: see Mohrmann 1952).

Phonetic Script and Education

Anyone who has attempted to teach phonetic script will be aware that even the best student does not find it natural, and has to be taught how to do it. Those who have already been taught how to write in the "correct" way, as everyone has who has been taught to write at all, is likely to find the traditional spelling of a word springing more easily to his ballpoint than the phonetic transcription of the same word. Yet discussion of both Medieval "Latin" and early Romance texts tends

to assume that writing an unfamiliar phonetic script is as natural to the human being as speech. On the contrary: in its initial stages, the operation of a novel spelling technique is a complex and demanding task. For example, early written French vernacular texts show evidence of diphthongization (e.g. *ciel*, *buona*, *belezour* in the *Canitène de Sainte Eulalie* of c.880); we can assume that these diphthongs correspond to speech habits of that time; the *Appendix Probi* of the seventh century (Robson 1963) shows no such forms, but this cannot be seen as evidence that diphthongization is a post-*Appendix Probi* evolution. It may have been only allophonic, but it certainly existed.

This situation has analogies in Modern English; Americans write *United States* as the British do, despite it being standard American for the first *t* to represent a spoken [d]. Thus the first *r* and the *d* in this phrase represent similar sounds, which are different from that represented by the two *ts* in *States* [steɪts]; similarly a ninth-century Leonese would write TOTUM representing a spoken [toto]. Even new vocabulary can be given old spelling. Neologisms in the U.S. ending in -ator or -aring, etc. (e.g. *deescalating*) are given a written *t* and a spoken [to] on the analogy of all the other comparable cases. The late eighth-century Northern French Glosses of Reichenau include items of "new" vocabulary (such as SPARNIAVIT, i.e. Fr. *épargner*, of Germanic origin), which are given unreformed Latin spelling even though they did not exist in Imperial times when that spelling corresponded more closely to pronunciation (Elcock 1975: 324-29).

Thus those who assume that the survival of old orthography implies the survival of old speech habits are in danger of appearing to underestimate the difference there can be between speech and spelling. What happens in a case such as Early Romance, as the two diverge with time, is that the rules for writing become more and more involved, including not only "silent" letters and syllables but even "silent" morphological forms. This is a reasonable explanation for the enormous stress on noun and verb endings in all early Medieval writing manuals (*Aries Grammaticae*). The same occurs in Modern France, where children can say [ʃi:t] indiscriminately for four grammatical persons, and are then bewildered at having to learn to write, and distinguish between, *chante*, *chantes* and *chantent*; for the many French children lacking the [e]/[s] contrast, [ʃi:te] can be spelt *chanter*, *chantez*, *chanté*, *chantée*, *chantés*, *chantées*, *chantai*, *chantais*, *chantait* or *chantaient*. It is hardly surprising that learning to write in the Early Middle Ages took so long; and it is logical to conclude that relative correctness of written Latin in pre-Reform Europe was as much a question of

"Correct Latin" = result of what? in EMS?

ASK: why are words like *ciel*, *buona*, *belezour* in the *Canitène de Sainte Eulalie* of c.880? For any of forms of "correct" pronunciation?

the effectiveness of education as it was of relatively evolved vernaculars. Scholars have talked of the "pure Latin" of seventh-century documents from Spain as if this were evidence of a greatly retarded process of linguistic evolution in the peninsula, rather than merely being the visible consequence of the higher level of education in Spain than in the rest of the Romance world in the seventh century.

As regards the emergence of new orthographies, general spelling reform can only occur as a positive act with a conscious purpose. It does not happen on its own, as pronunciation change does, and the social instability of the Early Middle Ages would not in practice have been conducive to the success of such a move, as Battisti (1960) has observed. The Carolingian reformers initially saw their linguistic task as that of re-establishing old norms; the subsequent emergence and invention of written Old French may well have been originally unforeseen.

If we are able to regard the emergence of vernacular orthography from the ninth century onwards as a spelling reform that recognizes changes that had long since occurred in everyone's speech, most of the problems that worried Muller resolve themselves; the phenomena attested in earlier texts can be seen not as questions concerning the coexistence of two separate dialects (the two-norm theory) but as questions concerning scribal practices and the nature of education in communities where the writing norms did not correspond closely to anyone's normal speech.

"Vulgar" Latin

The much-used term "Vulgar Latin" is multivalent and best avoided. Lloyd (1979) has discovered thirteen different meanings for the phrase. In this book, *Romance* is used to refer to the vernacular of any time after the end of the Western Roman Empire; *Imperial Latin* is used to refer to any variety of Latin before that time; *Latin* is used by the educated and is systematically and archaically distinct from contemporary Romance. This present book is examining the theory that this Latin is a Carolingian invention, with no previous existence, and no direct continuity with Imperial Latin. The concomitant suggestion of a one-norm Proto-Romance hypothesis is not intended to imply that there was any particular similarity between the norms of different communities.

The phrase *sermo vulgaris*, and assorted cognate expressions, as Díaz y Díaz pointed out long ago (1951-52), seem in pre-Carolingian communities to be applicable to the speech of everybody, not just of

the uneducated. The tendency to identify words of the *vulgus* with Proto-Romance, tout court, is untenable. Several of the words qualified as *vulgo* in glossaries have no direct Romance descendants; among others Díaz y Díaz mentions *cacida*, *caefio*, *conditum*, *fiscia*, *tereolo*, *verbosus*, *milus*. In this study Díaz y Díaz also performed the simple but necessary task of clarifying a remark made by Servius in the fifth century which was supposed in the nineteenth century to be a confirmation of the separate existence of a "Vulgar Latin". Commenting on Vergil's *Georgics* 3.147-8 (... *cui nomen asilo/Romulum est, oestrum Grai vertere vocantes*), Servius said "latine asilus, vulgus tabanus vocatur." Servius used *latine* where Vergil used *Romulum*; as Díaz y Díaz pointed out, *latine* here is used by Servius to refer to the local speech of Latium (and thus of Vergil), and *tabanus* is in fact the general word for a horsefly (as in Sp. *tabano*), being used by Varro and Pliny, among others. *Vulgo* refers here to the speech of all, including Servius.

In this case the anachronistic class sensitivity of the nineteenth-century German Schuchardt obliterated the geographical knowledge of the early fifth-century North African Servius; Schuchardt largely built the theory of a separate "Vulgarlatins" on such remarks, being among the first to use the phrase and responsible for rigidifying the concept (1866-68). As regards vocabulary, at least, the idea of a "vulgar" level is earlier than Schuchardt (see Coseriu 1977); the examination of the views of Isidore (later in this chapter) suggests that non-imperial vocabulary is all that he means to refer to with the word *vulgar*. Mohrmann (1961a) considered the theory from a Latinist's viewpoint with no hint of a belief in separate varieties of pronunciation. Geographical and social differentiation of vocabulary is normal in any complex language community, and has no implications in itself for pronunciation.

Menéndez Pidal (1926) saw spoken "Vulgar Latin" as still existing in tenth- and eleventh-century León, specifically distinguishable from both Romance and Latin, being the direct descendant of the speech of the late empire and thus opposable to Classical Latin as well as tenth-century Leonese. This view is still standard, but hardly defensible (see below, Chapter 4). Lloyd's conclusion (119) concerning the term "Vulgar Latin" is that "it is, in fact, a carry-over from the pre-scientific view of language, based on a tradition reflecting a lack of comprehension of the nature of linguistic evolution and of the complexity of social and stylistic stratification." He is right. The similar ambiguity of the term "Medieval Latin" may soon seem equally open to criticism, as Díaz y Díaz (1960), Bastarás (1960) and others have hinted. Pre-reform Medieval Latin was a written form of the vernacular,

different in kind from post-reform Medieval Latin, which was in theory a separate language. The phrase "Vulgar Latin", however, deserves to be banished at once from serious scholarly use, as have been *philologisti*, *humours*, and *the music of the spheres*.

The Evidence of Grammarians before 500

The pronunciation of Imperial Latin has been described by Allen (1970) and Bassols (1976). Allen makes it clear that speech did not correspond particularly closely to spelling, although even he may have underestimated the lack of congruence between post-medieval pronunciations of Latin and that of the Roman Empire. Kramer (1976) has produced a collection of quotations from grammarians concerning pronunciation, in which he apparently subscribes to the view that "Vulgar Latin" was largely immutable from Varro to Alcuin. What the contemporary grammarians say on this subject has not been seriously studied by philologists; the general assumption is that grammarians naturally recommended classical usage and opposed any phonetically evolved forms; although Janson (1979) points out that this is not so, he refrains from further discussion. The assumption is unjustified. A study of the grammarians in Keil's *Grammatici Latini* reveals that they are not much concerned with pronunciation at all. When they mention it their remarks are seldom easy to interpret precisely, mainly because written letters and spoken sounds are not theoretically distinguished, but some of them say enough to make it clear that they and their public use the vernacular of their time and place. If a dichotomy between "Latin" and (Proto-) "Romance" actually existed in these centuries (from the second to the fifth), the grammarians would have done what it has been customary for grammarians of the last thousand years to do, that is, they would have prescribed that "correct" pronunciation of Latin was to be that which corresponds sound for letter to the correct Latin orthography; they would have stigmatized and attacked any "vulgar", vernacular or evolved features. "Correctness", as advocated by these grammarians, would, if the traditional view were well-founded, combine a reverence for archaism with the prescription that the traditional spelling is to be used as the guide for speech.

The grammarians do not take this line. On the contrary, a reverence for archaic usage is no part of their approach. Aulus Gellius, for example, in the second century A.D., says with regard to the *h*: "inserabant eam veteres nostri ... nulla ratio visa est" (II 3; Kramer 1976: 48). Diomedes, in the fourth century, declared that "quod vulgo obsepio dicimus, veteres obsepio dixerunt" (Keil I 383.10-11; Kramer

1976: 26), allying himself roundly with the *vulgus*, by using the first person form *dicimus*, rather than with the *veteres*. As regards the "silent" letters such as *h*, *m*, and the *n* in *ns*, it is noticeable that although it was thought advisable to write them if that was the established practice, there was no recommendation that they should be pronounced.

a) The Second Century

Velius Longus, in the second century, is particularly clear concerning contemporary lack of congruence between speech and spelling. His comments on the intervocalic *h*, for example, show that it is to be written in *vehemens* and *reprehendit* even though "elegantiores et *vehementem* dicant et *reprehendit*, ... *prendo* enim dicimus, non *prehendo*" (Keil VII 68.14-17; Kramer 1976: 54); Velius is happy to speak the evolved form while writing the established orthography. *Elegans* is a word employed by Velius to praise a usage: e.g. "alimenta quoque per *i* elegantius scribemus quam alimenta per *u*" (Keil VII 77.8); elegance is not to be overdone, however — "nimiae rursus elegantiae sectatores non arbitror imitandos" (Keil VII 79.19) — for normal usage is his target. The usage of the *antiqui* is not recommended if it clashes with the normal usage of Velius' day: "*Mium* et *committum* quoque per *i* antiquis relinquamus ... nostris iam auribus placet per *e*, ut et *Mercurius* et *commercium* dicantur" (Keil VII 77.13-15). As regards *maximus* and *maximus*, the former is both *antiquum* and *rusticanum* (Keil VII 49.21), neither of which is a term of approval; the apparent belief of some philologists that the "rustic" was necessarily more evolved than the "urban" is unjustified. *Rusticanum* is also used by Velius to criticize the form *arrubus*: "mihi videtur nimis rusticana enuntiatio futura, si per *u* exulerimus" (Keil VII 68.6-7). For Velius Longus, if people want to write like the *antiqui* they can, but they are not to talk as they write: "concedamus talia nomina per *u* scribere iis qui antiquorum voluntates sequuntur, ne tamen sic enuntient, quo modo scribunt" (Keil VII 50.4-6). Velius is equally precise as regards word-final *m*: before a vowel, at least, it is a silent letter, written not spoken. He makes in this connection the general observation that spelling is *not* exactly a correspondence of the manner in which "we" talk:

Ingredienti mihi rationem scribendi occurrit statim ita quosdam censuisse esse scribendum, ut loquimur et audimus, nam ita sane se habet nonnumquam forma enuntiandi, ut litterae in ipsa descriptione posita non audiantur enuntiatae, sic enim cum dicitur

1. The editors of the Rome edition of 1587 were unable to believe this and omitted the *non*. Suggested translations of the longer extracts from grammarians are included in an appendix (263-70).

illum ego et omnium optimum, illum et omnium aequum in terminat nec tamen in enuntiatione apparet... confiterendum alter scribi, alter enuntiar... (Keil VII 54.1-13; Kramer 1976: 60)

... saepe aliud scribanus, aliud enuntians... (Keil VII 75.15)

Not even Quintilian in the previous century had recommended a consonantal [-m] in such circumstances: his comment "neque enim eximur, sed obscuratur" (9.4.40) suggests that written -m represented to Quintilian a nasalized vowel. No reader or speaker of Latin has followed Velius' prescription in the last thousand years, since the practice of giving a sound to every Latin letter has been routine; but Velius, and second-century vernacular speakers in general, were happy to write -m in some circumstances where they did not pronounce it. Velius prides himself on not confusing *orthographia* with *orthoepeia* (Keil VII 71.8-12); his *orthographia* is specifically not intended to prescribe pronunciations, for both he and his public talk their own language naturally anyway, and writing is the technique that has to be taught and learnt. "Proprium *orthographiae* est, quod in vocis enuntiatione nihil videmus ambiguum, at in scriptione tota haesitatio posita est, ut, cum dico *Troia*, per *i* unum an per *duo* scribere debeam" (Keil VII 72.2-4). Similarly, concerning the *n* of *ns*, Velius believes it is *elegantius* to follow Cicero's (spoken) use of *forensis*, *Megalesis*, *horrensia*, without the [n] (Keil VII 78.21-79.5; Kramer 1976: 64). Quintilian had said that "consides exempla *n* littera legimus" (1.7.29), allying himself in the first person with those who do not pronounce the [n] in such a word; Quintilian, Velius, and their public, at least, are not to blame by virtue of their literacy for the [n] in *Fr. pener*, etc. If we still seriously believe that Proto-Romance was the speech of only the uneducated, it is surely strange to see Quintilian reveal himself as "vulgar" in such a matter-of-fact tone. The comments examined here are more compatible with the view that this vernacular pronunciation is in fact the pronunciation of everyone in the first two centuries A.D.; that the educated Latin used by scholars ever since the Carolingian Renaissance was not the same as that used by the literate in the Roman Empire, and may never have been used at all by anyone before the Carolingian scholars invented it. Velius Longus has held the stage in this section because he is untypically explicit, but there is no need to suppose that his usage was untypical of that of his time.

b) *The Fifth Century*

When vernacular loses sounds entirely, such as [h], [-m], and [n] before [s] in Imperial Latin, their substance in the orthography makes the difference between speech and spelling immediately perceptible.

Grammarians are less likely to be aware of other kinds of change in progress, since allophonic variation is notoriously hard for native speakers to perceive. Imperial Latin developed considerably from the second century to the fifth, however, and by the time of the fifth-century grammarians some such developments are beginning to be noticed. These later writers occasionally mention vowel allophony in a way that makes it clear that the old quantitative vowel system was already superseded in their vernacular by that of "Proto-Romance". In Early Latin, there were distinctions of length of pronunciation between a long and short *a* and *e*, *i* and *e*, etc., but the evidence of all the later Romance languages suggests that such distinctions were no longer functional in speech by the fifth century. Consentius, for example (from Southern Gaul), noticed that the written letter *i* represents different sounds in the words *ire*, *habui* (and *tenui*), and *hominem*, and that the distinction is not explicable in terms of old-fashioned quantity alone. His terminology verges on the opaque, but if we consider these words in the light of their Romance descendants — e.g. *IRE* > Sp. *ir*, *HABUI* > Sp. *hube*, *TENUI* > Sp. *tuve*, *HOMINEM* > OSp. *omme* > *hombre* — it seems likely that he is referring to a closed [i] in *ire*, an open [i] = closed [e] in *habui* and *tenui*, and a tendency to a centralized schwa [ə] — as in the second vowel of English *father*) in *hominem*.

Romanae linguae in hoc erit moderatio, ut exilis eius sonus sit, ubi ab ea verbum incipit, ut *ire*, aut pinguior, ubi in ea desinit verbum, ut *habui*, *tenui*; medium quendam sonum inter *e* et *i* habet, ubi in medio sermone est, ut *hominem*. Mini tamen videtur, quando producta est, plenior vel acutior esse; quando autem brevis est, medium sonum exhibere debet, sicut eadem exempla, quae posita sunt, possunt declarare.

(Keil V 394.16-22; Kramer 1976: 26; Allen 1970: 48)

This final comment shows that his readers will see what he means if they pronounce these words for themselves; his readers too, it seems, are not archaizing unevolved "Latin" speakers either. Consentius is proud of the way he trusts his own observation of contemporary usage rather than merely accepting the authority of long-dead writers: "nos exempla huius modi debemus quae in usu loquentium animadvertere possumus" (approvingly quoted by Díaz y Díaz 1951-52: 215, n.4).

In the same century, Pompeius has similar advice for his readers.

When discussing the remarks made in the second century by Terentianus Maurus about "long" and "short" vowels, which would have been intelligible in that distant age when that distinction was probably still perceptible, Pompeius suggests that you can tell whether an *o* is technically

longa or *brevis* by the following practical test:

O longa sit an brevis? si longa est, debet sonus ipse intra palatum sonare, ut si dicas *orator*, quasi intra sonat, intra palatum, si brevis est, debet primis labris sonare, quasi extremis labris, ut puta si dicas *obit*. habes istam regulam expressam in Terentiano: *quando vis exprimere quia brevis est, primis labris sonat; quando exprimis longam, intra palatum sonat* (Keil V 102.13-18; Kramer 1976: 22; Allen 1970: 48)

For Pompeius, *longa* and *brevis* are technical terms unconnected with any distinction between a long and a short vowel. The distinction is instead based on relative manner of articulation, although neither *primis labris* nor *intra palatum* is clear to a modern reader; what was accidental and allophonic to Terentianus seems crucial and criterial to Pompeius. The [w] of the diphthong developed from [ɔ] is noticeably produced with the lips: this may be the meaning of *primis labris*.

According to reconstructed evidence from later Romance, in fifth-century Proto-Romance vernacular the originally short [ɛ] and [ɔ] are diphthongized when stressed, in most geographical areas, although since this is still simple positional allophony no special orthography is adopted to represent this in writing. Writers of modern English never write *pat* as *phat*, even though an allophonic aspiration is present. "Long" [ē] and [ō] do not suffer diphthongization, so a distinction between words containing the two phonemes is still in existence: it is no longer made in speech on the basis of length, but the traditional terminology of *brevis* (or *corruptus*) and *longus* (or *proctus*) remains. Grammarians tend to agree that the two letters *e* and *o* each represent two pronunciations. Modern linguists usually distinguish these as a more closed [ɛ] and [ɔ] opposed to a more open [e] and [o] which had begun to diphthongize, to [je] and to [wo], [we], or [we] respectively. (Closeness and openness refer to the distance between the tongue and the top of the mouth.) The grammarians of the time are unaware of the criterial importance of the relative aperture of the vocal tract, however. Three fifth-century commentators on Donatus, the North Africans Pompeius, Servius and Sergius, proclaim that the two pronunciations are distinguishable in that the open *e* or *o* (called *brevis* or *corruptus*) is pronounced as if it were a diphthong:

vocales sunt quinque. hae non omnes varios habent sonos, sed tantum duae, *e* et *o*. nam quando *e* corruptum est, sic sonat, quasi diphthongus: *equus*; quando productum est, sic sonat, quasi *i*, ut *demens*. similiter et *o*, quando longa est, intra palatum sonat: *Roma*, *orator*; quando brevis est, primis labris exprimitur: *opus*, *rosa*.

(Sergius: Keil IV 520.27-31; Kramer 1976: 22)

vocales sunt quinque, *a*, *e*, *i*, *o*, *u*. ex his duae, *e* et *o*, aliter sonant productae, aliter corruptae. nam *o* productum quando est, ore subulato vox sonat, ut *Roma*; quando corruptum, de labris vox exprimitur, ut *rosa*. item *e* quando productum, vicinum est ad sonum *i* litterae, ut *meta*; quando autem corruptum, vicinum est ad sonum diphthongi, ut *equus*.

(Servius: Keil IV 421.16-21; Kramer 1976: 22)

Allen (1970: 49) believed the diphthong referred to here as existing in the *e* of *equus* is the *ae*; but this is surely unlikely. *ae* has long since stopped representing [ai], as Allen himself shows; the comments of Pompeius (below) make it clear that he positively wishes to distinguish *equus* from *aequus*. In view of the development of Spanish *yegua* < EQUAM there can be little doubt that [je] is the diphthong to which Servius refers (Spanish is generally taken to be the closest surviving relative to the lost Romance of North Africa). *Demens* and *meta* involve the closed [ɛ], which in vernacular was indeed "vicinum" to short [i], as Servius and Sergius say – so much so that in Romance languages original short [i] and long [ē] develop into the same sound. *Opus*, with a short [ɔ], became O.Sp. *huevo*, so the labial diphthong with [w] is likely to be what Sergius is referring to: *rosa*, however, has for some reason not survived into Romance as a diphthongized form.

Pompeius actually offers us a diphthongized spelling. He is unexpectedly explicit on this subject in another context, that of "barbarisms". He says that one of the barbarisms people sometimes commit is to confuse technically *brevis* and *longa* syllables, but the confusion is exemplified in his study neither by mistaken aperture height, nor by length, but by the diphthongization of the wrong cases. Choosing a standard example of closed, originally long, [ɔ], the word *Roma*, Pompeius spells its mispronunciation with *uo*: *Ruoma*. It is ironic that the presence of diphthongization at the time, as firmly postulated by the reconstructors of Proto-Romance vernacular, is guaranteed by this spelling of a mispronunciation of a word which normally does not diphthongize:

Est alter [barbarismus] qui fit in pronuntiatio, plerumque male pronuntiamus et facinus vitium, ut brevis syllaba longo tractu sonet aut iterum longa brevior sonet; sicuti vellet dicere *Ruoma*, aut si vellet dicere *aequus* pro eo quod est *equus*, in pronuntiatione hoc fit.

(Keil V 285.5-9; Kramer 1976: 24)

The point here appears to be that it is *equus* that has the diphthongized pronunciation in normal vernacular, the barbarism being to give a diphthong to the normally monophthongal *aequus*. The barbarism (which often happens even, it seems, in the mouth of Pompeius himself: "male

pronuntiamus") is not that of talking some evolved "vulgar" non-Classical Latin, but that of not talking ordinary acceptable fifth-century southern vernacular. That vernacular would diphthongize OPUS and EQUUS but not ROMA, nor – probably – AEQUUS. (*Aequus* does not survive now, but the evidence of closure in the cognate Sp. *igual* or It. *uguale* < AEQUALEM suggests that this word was closed in its first syllable to at least as great an extent as *demens* or *meta*, and thereby "strengthened" itself out of diphthongization.) Not surprisingly to a linguist, Pompeius and his readers speak their own vernacular. We have cause to be grateful to Pompeius, for on the next page (in Keil) we find this long discussion concerning how to pronounce words spelt with *ri* or *di* and a following vowel (this section is here reproduced from Kramer 1976: 70, whose edition makes the point clearer):

sunt aliqua uita, quae uitare debemus. istas quinque res debes uitare: iotacismos, labdacismos, myotacismos, collisiones, harrus. iotacismi sunt, qui fiunt per *i* litteram, sicut ita dicit, *Tritus* pro eo quod est *Tris*<>*us*, *Auentius* pro eo quod est *Auentis*<>*us*, *Amanius* pro eo quod est *Amanis*<>*us*, quo modo ergo hoc fit uitium? definiamus illud, et uidebimus postea, quo modo cauere debemus, fit hoc uitium, quotiens post *ri* uel *di* syllabam sequitur uocalis, si non sibilus sit, quotiensque enim post *ri* uel *di* syllabam sequitur uocalis, illud *ri* uel *di* in sibilum uerendum est, non debemus dicere ita, quem ad modum scribitur *Tritus*, sed *Tris*<>*us*; media illa syllaba mutatur in sibilum, ergo si uolueris dicere *ri* uel *di*, noli, quem ad modum scribitur, sic proferre, sed sibilu profer. sed illud scire debes, quia tunc hoc facere debes, si media sit, si autem prima fuerit in prima parte orationis, etiamsi sequatur uocalis, non illam uertit in sibilum, ecce *dies* habet post se uocallem; debemus dicere *dies*<>*ies*, sed non dicimus, adde illi partem priorem, ut istam median facias, et dicis *meridies*<>*ies*; non possumus dicere *meridies*, ut scias quoniam tunc contigit hoc huius syllabae, quotiens media est, non, quotiens prior est.

"Non debemus dicere ita, quem ad modum scribitur...," "noli quem ad modum scribitur sic proferre." No instruction could be clearer, do not pronounce the word exactly as it is written. What is here called a "vitium" is precisely the pronunciation that is advocated as correct by the scholars of Carolingian times and later, viz. pronouncing written *ri* and *di* as [ri] and [di], and failing to palatalize intervocalic [i] to [ʃi] and [d] to [dʒ] before the semivowel [j]. The pronunciation advocated by Pompeius is that which still survives in Italian words such as *giustizia*; the pronunciation criticized is that which follows the spelling. The difference mentioned by Pompeius between the palatalizing *meridies*, with a [dʒ], and *dies*, with just a [d], is presumably due to the fact that there is a stressed [j] in *dies*

<> Sp. *dies*) but a semivocalic [j] in *meridies*. The spelling-pronunciation [meridies] without the [z] is something which cannot be said, "non possumus dicere". Pompeius not only *does not* talk in the normal post-Carolingian Latin manner here, he *cannot*. Holtz (1981: 236) comments that in general Pompeius used "African Latin", pointing out that he often used *habere* to form the future tense, in the normal spoken way.

The evidence suggests that the fifth-century grammarians, on the few occasions that they mention pronunciation, discuss, prescribe and use the normal vernacular of their time and place. In previous centuries, Cicero, Quintilian, Caper, Velius Longus and others did the same. They were writing for an audience who already knew how to pronounce and did not need to be taught. The evidence does not support the traditional view that the language of the educated classes had failed to undergo vernacular sound changes; and if the grammarians and their public did not speak such an underdeveloped archaic language, who did? Nor does it offer any comfort to the general view of Medieval Latinists, that the developments from "Classical" Latin to Romance only begin to any extent after 600 A.D. The evidence is not large. Much of it is opaque or ambiguous. But what there is supports the view that the grammarians themselves spoke the vernacular of their time, which had evolved as much, broadly speaking, as the reconstructionist philologists now claim, and more than Latinists and historians tend to believe: in short, that there is only one norm of pronunciation in an area (routine social and stylistic variations included), and that this norm can, if we wish, be called "Proto-Romance".

The Evidence of Legal Documents

The following three extracts from Italian legal documents were reproduced (among many others) in Sabatini (1965: 983-84):

- 1) from Colonia (near Grosseto), 762 A.D.:
... *per omnem annum iustitia ipsei case reddere debeam porco uno valente tremisse uno et uno pullo et quinque ovas et carnisia una valente tremisse uno et uno animale in mense magio valente tremisse uno, vinum et labore secundum consuetudinem ipsius case* ...
- 2) from Lusciano (near Bojsena), 762 A.D.:
... *per omnem annum de ipsa casa vel res reddere debeam uno animale annuino in mense magio, porco uno annuino in octummo, sex decimate de vino, grano siliigne bono modis quatuor, angaria quartas utilitas fuerit* ...
- 3) from Lucca, 765 A.D.:
... *prandium eorum tali sit per omnem septimanam: scaphilio*

grano, pane cocto et duo congia vino, et duo congia de pulmentario, faba et panico mixto, bene spisso et condito de uncto aut de oleo, et nullus de hereditibus nostris . . .

It is time to turn our attention to the manner in which practical literate men put to use their instruction in writing, *Grammatica*. One such group are lawyers, producing workaday documents such as the above. Sabatini has demonstrated that the language of notarial documents, from at least as early as the sixth century, falls into two categories. In the first place, those parts of the text which are off-repeated formulae, copied from a standard version, are generally more "correct". The existence of these sections is not evidence that the scribe naturally produced Latinate forms in his own vernacular; he had learnt to copy such expressions as part of the tools of his trade, as a lawyer might do today. These sections are often longer than the rest, and can thus give an impression that the language of the scribe was more archaic than in fact it was. (These formulae are often unclassical; e.g. above, *ipsei case, de ipsa casa, per omnem septimana*.) Sabatini calls these passages the "parti formulistiche"; they are distinguished from the rest, the "parti libere". These latter "free" sections deal with specific details of the individual transaction, which are intrinsically less amenable to prior formalization, and have in practice to be some kind of representation of the depositor's vernacular. For this purpose scribes worked with a "scripta latina rustica", consciously deployed for a practical end (cp. also Sabatini 1968). The fact that these free sections are not couched in authentic old-fashioned Imperial Latin is not the result of mere scribal ignorance, but of the intentional use of a less formal technique of writing. Under this perspective Merovingian or early Leonese lawyers can retain some professional self-respect, for this was a sensible and rational procedure. The once common patronizing approach to such documents, that "these barbarians couldn't cope", has been slightly refined — E. Lofstedt (1959:3), for example, describes the style of such documents as a "haphazard mixture" — but should be entirely discarded. Lawyers adopted "incorrect" forms for a practical reason.

Modern scholars have at times interpreted the general truth of a low level of education in much of the early medieval Romance areas (e.g. seventh- and eighth-century Northern France, ninth- and tenth-century León) as being a particular truth concerning the intelligence of individuals. Yet there is no reason to postulate a necessary variation in the intelligence of individual scribes proportional to the level of literacy in their community. If a man is writing obviously "incorrect" forms on purpose, rather than just scattering slips of the pen, it is pointless to

criticize him for inaccuracy. The lawyers are operating in an age when the connection between speech and spelling was becoming tenuous, and such practical adjustments are a reasonable compromise between stultifying pedantry and unacceptable imprecision.

The non-imperial innovations in these sections are of all kinds. They include new vocabulary, for example of Germanic origin, or in Spain of Arabic origin also; there is the limitation to the reduced case systems proper to the Romance of specific areas (for example, only the accusative form of plural nouns in Iberian texts of the tenth and eleventh centuries, where Romance nouns preserve only their accusative form); and they include experiments with occasional new graphic representations, such as the use of *ux* for [w], or of *g* for [j].

In the documents reproduced above, italics represent formulae and the rest is *libera*. Sabatini points out that *octurnio* (in no. 2) is a splendid confirmation of the postulated *AUTUMNIUS* form (> Sp. *otoño*; Italian *autunno* is a later borrowing from Latin). In Italy, *-CT-* became *-tr-* (*OCTO* > *otto*, *LACTEM* > *latte*), with the result that the [ot] pronunciation was sometimes spellable correctly as OCT. Hence the attempt to write it correctly as *octurnio*. Unknown to the scribe, [aw] > [o] was also a regular change (e.g. *CAUSA* > *cosa*), and in this case AU was the correct orthography. This form can only be interpreted as the attempt of a vernacular-speaking scribe to come up with official orthography. In the same passage, *decimate* represents the one case of the contemporary vernacular, which in Italy kept the originally nominative form (-e < -AE, in this case) whereas the old grammar would here require the -AS form. The -AS is written in *ovas*, but this is presumably either a misconstrued plural of the neuter plural OVA (> lt. *ova*), "eggs", or a variant form for "sheep", *oves*. *Octurnio* and *ovas* seem thus to be "mythical" forms, created on paper alone in an attempt to be correct. Other, correctly spelt, forms in the *parti libere* are thus likely to be the result of successful attempts to create official-looking orthography for vernacular forms. Further non-imperial usages visible here include -o for -UM (passim), *de pulmentario* for the simple genitive ("two measures of relish", no.3), and the Celtic *camisia* (no.1).

Sabatini is himself a proponent of the two-norm theory, in its diglossic guise, and assumes that scribes knew "Latin" as well as their vernacular; in his view they are consciously vernacularizing their Latin in the *parti libere* to aid subsequent reading aloud. The main thrust of his argument is that the various experiments adopted in the representation of evolved speech formed the practical basis of the subsequent techniques used in the elaboration of new Romance orthographies. But

his insight is more valuable than that. His discoveries are fully consistent with the hypothesis that lawyers did not know of "Latin" as a separate norm; that they learnt established formulae, but need not have used themselves the kind of language expressed in those formulae, nor even have known what they literally meant. At times the formulae are incorrectly transcribed in a manner that betrays their nature as set learnt phrases. Since the *pari libere* are the graphic representation of actual speech, and thus presumably simple enough for the lawyer to read back to the original depositor of the document if required, the main problem for a one-norm Proto-Romance theory of early medieval speech communities is not how to explain the existence of the *pari libere*, but how to visualize the *pari formulae* being read aloud in the absence of a Latinizing pronunciation norm alongside the vernacular.

This problem is dealt with in detail in Chapter 4 with reference to a document from tenth-century León, but the general answer can be summarized here. Anyone wishing to reproduce aloud sections of archaic language will have no problem in reading words that exist in the vernacular in any case; they will be read aloud as that vernacular word. *Veniunt*, for example, would be in Castille read as [βienēn]. Words that have died out would be given the pronunciation that analogously spelt surviving words had in the contemporary vernacular; e.g. intervocalic -r- will be pronounced as [d] in those areas where it voiced in the vernacular (e.g. France and Spain) but as [r] where it remained unvoiced in the vernacular (e.g. Southern Italy). Inflectional morphemes that are no longer current, such as the old passive ending -rur, would be treated in exactly the same way, i.e. in Spain, [dor], in Italy [tor]. The reader would not have any knowledge of how these words sounded half a millennium or more before; that would not prevent their being read aloud as if they were still in common use. The fact that the Latin syntax of these formulae might make the vernacular pronunciation quite meaningless is irrelevant, since in formulae intelligibility is not the point. Lawyers normally prefer their formulae to be opaque to laymen.

This is simple to illustrate with another analogy from Modern English. When reading aloud the Authorized Bible (cf 1611) in church, modern Englishmen give to archaic words and morphemes (such as the -est of *thou makest*, or the -eth of *doeth*) the pronunciation which they would have had were they not archaisms, without worrying about the early seventeenth-century phonetic realization of what is now the /θ/ phoneme, in exactly the same way as they also give those words and morphemes that are not archaic a normal modern twentieth-century vernacular pronunciation. The -est of *makest* will sound exactly like the

-est of *greatest*.

The objection could be raised that this hypothesis of vernacular pronunciation even of the formulaic sections of notarial documents, would still have made the resulting speech sufficiently distinctive for it to count as a separate non-vernacular language, for syntactic and lexical reasons. This argument does not follow unless we always wish to view the lawyers of any community as speaking a literally different language from that of their clients.

Generally there is good reason to see the forms used in legal documents as examples of a separate style or register rather than as a separate language. For example, Crystal and Davy (1969: Chap. 8) have produced an illuminating stylistic study on "The Language of Legal Documents" in English. This is only concerned with modern Britain, but many comments could apply to the documents of early Romance:

There is a strong motivation for any lawyer to turn to a form of words that he knows he can rely on rather than take a chance on concocting something entirely new ... much legal writing is by no means spontaneous but is copied directly from 'form books', as they are called, in which established formulae are collected ... it is a form of language which is about as far removed as possible from informal spontaneous conversation ... (194)

It is especially noticeable that any passage of legal English is usually well studied with archaic words and phrases of a kind that could be used by no one else but lawyers. (207)

This archaism, mainly in syntax but also in vocabulary and morphology (e.g. *witnesseth*), is a known and provable phenomenon; yet it would be absurd to propose that modern British lawyers normally converse with each other according to a pronunciation system which is centuries older than that of the rest of the community. Indeed, they clearly pronounce archaisms according to the phonological system of their modern vernacular. If that is the case now, there is no reason to propose that lawyers in early medieval Romance communities need have spoken anything other than the phonology of the community they were in. We can regard the lawyers of those centuries as drawing up documents in a traditional and practical manner, but we can hardly expect them to be capable of linguistic feats which border on the impossible, such as successfully imitating the phonology of a far distant era. A professional linguist of our time finds it hard to copy a long-defunct phonology accurately; a working medieval lawyer, in a pre-philological age, would have found it impossible in practice to imitate Imperial Latin even if it could have occurred to him to try. In short, the language of legal documents contains no evidence to point to the

coexistence of an archaizing "Latin" pronunciation alongside the vernacular of their region, since it can be naturally explained in other ways. (For the semantics of legal language, see Baldinger 1980: 14c.)

In a more recent study (1978) Sabatini has discussed a seventh-century papyrus from Ravenna which contains the same Latin text in both the Latin and the Greek alphabets. He mentions that, for example, Latin *huc chartule* is written in the Greek alphabet as the equivalent of *ouki caretule* (sic), *donationis* as *donazione*, *portionis* as *porzione*, *fundi* as *fondi*. The significance of such transcriptions is not lost on Sabatini; the scribe using the Greek alphabet interpreted the written *-is* as an [e] in speech, the written *u* of *fundi* as a spoken [o], the written *ti* as either [tsj] or [dʒj], the *h* of *chartule* as representing no sound at all. Here is a member of the legal profession who spoke the vernacular of seventh-century Ravenna. Sabatini's conclusion (453) is surely right; the first "Romance text" later represented more than anything else an orthographical reform. But despite all the evidence Sabatini has adduced, not even he has yet proposed (in print, at least) the hypothesis of a single early Romance pronunciation system, without a coexisting Latin system, for lawyers in early medieval communities; the hypothesis seems natural to a historical linguist and is quite compatible with the surviving evidence.

The Evidence of "Rhythmic" Poetry

Another kind of text which might be expected to yield us some information concerning the relationship between speech and writing in the Early Middle Ages is the variety of poetry known as "rhythmic". This evidence is slippery, for several reasons. Firstly, much poetry is consciously written in imitation of earlier models. Secondly, there is no need to assume that even metres ostensibly based on an equal number of syllables per line always require these mathematics to be satisfied exactly. Thirdly, the same line of music can be used for texts of differing numbers of syllables. Fourthly, hymns in particular, but other songs also, often induce performers to pronounce in a manner that would be regarded as strange outside a musical context. And fifthly, modern editions of surviving texts have often at some stage been "emended" in accordance with later scholars' preconceptions concerning what these texts ought to have been. Even so, it is instructive for our present purposes to consider the genre.

The techniques of post-Classical poetry are usually divided into two categories. "Quantitative", or "metrical", verses use the old traditional metres based on distinctions of vowel quantity and syllable

length. "Qualitative", or "rhythmic", poetry takes no account of the ancient distinctions of length; instead it imitates the old forms by substituting a vernacular stress pattern for the original length pattern. Stress concerns the greater effort given to the pronunciation of some syllables than is given to others; some monosyllables (e.g. *et*) and a few disyllables (e.g. *quasi*) were usually unstressed, "atonic", but otherwise each word in Proto-Romance seems to have had a predetermined stress, and rhythmic patterns could be created by the arrangement of the stresses of the words within the lines. As Norberg (1968) convincingly showed, much poetry is based in this way on accentuation patterns that we can plausibly envisage as being those of normal vernacular.

The production of metric Latin poetry on the original quantitative basis is a *recherché* pursuit of the learned, an esoteric accomplishment of antiquarians, and has been so ever since the quantitative distinctions ceased to have any counterpart in ordinary speech. The existence of a few early medieval quantitative verses cannot be taken as evidence that phonemic length persisted in the speech of their composers any more than it is evidence of phonemic length in the speech of nineteenth-century scholars who also dabbled in the same pastime; it has to be seen now, as then, as a symptom of a sophisticated education system. For example, seventh-century Spanish poets used classical metres for much of their extant work, and, as Raby observed (1934: 153), this poetry "derives its form and expression from what remained of the methods and practices of the ancient schools". Visigothic Spain was educationally the most sophisticated of the seventh-century Romance realms, and the quantitative verse there produced is visible evidence of this preeminence. Norberg (1965) suggested that after the length distinctions fell, metric verse was recited with the normal vernacular stresses on the words; otherwise, the only available technique would have involved stressing the long syllables and giving no stress to the short syllables, and in the many cases where the consequentially produced stresses did not correspond to those of the vernacular words involved, that can only have been thought of as a peculiarity of the rules of such sophisticated creations. When the rules of metrics were known, it was because they had been studied in authoritative works, and such a result as odd-sounding lines would thus be thought a sign of erudition rather than eccentricity.

Rhythmic verse, however, usually was destined for oral performance in an intelligible manner; if any evidence concerning pronunciation is to be discovered in poetry of the period, it will be here. What classicists once considered to be the ignorant miscegenations of

uncouth barbarians have now been reinterpreted as the results of techniques of composing in a poetic style of the vernacular. If such poems look "Latin" on paper, that is a consequence of their being written before the orthographic innovation consequent upon the Carolingian reforms.

Rhythmic poetry is not, however, the same as popular verse. Popular songs and verses undoubtedly existed then, as they have in all known cultures, but by their very nature have left little trace. Rhythmic poetry was a written genre. Many poems of this type that have survived were shown by Norberg to belong to carefully defined schemata and patterns that needed to be learnt. The number of syllables was the most important criterion for a line of "rhythmic" verse; the academic definition of a syllable was included in *Grammatica*. "Rhythmic poetry" is thus an educated genre with its own standards, rather than a collective label for all verse that fails to conform to quantitative metrics: "A rhythmic poem is one in which the old system has been replaced by a new one, not a poem whose characteristics are the absence of rules and anarchy" (Norberg 1958: 94). It is a simplification, therefore, to call all non-metric verse "rhythmic".

Regularity was based on the number of syllables per line, and in the grammatical tradition syllables were defined on the basis of written vowels. The only features of pronunciation in these poems that might not have accorded with normal vernacular thus concern vowels; in particular the occasional avoidance of syncope of short vowels and a routine tendency to leave vowels in hiatus. This reluctance to elide is often extended to an apparently deliberate avoidance of having a word-final vowel followed by a word beginning with a vowel at all. These phenomena probably have something to do with the normal style of song performance, which is usually slower than spontaneous speech.

Some rhythmic verses imitate most of the structural aspects of an old form. In these, patterns of stressed and unstressed syllables replace the old patterns of long and short ones, and caesuras are still observed at the mid-points of longer lines. Rhyme and assonance begin to be used regularly in addition, and the formal complexities are often increased by the inclusion of such intellectual games as acrostics and telestichs. In consequence, many of these compositions represent an intellectual achievement, whatever we may think in this Romantic age of their poetic value. They are signs of changes in procedure and in taste, but are nonetheless the conscious products of skilful writers. Other rhythmic verses disregard the strict pattern of stressed and unstressed syllables except before the caesura and at the end of a line; in

these the number of syllables (i.e. of written vowels and Classical diphthongs) per line remains theoretically constant, and the caesura is in the traditional place, preceded by a paroemion. These verses are the intellectual forerunners of the later Romance verse systems, many of which follow a similar pattern and come to be identified through the same criteria of rhythmic pattern and regularity of line length (e.g. the hendecasyllable).

For example, the ninth-century Spanish hymn to St Jerome (*AH* XXVII 126) begins as follows:

1. Christus est virtus, patris sapientia,
Cunctos qui replet spiritali gratia,
Ut possint probe digere normam
Et proximum illustrare opaca,
Ut digne possint fructu caelestia.
2. Ipsius dono perfatus egregius
Olim hic vates nomine Jeronymus,
Omnibus notus doctarum fontibus,
Cunctos irigans ex alnis dogmatibus,
Ut sol resplendet in ortu igniconus.
3. Hic procul cuncta saeculi negotia
Percalcans pede velut spurcissima
Dedecorosa respuitque saecula,
Alens inopum egenaque viscera,
Sibi aeterna acquires stipendia.

(Blume 1897: 180)

The basic pattern here is of 5 syllables plus 7 syllables; by the ninth century this rhythmic "iambic" line is as traditional as any other. The caesura is observed throughout the seventeen stanzas. Elision is conspicuously not required, and the music may have also had the effect of relengthening into full syllables some vowels that were in the process of syncope in normal speech (e.g. the *u* in *saecula*, 3.3). The prothetic *e* that Blume adds before *spurcissima* (3.2) might be unnecessary in view of the possible emendation of *velut* to *veluti*, but in the following lines it seems unavoidable:

- 4.2 Et sese valde (e)stringit ac regulam
- 5.2 Gentiliumque summo cum (e)studio
- 6.4 Obsequiorum tantum ut (e)speritum
- 14.2 Te criminatur (e)stultorum factio
- 17.4 Iugiter, semper, per aevi (e)spatio

This is not a regularity, however: the prothesis is not required in other lines such as 1.2 and 3.5 above, and the metric licence common to hymn-singing in most languages seems to have been exploited here.

Rhyme is another possible indication of the composer's speech: in this poem, if we can assume that every line in a stanza is meant to rhyme, we can deduce that [m] is not pronounced from *normulum* in 1.3, which rhymes with four words in -a. It also seems that -u and -o represent the one sound [o], for stanza 6 has the following five rhyme words: *corpusculum*, *ferculo*, *vulgulum*, *spiritum*, *turnulo*. *Spiritu* also rhymes with four -o words in stanza 7, and with three -o words (and *Deus*, in Blume's emendation) in stanza 17; there is no support here for the normal view that the -u in Modern Spanish *espíritu* derives from Early Medieval "learned" ecclesiastical pronunciation of the word.

Within hymns and comparable performed verse, the preservation of syllables that otherwise might have been disappearing seems to be the only reconstructable feature distinguishing hymns from vernacular. The presence of an internal vowel in *saeculo* in hymns may seem to be attested, but that in performance that vowel was [u], or that the congregation pronounced the word like this in any other context, is an unjustifiable conclusion. English hymns can give the word *heaven* any number of syllables from one to four (or more), but that does not affect our pronunciation of it elsewhere; it is a feature of the style of song-performance, [s]leglo could similarly be made to cover three notes.

The defining feature of a rhythmic verse form thus, with time, became a regular number of syllables (written vowels) per line. This might appear to have implications for our reconstruction of the number of syllables with which words were pronounced, but because of the common feature of artificial slowness of song-performance, it is not of great help in many cases. The argument does not lead far in communities where most vernacular words of this time seem likely to have still had as many syllables as their Imperial etyma, such as Italy, Visigothic Spain, or even Moslem Spain. It might be relevant to ninth-, tenth- and eleventh-century Northern Spain, although the available material is small and awkward to analyse for other reasons (see Chapter 4); it could well be relevant to pre-Carolingian Northern France, though the reforms arrive here before other areas and the available material is similarly not large. Even so, Norberg has shown convincingly that many features of specifically vernacular speech have to be postulated in his analyses in order to account for apparent syllabic anomalies, even after the traditional assumption of the existence of a "Latin" speech has led to the distorting introduction of palaeographic anomalies uninvited "emendations". The formation of semivocalic glides from what were originally unstressed close vowels, with the consequent formation of monosyllabic diphthongs instead of two adjacent syllabic vowels, is commonly

attested: "the oldest example of a rhythmic poem, Saint Augustine's alphabetical psalm, has several examples such as *ecclesiam*, *gladium*, *rescō*, *petrū*, *fiērī*, *sūm*" (Norberg 1958: 29). The prothetic vowel added in vernacular to initial [s] followed by a consonant (as in the examples quoted above) is required in many lines, particularly from Spain, in order to preserve syllabic regularity (Norberg 1958: 31; Blume 1897: 54). And despite the general effect that song has on the avoidance of syncope, there are also cases of syncope of unstressed close vowels in internal syllables in long words. Many of these poems are composed to be sung to preexisting melodies; they are not necessarily isosyllabically organized, with a syllable per note, as was to happen in the Carolingian church (most notably in the sequence), so there is no reason to suppose that the music would in practice inhibit natural syncope, epenthesis, or any other development, in subsequent performance, however the author originally expected and envisaged its reproduction (cp. Norberg 1954: 25-29; 1958: 186-87). The surviving texts are not the records of actual performances.

The evidence of rhyme and of vowel assonance also supports the conclusion that this verse was normally performed with the ordinary vernacular sounds of the performer. Norberg points out, for example, (1958: 48), that even an Archbishop, the seventh-century Eugenio de Toledo, assonated *delectatio* with *solacium*, *recogito* with *transeunt*, etc., following the regular vernacular pattern of collapsing the old distinctions between (i) stressed originally short /i/ and originally long /ō/, and between both of these and originally short /ō/ in unstressed positions, and (ii) stressed originally short /i/ and originally long /ē/, and between both of these and originally short /ē/ in unstressed positions, into one /o/ and one /e/ phoneme respectively. (Stressed /ē/ and /ō/ were the ones to diphthongize). Consequently the same sound is represented by the *o* of *delectatio* and *recogito* and the *u* of *solacium* and *transeunt*; and the same sound is represented by the *i* of *recogito* and the *e* of *transeunt*. The lines in question are from Eugenio's *Carmen* XIV 25-28:

Potus cibique nalla delectatio
Lamenta sola conferunt solacium.
Haec taediosa mente dum recogito,
Libet, relictis omnibus quae transeunt.

(Vollmer 1895: 243)

In France, Venantius Fortunatus assonated *concinir* with *carnine*, *redditum* with *prospero*, etc., in the same way: e.g. *inpleta sunt quae concinir* / *David fidelis carnine* (MGH *Auctores Antiquissimi* IV 34;

Norberg 1968: 45). This collapse of atonic /ī, ē, ē, and /ō, ō, ū/ and of tonic /ī, ē, ē, and /ō, ō, ū/ is found in assonance patterns of Latin poetry from all Romance-speaking communities, but not in the poetry composed in the British Isles, where there was little native Romance vernacular after the fifth century and the poems were performed according to a largely artificial usage kept alive through spelling-pronunciations. The Anglo-Saxon habits may in this respect have been roughly similar to Muller's envisaged language of the fifth to the early eighth centuries, but they did not coexist with an evolving Romance in the same community (see below, *re Bede*).

The evidence of rhythmic poetry from pre-Carolingian times suggests that it was originally expected to have the normal vernacular sounds of the time and place where it was composed. Neither this, nor the erudite metric verse, can be adduced as evidence in favour of the hypothesis that their authors regularly used an archaic Latinate speech systematically distinct from the coexisting Old Romance.

Some of the surviving Merovingian examples have baffled even Norberg with their apparent lack of organization on recognizable patterns: e.g. King Chilperic's hymn to St Médard, Theophilus of Corbie's hymn XII, parts of the versified life of St Eligius (Norberg 1954: 32, 52, 70 n.20).

The third of these was examined by Meyer (1936: 234-42), whose study suggests that the author may have thought that metric hexameters are lines containing six word-stresses. By permitting two accents on some long words, and ignoring the atonic words, such lines as the following can become "regular" after a fashion (the accents are Meyer's); e.g. lines 495-500:

Satis fecisse me réor	succincto cāmine pléctro;
plura nam réferre grāvor,	necesse quōque nec opīor.
Hæc pāuca hexāmētis	reciprocate stūdiū versiculis,
āctūdētibz dīgits	tānti amorē antēstīis.
Chrīstēm hæc mētrica	volūtem immōque perpaucā
ratiōne compōnere	non differam breviter explanāre.

It is possible that the poet, probably writing soon after Bishop Eligius' death in 660, was working on a regular proto-Romance oral basis of some kind. For example, the second half of 1.497 could become a rhythmic hexameter, with the long/short pattern replaced by a stress/unstressed one, if the *-are* of *reciprocate* could be seen as a single stressed syllable, the *-ui* as a single unstressed syllable, and the *u* of *versiculis* as absent, all of which are quite plausible reconstructions for the Romance of the time. It would in this way be possible to turn much of this into a popular-stress-based poem, although there are too many imponderables

for any such conjecture to be convincing in detail. The general point holds, though: whatever these verses are, it seems absurd to suppose that their authors were not, in fact, operating with Merovingian vernacular, which was written down in the only available spelling, the traditional Roman. More successfully elaborated poems, on the other hand, are signs of well-educated poets rather than of bilinguals.

The Evidence of the Visigothic Liturgy

Liturgy is a field that the philologist cannot ignore. The hispanic liturgy is fuller and wider than probably any other in the West; through many sources we know that the most notable literary figures of Spain were involved in preparing the texts . . .

(Díaz y Díaz 1976: 50)

Collins (1977: 33) has stated baldly but accurately that most of the current accepted wisdom concerning Visigothic Spanish society is based on "whimsical concatenation of fantasy". A large corpus of texts survives from the period 589-711, however, including the many books of the Visigothic (or "mozarabic") liturgy elaborated by the best group of scholars in Europe at the time, led by Isidore of Seville. If there is to be found anywhere in Romance Europe between 500 and 700 evidence of a "learned" pronunciation norm distinct from the vernacular, one would expect it to have existed in the performance of the Visigothic liturgy. The texts are long and provide a full picture of liturgical practice; they are filled with rhetoric and elevated style, such as clausular rhythm and high-sounding vocabulary.

Rhetorical skills, literary style and elevated vocabulary, however, are all teachable and learnable in accordance with the techniques of written composition incorporated in the grammars used in Visigothic Spain. Whether they, or arcane and archaic pronunciations, existed in the spoken vernacular of those who prepared and used the liturgy, is open to doubt. It used to be assumed that any spelling or grammatical errors found in the texts were errors introduced by subsequent copyists; Férotin (1912: xix), for example, commented adversely on "the incorrectness of mozarabic manuscripts and the ignorance of the copyists". As a consequence, the inherited textual versions ran the risk of being refully "corrected" by modern editors as if they were school-boy compositions (see the strongly-worded comments of GIL 1973b). Fortunately the evidence remains; Díaz y Díaz studied the language of the manuscripts and demonstrated that most of the apparently "incorrect" features of these texts are faithful to the original versions

rather than ascribable to the negligence of later copyists? Most of these "in-corrections" seem in fact to be forms used in seventh-century Spain, as far as can be reconstructed from later evidence. This is even visible in the vocabulary, which the authors were in general keen to keep as "elevated" as possible.

As regards morphology, many reconstructable features of seventh-century Spanish vernacular are perceptible in the texts. For example, we can guess that originally fourth-declension nouns have been by now reallocated to the second (cp. GRADUM, MANUM > Spanish *grado, mano*); this is here attested e.g. by the liturgical form *factos* (Orac. 969), accusative plural, instead of the Imperial *factus*. We can guess that originally neuter plural forms in *-a* have by now been reinterpreted as feminine singulars; the *q*-forms are thus available for pluralization with an etymologically ridiculous *-s*. This is attested in the liturgy e.g. by the form *erratas* (Orac. 361.21); the Imperial *errata* was already a plural of the neuter *erratum*. Such reinterpreted forms are also capable in the liturgy of being the subject of verbs in the singular (*nos subsequatur munera pietatis*, Orac. 989), whereas Imperial Latin neuter plurals took plural verbs. We can guess that most verbs of the third conjugation, with infinitives in *-ere* with an open [e], have by now been reallocated either to the second conjugation, with infinitives in *-ere* with a closed [e], or to the fourth conjugation, with infinitives in *-ire*; Spanish has no distinct third conjugation, and the verbs in question are now either conjugated as *-er* verbs (e.g. VINCERE > *vencer*, CAPERE > *caber*) or as *-ir* verbs (e.g. FUGERE > *huir*, SCRIBERE > *escribir*). This is attested in the liturgy by e.g. the present-tense form *vinces* (Orac. 413), instead of Imperial *vincis*; the form *capuisse* (Orac. 686) (which was the past infinitive of *-ERE* verbs: cp. Modern Spanish *cupe* < *CAPUI, not < CEPT) rather than the Imperial *cepisse*; and the form *fugire* (Sac. 288.34) rather than the Imperial *fugere*. We know from the evidence of Spanish that many irregular verbs lost irregular forms in favour of regular forms created by analogy with the normal pattern; in the liturgy we find, for example, present-tense *auferes* and *auferet* (Orac. 95.26; Sac. 50.9) for Imperial *auferi*, *auferet* (cp. Spanish *sufre* < *SUFFERET rather than *SUFFERT*). The loss of the passive caused many deponent verbs to be reallocated to active forms (e.g. MORI > *morir*); this is attested in the liturgy by e.g. *egredere* (Orac. 46.2).

2. Díaz y Díaz (1965). Printed editions of the texts used in his study are Vives (1946) = Orac.; Férotin (1904) = Orac. and (1912) = Sac.; Brou and Vives (1959). The subsequent examples all come from Díaz y Díaz's study unless otherwise stated.

rather than Imperial *egredi*. We know that medieval Spanish had a love of compound prepositions with *de*, such as *después* < DE POST, *después* < DE EX POST, *desde* < DE EX DE; in the liturgy we find, as single words, *depost* (Orac. 152.24) and *desub* (Sac. 128.40). We can see the latter as evidence for a seventh-century vernacular [dezó], *desó*, even though *so* eventually survives in preference (as in *so pena de excomión*) and *deso* is not attested as such. We know that Spanish nouns normally derive from the accusative form of the original etymon; that the accusative was probably the only surviving form even as early as this can be seen by the way that prepositions which once required other cases take accusatives in the liturgy: e.g. *de insidias* (Orac. 240.1) rather than *de insidiis*, the Imperial ablative. These, and other, features are sufficient to show that the evolved vernacular form in question is the one used in the morphology of the normal speech of those who prepared the text; despite the lengthy lists of old morphology found in all their manuals of "correct" writing, they were not able to avoid their natural usage on every occasion. In contrast, the relatively insignificant probable archaisms of morphology that occur in these texts, such as the synthetic passive forms and imperatives in *-to* that the reconstructionists of Proto-Romance would prefer to think did not exist in the normal vernacular at this time, are simple to account for as being the result of education rather than symptoms of the writers' own usage.

Similarly, the spelling is sufficiently non-standard to suggest that the "correct" forms are the result of strict training rather than a simple phonetic transcription of archaic pronunciation (as English written *knight* reflects a well-taught speller rather than a producer of [knixt] or [knight]). It seems probable, as we attempt to reconstruct the vernacular of the time, that between two vowels the original [t], [k], [f] and [p] — traditionally spelt respectively *t*, *c* (or *qu*), *f* (or *ph*) and *p* — have become in speech the voiced counterparts of these sounds: i.e. [v] > [d], [k] > [g], [f] > [v], [p] > [b]. The spelling taught and learnt has not changed, of course, so the norm is — if the reconstruction is correct — for sounds in many words pronounced with a [g], [d], [v], [b] to be written as *c* (or *qu*), *t*, *f*, *p*, respectively. This produces uncertainty in both directions; words that used traditionally to be spelt with *c* (or *qu*), *t*, *f* (or *ph*), or *p* can be misspelt with a *g*, *d*, *v* or *b* because that is how they are now pronounced; conversely, words that traditionally were spelt with *g*, *d*, *v*, or *b* can be misspelt with a *c*, *t*, *f* or *p* precisely because so many words are indeed at this time written with *c*, *t*, *f* or *p* and pronounced with a [k], [d], [v] or [b]. The first type of misspelling even occurs in Greek borrowings and technical

ecclesiastical terms such as *eglesiae* (e.g. *Orac.* 1030: Greek ἐκκλησία > Latin *ecclesia* > Spanish *iglesia*), *paraglitum* (*Orac.* 1036: "paraclete"), and *psalmograue* (*Orac.* 705) – in which the penultimate letter *u* represents what was normally spelt as *ph*, Gk. ψαλμῳγῆ; other examples include *memedipsum* (*Ord.* 253: i.e. MEMET PSM: > Sp. *mi mismo*, cp. Italian *medesimo*), *neborium* (*Ord.* 417: NEBORUM; in Spanish the [b] has since disappeared regularly before a syncopated internal vowel and a dental, > *nieto*), *preuato* and *prouina* (*Ord.* 440 and *Sacr.* 116.31: PREFATIO and PROFANA; Sp. *prefacio* and *provano* are fifteenth-century borrowings). The similar neutralization of the [t]/[d] distinction at the end of words leads to written forms such as the common *reliquid* for *reliquit*. The second type of misspelling, in which *e* and *t* are used to represent a sound which had never been anything other than [g] and [d], is exemplifiable with *clorize* (*Orac.* 120: for *glorize*), *discent* (*Sacr.* 609.2: for *ghiscent*), and the common *aliut* (for *aliud*), *at* (for *ad*). The problems concerning the manner of written representation of final [d] are a perpetual preoccupation of the early medieval grammarians as they attempt to prescribe when to write *it* or *id*, *quit* or *quid*, *ut* or *ad*, etc.; it is one of the few linguistic features discussed by St Isidore himself. It does not look as though these scholars spoke a dialect in which written final -t and -d represented distinct sounds. (The old hypothesis that the Spanish of Modern Spain preserved unvoiced intervocalic [t], [p], [k], has been by now exploded; see e.g. Corriente 1978, Galnás 1977).

Other evolutions of the sound system that have occurred by the seventh century in the vernacular of Spain are also perceptible in its transcription in the liturgy. The positional instability of the [r] sound in Iberia (e.g. *quebrar* < CREPARE, *entrego* < INTEGRO, etc.) is visible in written *prespicuus* (*Orac.* 841) and *prescruator* (*Sacr.* 329), for *perpicuus* and *perscruator*. The common Spanish metathesis of [r] and [l] (e.g. *peligro* < PERICULUM) is visible in written *fragabit* (*Sacr.* 461.23), for *fragabit*. The intrusion of unexpected [r] after [t] (e.g. *estrella* < STELLAM) is perceptible in *retorsere* (*Sacr.* 118.7), for *retorsere*. The prothetic [e-] of words beginning with [s] plus a consonant is not usually transcribed in liturgical manuscripts any more than in hymns, being an allophonic and predictable variant of /s/, as [we] was of [b] and [je] was of /j/, but the presence of the extra vowel in speech is deducible from the common converse misspelling of *sie* (*Sacr.* 276.15, etc.) for original *isre* (by now pronounced [ese]), as in Sp. *este*, on the analogy of the correct spelling *sio* for contemporary *esio* (STO > *esio* > *esio*). Even the use of clausal *thyns* has

been shown to require a basis in the vernacular prothesis; the syllable apparently missing in the clausula *coronados status* (*Orac.* 8 and 674), for example, is discoverable in a pronunciation comparable to modern Spanish *estatus*.

This study by Díaz y Díaz, ignored by Romance philologists, ought to have had the effect of removing one of the last legs on which the two-norm theory can rest. If the liturgy of the most literate area of seventh-century Europe required and implied a contemporary vernacular pronunciation from both its authors and its performers, who is there left in the whole of seventh-century Romania to whom the hypothetical existence of a Latinized norm can apply? The common admission that only a very few men of the time "knew Latin" needs to be revised into a firm recognition that nobody of that time "knew Latin", neither imperial vernacular nor the artificial language used since the twelfth-century Renaissance: Latin as we know it now had not yet been invented. What existed was the vernacular Old Romance of seventh-century Spain; and when that was written, traditional orthography and the prescriptions of the "grammatical" manuals of writing succeeded in making the written version an inexact copy of the spoken. It was only the assumption that the clerics must have spoken "Latin" that prevented modern scholars from perceiving that they could not have spoken "Latin"; and as Collins observed, with his customary lapidary directness, "interpretations based on *a priori* assumptions as to how clerics think and behave must be regarded with some scepticism" (1977: 35).

As regards the Roman liturgy used elsewhere, the extensive researches of Mohrmann (e.g. 1961b) have effectively discounted the possibility that the Roman liturgy of pre-Carolingian times was pronounced differently from the vernacular of the congregation, whatever its stylistic or lexical level. Saint Augustine himself said, "melius est reprehendat nos grammatici quam non intelligent populi" as regards grammatical comprehensibility; and although Mohrmann does believe in a kind of bilingualism in the Christian hierarchy, she noticeably refrains from extending this to support for the theory of separate manners of pronunciation. In all her voluminous work there appears no hint that liturgical or Christian Latin involved a non-vernacular phonetic system. We can be sure that she would have noticed it if it had.

Early medieval liturgists seem no more to be aware of the modern consensus that they did not speak contemporary Romance than were poets, lawyers, or the fifth-century grammarians; the crucial evidence has effectively been narrowed to the testimony of the

grammarians of 500-770 A.D. If they cannot be shown to demonstrate the existence of the two norms at that time, the suggestion that Early Romance alone existed then can surely be resisted no more.

The Evidence of Grammarians, 500-770

Most grammatical training from 400 to 800 A.D. was based on the works of Aelius Donatus, written in the mid fourth century, and his subsequent commentators (Keil 1864: 353-402; Holtz 1981). Donatus was giving instruction on how to compose acceptable written works to a community for which Imperial Latin was the vernacular, already known, and consequently he hardly mentions pronunciation. These works were the basis of education in the four following centuries and beyond, but can hardly have been of much practical use to anyone wishing to teach old-fashioned pronunciation to Romance-speaking pupils. If the distinction between Latin and Romance speech existed at all between the Roman and Carolingian empires, it would have been likely to be reflected, consciously or unconsciously, in the works composed by grammarians of that period; in particular, in Cassiodorus of Vivarium (Italy), Isidore of Seville, and Julian of Toledo, who, unlike Bede and the insular grammarians, were writing for Romance-speaking clients.

Cassiodorus

In 573, apparently at the age of 93, the Italian scholar Cassiodorus (c.480-575) wrote a treatise *De Orthographia* (PL LXX 1240-70; Keil VII 143-210. Cp. also Cappuyns 1949; Roger 1905: 175-86; Riché 1962: 204-12). This was intended as a guide to the monks in his scriptorium at Vivarium in the art of manuscript copying, and pays particular attention to the circumstances in which the correction of texts is deemed permissible. It is based on the comments of eight previous authorities, including the first two books of the *Institutiones* of Priscian, which seem to be otherwise unknown in Romance Europe before the late eighth century. *De Orthographia* is an accurate title for this work, since the desire for acceptable spelling is the incentive for producing it, but there are occasional references to pronunciation too.

The pronunciation Cassiodorus implies for himself and his monastic audience seems, so far as we can tell, to be simply that of his time and place. This was not greatly evolved from the imperial vernacular, but there are enough indications to show that he is not prescribing any non-vernacular "literate" pronunciation patterns based on the spelling. For example, in the following case Cassiodorus specifies

that the pronunciation and the spelling ought to be different: although *tantus* is, as a matter of ordinary vernacular practice, pronounced with [n], it is, for etymological reasons, to be written *tantius*:

Tantius et *quantus* in medio *m* habere debent, *quam* enim et *tam* est: unde *quantitas*, *quantus*, *tantus*. Nec quosdam moveat, si *n* sonat; jam enim supra docui *n* sonare debere, tametsi in prima scriptura *m* posita sit. (PL 1245C; Keil VII 152.3-5)

Similarly, he explains that although, as another matter of simple vernacular fact, the final written *-m* before a word-initial vowel represents no sound, it is to be written even so:

igitur si duo verba coniungantur, quorum prius *m* consonantem novissimam habeat, posterius a vocalibus incipiat, *m* consonans perscribitur quidem; caeterum in enuntiando durum et barbarum sonat. (PL 1243B; Keil VII 147.27 - 148.2)

This prescription is reiterated more baldly later (PL 1267C; Keil VII 206.17-18): "*M* litteram, ad vocales primo loco in verbis positas si accesserit, non enuntiabimus."

Much of the *De Orthographia* (PL 1252D - 1263A; Keil VII 167-199) concerns the prescriptions of Adamantius Martyrius regarding when to write *b* and when to write *v*. The implication of the tone adopted is unmistakably that with some words neither Cassiodorus nor the scribes to whom the instructions are directed can see any phonetic distinction correlated with this orthographic distinction. The problems over when to write *b* are discussed in a similar manner (PL 1263A - 1265A; Keil VII 199-202). Where there may be cases of possible alternative pronunciations as well as possible alternative spellings, current practice (*consuetudo*) is the arbiter: concerning *maximus* or *maxumus*, melius tamen est, et ad enuntiandum, et ad scribendum, *i* litteram pro *u* ponere, in quod iam consuetudo inclinavit. (PL 1244D; Keil VII 150.16-17)

The current normal pronunciation determines the spelling:

quando autem fiant, quando non, sono internoscimus: *accedo* duo *cc*, *attuli* duo *tt*... in his non solum propter lenitatem [levitatem: Keil] consonantes mutantur, sed etiam quod nullo modo sonare *d* littera potest. Ubi sonat, et ibi scribitur [est ubi sonet et ubi scribatur: Keil], cum *f* consonanti adiungitur; ut *adfluo*, *adflui*, *adfectus*; contra *b* non sonat, *offui*, *offero*, *offendo*. (PL 1245B; Keil VII 151.13-17)

This is a simple statement that in the speech of his time there is a [d] - or, more probably, an unvoiced [t], in the event - in such *-df-* compounds as *adflui*, but not in *accedo* or *attuli*; nor is there a [b] in *offui*. This is fully compatible with reconstructable evidence of sixth-century Italian Romance, although [df] became [f] or [ff] over the next three

centuries. Phonetic factors make it plausible for [bʃ] (or [pʃ]), two labials, [dʃ], two dentals, and [dʃk], dental plus velar, to simplify earlier than the [dʃ] cluster.

In his preface, Cassiodorus discusses *pronuntiatio*. This does not, however, mean "pronunciation", for which his term is *enuntiatio*, and the verb *enuntio*. *Pronuntiatio* refers to the manner of delivery, as it had done for centuries (e.g. in Cicero, *De Inventione* I vii 9: "pronuntiatio est ex reſum et verborum dignitate vocis et corporis moderatio"). This is the requirement to make a reading aloud intelligible through appropriate intonation, speed and pauses:

... ut quae incipis bene discere, ad finem perfectionis inoffensa debeas pronuntiatione perducere. Nam si vobis adst capendi desiderium, quae prius per moram quaesistis, protinus inoffensa velocitate transcuratis. Gloriosum profecto studium ... quod loqui debeas, competenter scribere, et quae scripta sunt, sine aliqua erroris ambiguitate profiteri. (PL 1241C; Kell VII 145.3-8)

illud etiam vos magnopere credidi commemorandos, ut distinctiones sensuum sollicita mente perquirere ac ponere debeatis, sine quibus neque legere quidquam competenter neque intelligere praevaleamus. (PL 1242A; Kell VII 145.28-30)

The following section concerns pauses, *distinctiones*, *seu posturas*. These remarks on *pronuntiatio* concern the natural good habits to adopt when reading aloud, how to make the texts intelligible, by competent presentation of the ordinary pronunciation, but not how to make it "Latinate" rather than Romance. If the two norms did coexist here, it is strange that Cassiodorus' comments include no reflection of their possible confusion; reinforcement of the "Latin" version would have required considerably more precise specification than Cassiodorus gives it.

There are further indications that Cassiodorus and his audience used their own vernacular in his *Institutiones Divinarum et Saecularium Litterarum* (Mynors 1937; PL LXX 1105-50). For example, he mentioned *euphonia* in *Inst. Div. Litt.* xv 9 (PL 1129A), but as Cappuyns pointed out (1949: col.1405), "he is not only thinking of particular spellings, grammatical forms or syntactic arrangements, but of whole phrases whose function is merely decorative". This Chapter xv is entitled *Sub qua cauteſta reſeſt debent coeſtſta auctoſitas*, and concerns the appropriateness of emending scriptural or patristic works that appear to have been wrongly copied. For example (xv 9):

In verbis quae accusativis et ablativis praepositionibus servantur, situm motumque diligenter observa, quoniam librari grammaticae

artis expertes ibi maxime probantur errare. Nam si *m* litteram inconvenienter addas aut demas, dictio tota confusa est. (PL 1128C)

This last remark is not a reference to any prescribed Latinate pronunciation, but a reference to the morphological confusion that can be caused by wrongly transcribing or failing to transcribe the final *-m* that distinguishes in writing between the accusative and the ablative case of many nouns, with the result that the reader may be unable to understand the expression (*dictio*) concerned. This seems to confirm that even the educated no longer naturally use the old case system or pronounce anything corresponding to the accusative *-m*, and that they regard the letter *m* at the end of a noun as a written signposting device to indicate the syntactic relationships of nouns rather than the representation of a live phonetic entity. In general, discrepancies between speech and spelling can be taken in native speakers' stride. The forms preferred *respectu euphoniae* or *propter euphoniā* are the respectably Classical compounds *illuminatio*, *irrisio*, *immutabilis*, *impudens*, *improbatus*, and *quicquam*; orthographic practice is prescribed to suit normal pronunciation, "ne articulatione vocis pulchra modulatio, peregrinis litteris maculata, absque potius et indecora reddatur" (1129A). The final consonants of *in* and *quid* had a potentially destructive psychological validity, and an overanxious monk might have caused confusion by writing *invisio* when he and his audience naturally said [in-].

Neither this Chapter xv, nor Chapter xxiv, *Quo studio scriptura sancta cum expositioſibus legenda sit*, nor Chapter xxx, *De antiquariis et commemoratione orthographiae* (in which he apparently refers to an earlier, now lost, *De Orthographia* of his own), suggest that Cassiodorus was aware of the existence of regular phonetic differences between the speech of his monks and that of their secular neighbours. On the contrary, it seems that they used the vernacular of their time and place. Cassiodorus is simply reinforcing the traditional spelling requirements in a community who might be tempted by their vernacular to cut corners. If he tells them to be sure to write the *-m* on *solem*, that is equivalent to telling modern typists not to miss the *k* off *knife*; it does not imply that anyone pronounced [m] then or pronounces [k] now.

Gaul appears not to have harboured many sixth-century grammarians. Indeed, Caesarius, bishop of Arles from 502 to 542, makes it clear that it was hard work encouraging literacy at all. His sermon *De Adulterate Legenda* is an attempt to promote the knowledge of reading and writing in general, including the reading of sacred works, with no hint that a specialised pronunciation is required to do so. As regards the

reading of lessons from the Old Testament, Caesarius prescribes that the reader should make the points as intelligibly as possible for the assembled congregation (Morin 1953: sermons VIII, I 41-45 and LXXXIX, I 865-69). This would be impractical in sixth-century France if the reading pronunciation were Latinate, as Lot (1931) pointed out. The congregation would not have understood.

In the seventh century, attention turns to Spain.

Isidore of Seville

Modern histories of Medieval Europe tend to regard the sixth and seventh centuries as a dismal low point in European culture. This can only come from a deep-rooted and apparently long-standing instinct on the part of modern historians that Spain is not part of Europe. Latinists seem to feel the same; Mohrmann, for example, is merely one of many modern medieval Latinists to have devoted almost no attention at all to Spain.

The Visigothic Renaissance in Spain of the late sixth and seventh centuries was a constructive revival of Latin culture in a Christian context. Spanish scholars of the age appear as the epitome of cheerfulness in comparison with writers further north-east. The Byzantine Empire touched the Mediterranean edge of the Peninsula until 629, and much of the North African coast until the Moslems came; the cultural contexts this involved may have been of the North African Latin heritage rather than anything particularly Byzantine (see Scudieri 1959). Gothic was almost certainly not spoken by the seventh century, and both Goths and Romans consciously wished to preserve the old Roman culture (Wallace-Hadrill 1967: chap. 6).

The most important representative of this Renaissance is Isidore, bishop of Seville and then archbishop of Toledo (c. 560-636). In retrospect, his great work, *Etymologiae*, sive *Origines* (Lindsay 1962), has been cast in the rôle of a hinge between Classical and European Medieval culture; at the time Isidore was naturally unaware of the future use to which his work was to be put, and was concentrating on collecting as much of the knowledge and practice of Classical civilisation as he could for the benefit of his contemporaries. Until Fontaine's book (1959), Isidore's work was not normally seen in context, partly because the context was not easy to assess. It is probable, however, that the school system of the Roman Empire cannot have lasted between the early fifth and the mid sixth century in anything more than a rudimentary form. Díaz y Díaz pointed out (1976a: 23) that in the late sixth century there was nobody in the diocese of Cartagena who met the educational

standards for ordination laid down by the Pope, Gregory the Great. Isidore and the other scholars of the age were not, it seems, the last outcrop of the old system, but a new generation positively keen to salvage and restore much of the old culture that was largely unavailable in Spain. This included the study and discussion of available grammarians: as Riché said (1962: 305), "the superiority of Visigothic Latin over Merovingian can only be explained as the result of a serious study of the grammarians". In a general sense, Isidore is the most "Latinate" scholar between the Roman and Carolingian Empires, although, as Rodríguez Pantoja (1974) made clear, Fontaine and Lindsay greatly overestimated the Latinity of his orthography. For present purposes, his work is worth choosing for detailed examination.

a) The *Differentiae*

In Isidore's work there are several different kinds of evidence concerning the nature of his own speech and that of his community. His first remarks of a linguistic nature were in the *Differentiae* (PL LXXXIII 9-97), of which the first half, *De differentis verborum*, is a work in the traditional genre that included, for example, the *Orthographia* of Capar (2nd century) and the brief imitation by Agrosolus (5th century) (Keil VII 92-105). This is a list of 610 pairs, or occasionally trios, of words whose meaning needs to be differentiated. Usually the ambiguity is the result of semantic contiguity, e.g.:

313. Inter *inquirere* et *quaerere*. *Inquirimus* ea de quibus dubitamus, *quaerimus* ignota.

426. Inter *pigrum* et *torporem*. *Torpor* dormitantis est, *pigrum* vigilantis.

594. Inter *vocem* et *sonum*. *Vox* est hominis, *sonus* crepidinis.

Occasionally, however, the ambiguity is the result of phonetic indistinctiveness of two words whose meaning could not be confused, e.g.:

593. Inter *vae* et *ve*. *Ve* sine a conjunctio conjunctiva est. *Vae* cum a, interjectio dolentis est.

602. Inter *vivit* et *bibit*. *Vivit* de vita, *bibit* de potione.

If Isidore did not himself join in the reconstructable Romance Spanish pronunciations of his time, in which *vae* and *ve*, *vivit* and *bibit* could have represented the same sound sequences, it is hard to see why these distinctions needed to be made. Such examples as these, implying homonymy, are sometimes repeated in his *Origines* 127.

b) *Origines* I

The first chapter of Isidore's *Origines* is a grammar, based largely on the African commentaries on Donatus. As Fontaine showed

(1959: Part I chaps. 2, 6), this is not in arrangement a work of pedagogical intent; Isidore gives the impression that his readers already know the language under discussion. The most didactic section is the brief chapter on *orthographia* (I 27), mostly concerned with orthographical distinctions between homonyms, in the manner of the *Diffrēntie*. The fact that most of his examples are taken from elsewhere led Fontaine to conclude, sadly, that this work cannot be used as evidence for dialectal traits of Southern Spanish pronunciation. Even so, it is clear from the examples he chooses that Isidore pronounces vernacular homonyms homonymously, regardless of orthographical distinctions.

nam sicut ars tractat de partium declinatione, ita orthographia de scribendi peritia, ut puta *ad*, cum est prepositio, *d* litteram; cum est coniunctio, *t* litteram accipit. *Maad*, quando adverbium est, negandi, *d* littera terminatur et aspiratur in capite, quando autem coniunctio distinctiva est, per *t* litteram sine aspiratione scribitur. (I 27.12)

These are old precepts. *Ar* and *ad* appear in *Diffrēntie*; Isidore probably took them from Cassiodorus (*PL* 1246C-D; Keil VII 154.13-16 and *PL* 1249A; Keil VII 158.20-22; cp. Lehmann 1913), who took them from Velius Longus (Keil VII 69.20-24) and Papinianus. But the fact that the examples are taken from elsewhere cannot be taken to mean that he does not regard them as relevant, for it is his practice to use his sources constructively rather than merely copy them. His phrasing here suggests that he sees [ad] and [awd] as one word each; in the first case, if it means "to" it is spelt with a *d*, if it means "but" it is spelt with a *t*; in the second, if it means "not" it is spelt with a *d* and an *h* (*aspicitur* always refers to the letter, not the sound), if it means "or" it is spelt with a *t* and no *h*. Neither Isidore nor the scribes he is addressing seem to make a phonetic distinction here. There are five other similar examples concerning *-t* and *-d*, such as:

id pronomen neutri generis per *d* scribitur, ab eo quod est *is*, *ea*, *id*, quia facit *idem*. Quod si verbum est tertiae personae, per *t* notabitur, ab eo quod est *eo*, *is*, *it*, quia facit *itur*. (I 27.12)

Other distinctions that we can most logically assume are orthographic alone include that between *ae* and *e*. *Laetus* (I 27.14), for example, is said to need the written *a* because of an etymological connection between *laetitia* and *laetudo*. The original [ai], *ae*, and "short" [e], *e*, had been indistinguishable in vernacular for several

centuries. Concerning *-s* and *-us*, where the [n] had dropped long since in ordinary speech, Isidore regards the *n* as a "silent" letter; *formosus*, he says, "sine *n* scribitur, quia a forma vocatur" (I 27.9), taking the example from Cassiodorus (*PL* 1250A; Keil VII 160.12-15) but adding his own reason. He also takes over Cassiodorus' mention of an *m* in *lunus*, etc. (see above), but rephrases it into the past tense with *habebant* (I 27.25).

This chapter (I 27) deals with spelling alone. Isidore is not concerned to prescribe pronunciation; his audience can do that naturally in any case, it seems, and Isidore feels no need to reinforce or even mention any distinction such as that postulated by modern scholars between his own supposedly "Latin" speech and the Romance of the majority of the community. For example, Fontaine said (1959: 93) that Isidore was "combating the mistakes of a spoken language in full evolution"; but there is no evidence for this hostility other than Fontaine's assumption that Isidore must have been hostile. Such evidence as there is is to the contrary. For example:

nam cum *lusitana* sonum *z* littera exprimat, tamen, quia Latinum est, per *t* scribendum est. Sic *militia*, *meditia*, *neguitia*, et cetera similia. (I 27.28)

Z was a Greek written letter, representing [ts] or [dz]; early Romance developed this sound in such words as *malitia* (> Spanish *maldad*), although original Imperial Latin had no such affricates, and the letter *z* was quite often used. Isidore disapproves of that spelling, not because it represents an unacceptable sound — he concedes that it is indeed a 'z' sound — but on the historical grounds that these are not Greek words. The example of *lusitana*, but not this reasoning, probably comes from Q. Papirius (Keil VII 216.8-14). Isidore is later to use the *z* himself in a famous passage (e.g. Fontaine 1959: 92 n.3; Kramer 1976: 72): "solent ita dicere *ozie* pro hodie" (XX 9.4). This is reconstructably simple Proto-Romance; Italian had affricated [-dʒ-] (e.g. *hodie* > Modern *oggi*) and Spanish had not (*hodie* > *hoy*), whereas both affricated [-tʃ-] as in *MALITIA*. It also seems certain that he used the prothetic [e-] on words beginning with [s] plus a consonant, for he later gives etymologies for *scarra* ("jester" or "parasite") and *scarrus* (a kind of sea-fish) based on *esca* ("bait", or simply "food"). The first of these is even catalogued alphabetically under the letter *i*:

iscurra vocatur quia causa escae quempiam consecretur (X 152) *escarrus* dicitur eo quod solus escam ruminare perhibetur (XII 6.30)

It would be hard to find clearer indications that Isidore did not himself use a systematically archaic pronunciation norm. He hardly

3. Jiménez (1961: 484), shows that the scribe of the León MS spells *is* *istis* here anyway.

mentions phonetics; Fontaine was right to point out that "unlike an authentic grammarian such as Varro, Isidore hardly bothers with methodical or objective comments on phonetics itself" (1959: 112). This can only be because it is of no great interest to him. It is fair to conjecture that had there been a systematic duality of phonetic norms, his voracious intellectual curiosity would probably have led him to mention it in considerable detail.

In I 32, Isidore discusses *Barbarismus*.

Barbarismus est verbum corrupta littera vel sono enuntiaturum. Littera, ut *floriet*, dum *florēbit* dicere oporteat; sono, si pro media syllaba prima producat, ut *latebrae*, *tenebrae*. Appellatur autem *barbarismus* a barbaris gentibus, dum latinae orationis integritatem nesciunt. (I 32.1)

This is substratum theory. Barbarisms are mistakes committed by non-native Romance speakers when learning to speak *latina oratio*; although it might seem at first sight that Isidore is defending "correct Latin" against Romance, he is not; for *latina* is never opposed to "Romance" in Spain and the features criticized are not Proto-Romance at all. The particular examples of I 32.1 are not particularly illuminating. *Floriet* for *florēbit* is, to the eyes of a modern linguist, a morphological rather than a pronunciation "mistake" – using the normal -IRE future form, instead of the -ERE, for *FLORĒRE*. The Modern Italian is *florire*; yet the later Spanish is not *florir* but *florecer*, and the Proto-Romance future of the seventh century was in any event formed from the infinitive and HABEO (e.g. DARE HABES > Sp. *darás*), so the castigated form is not likely to have been a native Romance one. The mistaken conjugation of FLOREO is thus the opposite of the contemporary liturgical use of CAPĒRE and FUGIRE mentioned above, which are both acceptable and Old Romance. *Latebrae* and *tenebrae* are examples taken from Donatus. In Imperial Latin these words apparently had alternative pronunciations: *tēnēbrae* ~ *tēnēbrae*, *lātēbrae* ~ *lātēbrae*. The source of Spanish *trinebras* is presumably the non-Classical *TENĒBRAE, since the [je] diphthong arises from a stressed "short" [e]. *Producitur* is probably used here to refer to stress; the use of *pro*, "instead of," suggests this, in that words can have only one stress, but any number of long vowels. If so, the barbarism is unacceptable in seventh-century Spanish, but would not have been in Imperial Latin (*tēnēbrae*). If Isidore is referring to length here, [te:n-] would indeed have been a barbarism in any kind of Latin; but Isidore tends to use for "short" and "long" *brevis* and *longus*, as in the following section, which reflects traditional preoccupations:

Pronuntiatio autem fit in temporibus, tonis, aspirationibus et reliquis quae sequuntur. Per tempora quippe fit barbarismus, si pro longa syllaba brevis ponatur, aut pro brevi longa. Per tonos, si accentus in alia syllaba commutetur. Per aspirationem, si adiciatur *h* littera ubi non debet aut detrahatur ubi esse oportet. Per hiatum, quotiens in pronuntiatione scinditur versus antequam compleatur, sive quotiens vocalis vocalem sequitur, ut *Musae Aonides*. Fit barbarismus et per motacismos, [motacismos] et labdacismos. Motacismus est, quotiens *m* litteram vocalis sequitur, ut *bonum aurum*, *iustum amicum*; sed hoc vitium aut suspensione *m* litterae, aut detractioe vitamus. Iotacismus est, quotiens in iota littera duplicatur sonus ut *Troia, Maia*, ubi earum litterarum adeo exilis erit pronuntiatio, ut unum iota, non duo sonare videantur. Labdacismus est, si pro una *l* duo pronuntientur, ut *Afri faciunt*, sicut *colloquium pro colloquim*; vel quotiens unam *l* exilius, duo largius proferimus. (I 32.3-8)

The *barbarismi* include the use of hiatus rather than elision in, e.g. *Musae Aonides* – a "barbarism" apparently committed in contemporary hymns, but, as we know, the hymn-singing style can legitimate avoidance of such surface constraints –, and the pronunciation of [m] in e.g. *bonum aurum*, neither of which can have been normal seventh-century Spanish usage; the reinforcement of the [j] sound in e.g. *Maia* is incorrect old Spanish (but correct Old Tuscan: *MAIUM* > Sp. *mayo*, *MAIUS* > Ital. *maggio*), and the pronunciation of a double [ll] in e.g. *colloquim*, "as Africans do," is not correct old Spanish either (CONLOQUUM > *coloquio*); so the barbarisms castigated are all, in fact, reconstructably unacceptable as attempts at the vernacular. In a later summary (I 34.2) Isidore's example of *barbarismus* is "ut si tertiam syllabam quis producat in *ignoscere*"; this may refer to either stress-shift or vowel-closing or both (IGNOSCĒRE > IGNOSCĒRE), and is in modern eyes another morphological reanalysis rather than a phonetic change. This might be Proto-Romance; IGNOSCERE does not survive, but this shift is reconstructable for COGNOSCERE (> *conocer*). Apart from this morphological example, Isidore does not castigate Old Romance as the wrong pronunciation. Had there been two norms, surely Isidore of Seville, if anyone, would have used the "Latin" and mentioned evolved *vīta* as to be avoided; but the evidence is that he did not, and that he spoke in vernacular pronunciation. Fontaine is surely wrong to suggest that Isidore was "combating" the vernacular speech of his community.

c) De Lectoribus

Isidore's *De Ecclesiasticis Officiis* II 11 (PL LXXXIII: 791-92), entitled *De lectoribus*, deals with the proper manner of reading aloud the word of God. He mentions in detail the care for proper audibility,

accurate intonation, appropriate pauses, as Cassiodorus had before him; there is no hint at all that a special non-vernacular pronunciation is involved here, as it was later in Carolingian times. Nothing phonetic is thought relevant. The chapter is here reproduced in full:

CAPUT XI De lectoribus

1. Lectorum ordo formam et initium a prophetis accepit. Sunt igitur lectores qui verbum Dei praedicant, quibus dicitur: *Clama, ne cesses, quasi tuba exalta vocem tuam (Isai. LVIII)*. Isti quippe, dum ordinantur, primum de eorum conversatione episcopus verbum facit ad populum. Deinde coram plebe tradit eis Codicem apicum divinarum ad Dei verbum annuntiandum.
2. Qui autem ad huiusmodi provehitur gradum, iste erit doctrina et libris imbutus, sensumque ac verborum scientia perornatus, ita ut in distinctionibus sententiarum intelligat ubi finitur iunctura, ubi adhuc pendet oratio, ubi sententia extrema claudatur. Sique expeditus vim pronunciationis tenebit, ut ad intellectum omnium mentes sensusque promoveat, discernendo genera pronunciationum, atque exprimendo sententiarum proprius affectus, modo indicantis voce, modo dolentis, modo increpantis, modo exhortantis, sive his similia secundum genera propriae pronunciationis.
3. In quo maxime illa ambigua sententiarum adhibenda cognitio est. Multa enim sunt in Scripturis, quae nisi proprio modo pronuntientur, in contrariam recidunt sententiam, sicut est: *Quis accusabit adversus electos Dei? Deus, qui iustificat (Rom. VIII 33, 34)?* Quod si quasi confirmative, non servato genere pronuntiationis suae, dicatur, magna perversitas oritur. Sic ergo pronuntiandum est, ac si diceret: *Deusne qui iustificat? ut subaudiatur non.*
4. Necesse est ergo in tantis rebus scientiae ingenium, quo proprie singula, convenienterque pronuntientur. Propterea et accentuum vim oportet scire lectorem, ut noverit, in qua syllaba vox protendatur pronuntiatis. Plenumque enim imperiti lectores in verborum accentibus errant, et solent irridere nos imperitiae hi qui videntur habere notitiam, detrahentes, et iurantes penitus nescire quod dignus.
5. Porro vox lectoris simplex erit, et clara, et ad omne pronuntiationis genus accommodata, plena succo virili, agrestem, et subrusticum effugiens sonum, non humilis, nec adeo sublimis, non fracta, vel tenera, nihilque femininum sonans, neque cum motu corporis, sed tantummodo cum gravitatis specie. Auribus enim et cordi consulere debet lector, non oculis, ne potius ex seipso spectatores magis quam auditores faciat. Vetus opinio est lectores pronuntiandi causa praecipuum curam vocis habuisse, ut exaudiri in tumultu possent. Unde et dudum lectores praecones vel proclamatores vocabantur.

Section 4 deals with the need for accurate stress-positioning on unfamiliar vocabulary, but does not discuss the phonetics. (*Subrusticus* is "boorish" or "fivolous".) Neither this chapter, nor Chapter 10 *De lectoribus*, nor any other chapter of the *De Ecclesiasticis Officiis*, mentions or hints at any need to use an official archaic pronunciation rather than vernacular. Isidore was the most educated scholar of his age, keen to preserve the old culture, but his remarks give little weight to the modern assumption that the church of his time is supposed to have preserved archaic speech.

d) Isidore's *Comments on Linguistic Variety*

This reticence on phonetic matters is not the result of linguistic insensitivity on Isidore's part. On the contrary, he has many comments to make in the *Origines* concerning varieties of speech (cp. Sotér 1930). For example, Isidore is sometimes concerned to distinguish present-day usage from that of the past. The relevant word used here is often *veteres*: e.g.

Erat enim apud veteres hoc signum meretriciae vestis, nunc in Hispania honestatis
(XIX 25,5)

Porus . . . hunc veteres a baloandis mercibus paucis vocabant.
(XIV 8,40)

Veteres take a past tense. So, usually, do *Romani*:
Ideo autem Romani aquam et ignem interdicebant quibusdam
damnatis.
(V 27,38)

On one occasion, *Romani* take a present tense, but since it is exemplified by a quotation from Horace (*Carmen* III 18,1) this present can only be a historic present (VIII 11,103-04). *Romani* and *veteres* have usages that do not survive. The words *Latini*, *Latine* and *Latina lingua* are usually used with much wider reference, contrasting Latin-Romance to other languages entirely. (Occasionally Isidore takes over another scholar's reference to *Latina* as the speech of Latium; see below.) This was evident in the already quoted comment on the causes of *barbarismus* (see above). The other language to which *Latina* is contrasted is often Greek:

Plinios, quam Latini heretam lucernam vocant
(XVII 9,73)

Strychnos, quae Latine herba salutaris vocatur
(XVII 9,78)

Chronica . . . Hieronymus presbyter in Latinam linguam convertit.
Xocho enim Graece, Latine tempus interpretatur.
(V 28,1)

These words continued to have the wide meaning, of the Latin-Romance tongue as opposed to other languages entirely, in Moslem Spain, where *Latina* is often contrasted specifically to Arabic, and means the vernacular old Spanish (see Chapter 4). Christians in Moslem Spain were educated in the Visigothic tradition, and this meaning for *latina* is

part of it. Thus it is that, within the *Origines*, the *vulgus* can be said to speak *Latine*:

Framea ... quam vulgo spatam vocant ... alii spatam Latine autumant: dictam eo quod spatiosa sit ... (XVIII 6.3.4)

Within the speech of his own time, Isidore is aware of national differences of vocabulary.

Totes Gallica lingua dicuntur, quas vulgo per diminutionem tusillas vocant (XI 1.57)

Tucos, quos Hispani ciculos vocant (XII 7.67)

Unde et eos Hispani et Galli tucanos vocant. (XVIII 7.7)

His only apparent comments on contemporary geographical variation in phonetics are the previously mentioned remarks on African *colloquium* (I 32.8) and Italian *ozie* (XX 9.4); otherwise, the examples are of vocabulary alone. Within Spain, he mentions the particular usages of his native province:

Actus quadratus ... hunc Baetici arapennem dicunt (XV 15.4)

Actum provinciae Baeticae rustici acnam vocant. (XV 15.5)

On other occasions, Isidore discusses variation without being specific about who chooses which words; *multi* say this, *alii* that, *quidam* the other:

Molochinia ... quam alii molocinam, alii malvellam vocant. (XIX 22.12)

Vulgus is a key term for present purposes. Isidore defines *vulgus*: "vulgus est passim inhabitans multitudo, quasi quisque quo vult" (IX 4.6). This seems clear enough: "everyone and anyone". The usage of the *vulgus* is often adduced; with one exception it is given a present tense. The exception is based on a remark in Pliny's *Natural History* XXXIII 12.40, *virtolae Celtice dicuntur* ("bracelets"), which is by Isidore's time no longer normal: "Armillae ... Unde et quondam vulgo virtolae dicebantur" (XIX 31.16). Such uses with a present synthetic passive are usually taken from a source.

If *vulgus* is the subject, it often takes a plural verb:

Palumbes ... quas vulgus citos vocant. (XII 7.62)

Occasionally it takes a singular verb:

Regiones ... quas vulgus conventus vocat. (XIV 5.21)

And occasionally *vulgo* appears adverbially with a general third person plural:

Abactor ... quem vulgo abigeium vocant. (X 14)

In all grammatical circumstances, *vulgus* is most naturally taken to mean "everyone here and now", referring to the normal usage of Isidore's time and place; it is not used to contrast any "Vulgar Latin"

with anything else contemporary. Isidore uses *vulgus* vocabulary himself, without apparent self-consciousness: e.g. the words *lunaticos* and *burgos* are said to be used by the *vulgus*, but are also used by Isidore elsewhere:

Epilemsia ... hos etiam vulgus lunaticos vocant (IV 7.5-6)

Caducus a cadendo dictus. Idem et lunaticus, eo quod certo lunae tempore patiatur (X 61)

Crebra per limites habitacula constituta burgos vulgo vocant (IX 2.99; also IX 4.28)

Burdigalim appellatam ferunt quod Burgos Gallos primum colonos habuerit. (XV 1.64)

The following passage is particularly indicative:

Solidum nuncupatum, quia nihil illi deesse videtur; solidum enim veteres integrum dicebant ... hunc, ut diximus, vulgus aureum solidum vocant. (XVI 25.14)

In this case, a previous remark without the word *vulgus* ("solidum nuncupatum") is taken up again ("ut diximus, vulgus aureum solidum vocant") in such a way as to imply that the unspecified agent of such a generic passive as *nuncupatum* is understood to be the *vulgus*.

Isidore's use of the first person plural suggests that he too is happy to use terms the *vulgus* use. The following remarks come in successive sections:

Poderis est sacerdotalis linea ... quam vulgo camisiam vocant (XIX 21.1)

Camisias vocari quod in his dormimus in carnis, id est in stratis nostris. (XIX 22.29)

There is often no obvious difference in meaning or tone between Isidore's use of *vulgus* and of the first person plural:

Lactuca agrestis est quam serralliam nominamus, quod dorsum eius in modum serrae est (SERRALLIAM became Spanish *carraja*, "sow-thistle"). (XVII 10.11)

Viticella herba a Latinis appellata quod ... apprehendat corymbis, quos anulos appellamus. (XVII 9.92)

The majority of cases in which the *vulgus* are specified involve a lexical usage not normally attested in the old texts that Isidore has inherited. The number of occasions in which the *vulgus* form is said to have unacceptable pronunciation is very few; on all such occasions the word *vulgus* is accompanied by either the word *imprudens* or the word *corrupte*, and it is simple to see that the undesirable phonetic element is referred to with the *imprudens* or the *corrupte* rather than with the *vulgus*:

Corus ... quem plerique Argesten dicunt, non ut imprudens vulgus Agrestem (XIII 11.10)

Rhododendron, quod corrumpit vulgo Iorandrum vocatur

(XVII 7.54)

(or, according to other manuscripts, *Rodandrum*, *rodandrus*, *Iorandrum*, *Iaurandus*, *Iaurandrum*; Sofer 1930: 99)

Citocacia vocata quod ventrem cito depurgat; quam vulgus corrumpit citocociam vocant

(XVII 9.65)

(or, according to other manuscripts, *vitocociam*, *citocam*, or *citocacium*; Sofer 1930: 53)

Sagina, quae corrumpit vulgo salina dicitur.

(XX 16.5)

Such imprecision and corruption can occur without it being specifically attributed to the *vulgus*:

Gubellum corrumpit a globo dictum per diminutionem, quasi globellum.

(XIX 29.6)

On two occasions Isidore allies his own usage with the *corrumpit* form:

Phaselus est navigium quem nos corrumpit basellum dicitur

(XIX 1.17)

Propina Graeco sermo est, quae apud nos corrumpit popina dicitur.

(XV 2.42)

On the other hand, there are no occasions on which a *vulgo* usage, without *corrumpit* or *imprudens*, is contrasted with a phonetic variant of the same word rather than with a different lexical item. The following is a morphological variant, not a mispronunciation:

Capitulum est quod vulgo capitulare dicitur.

(XIX 31.3)

In general, however, Isidore is only concerned to mention variations in items of vocabulary (cp. Rabanal 1970: 189). Items that are ascribed to the *vulgus* seem to be unashamedly part of his own repertoire, and such ascription in itself is not intended to imply anything at all of a phonetic nature. The words of the *vulgus* are regularly said to be so because they are seventh-century usage, and not authoritatively attested ancient forms (see Sofer 1937); these forms of seventh-century Spanish are thus indeed contrasted to the "Classical Latin", but to the Latin of the distant past rather than to the usage of a contemporary learned elite.

In IX 1 Isidore briefly discusses languages. His entire comment on Latin is as follows:

Latinas autem linguas quattuor esse quidam dixerunt, id est Priscam, Latinam, Romanam, Mixtam. Prisca est, quam vetustissimi Italiae sub Iano et Saturno sunt usi, incondita, ut se habent carmina Saliorum. Latina, quam sub Latino et regibus Tusci et ceteri in Latino sunt locuti, ex qua fuerunt duodecim tabulae scriptae. Romana, quae post reges exactos a populo Romano coepit, quae Naevius, Plautus, Ennius, Vergilius poetae, et ex oratoribus Gracchus et Cato et Cicero vel ceteri effluerunt. Mixta, quae post imperium latinus promiscuum simul cum moribus

et hominibus in Romanam civitatem inrupit, integritatem verbi per soloechos et barbarismos corrumpens.

(IX 1.6-7)

There are four divisions of "Latin language", viewed as a temporal progression. *Prisca* is the pre-historical. *Latina*, used in the narrow sense, is that of Latinum. *Romana* comes after the expulsion of the Kings. Finally, *mixta* dates from the days of the empire, and a more cosmopolitan speech community. Fortunately, Isidore has just previously defined the word *mixta* for us in the context of Greek, where it is the translation of *κοινή*: "*κοινή*, id est mixta, sive communis quae omnes utuntur", a common language used by all, opposed here neither to high-class language, nor to archaic language, but to the regional Greek forms, Attic, Doric, Ionic and Aeolic (IX 1.5). As regards Isidore's comments on Latin, this cannot be interpreted to mean that Isidore was so keen to avoid regional usage that he used a distinct Latinate *κοινή*; nor can it be interpreted as Maurer interprets it (1962: 96 n.143), apparently unaware of the previously explained Greek model, to mean that *mixta* is "Vulgar Latin" coexisting with a conceptually separate "Classical Latin" from imperial times; it is clear that all regional varieties come under the one general heading of the one common language in use for at least five centuries, called *Latina mixta* precisely because it incorporates diversity and variety within it, and specifically stated to be later than the other kinds of Latin, rather than still coexisting with anything more respectable. Isidore himself certainly regards himself as speaking the language of the community that he is in: Cum autem omnium linguarum scientia difficilis sit cuiquam, nemo tamen tam desidiosus est ut in sua gente positus suae gentis linguam nesciat. (IX 1.10)

"No one is so feeble that he cannot speak the language of his own community." It is surely preferable to conclude that Isidore also speaks and uses the language of his time and place.

e) *Vulgus* and *Rustici*

The reconstructionists cannot claim to reconstruct most of the vocabulary of seventh-century Spanish, on the basis of thirteenth-century Spanish, as accurately as they claim to be able to reconstruct the phonology in its essential features. Vocabulary comes and goes. Often, the words said to belong to the *vulgus* in the *Origines* can be postulated as probably existing in the Romance of seventh-century Spain, on the basis of the later developed form's existence: e.g.

COLOMELLOS	> Sp. <i>colmillos</i>	(XI 1.52, "lucks")
IRICUM	> Sp. <i>erizo</i>	(XII 6.57, "hedgehog")
FURCA	> Sp. <i>horca</i>	(V 27.34, "gallows")

- SARNAM > Sp. *sarna* (IV 8.6, "impetigo")
 SUILLOS > Sp. *sillos* (XII 6.12, "sturgeon")

But other words said to belong to the *vulgar* seem subsequently to have died out, as vocabulary often does. The following have no Romance descendants (according to Meyer-Lübke 1935, at least): *RASILUS* (XIX 22.23), *REPTOS* (XIX 23.4), *SCOTICA* (V 27.15), *STINCUM* (XVII 9.43), etc. But in addition, as Sofer stressed, there are many non-Classical words used by Isidore himself, without being specifically attributed to the *vulgar*, which a reconstructionist would like to postulate as existing in the ordinary vernacular of his time. This causes neither surprise nor problems if we can admit that Isidore is writing his own seventh-century vernacular anyway. For example, the following words are the etyma for subsequent Spanish developments:

- CATENATUM > Sp. *candado* (XX 13.4, "lock")
 INCINCTA > Sp. *encinta* (X 151, "pregnant")
 MERENDA > Sp. *merienda* (XX 1.12, "afternoon snack")
 PROSTRARE > Sp. *postrar* (XVIII 42.2, XVIII 56, "prostrate")
 TULUS > Sp. *tío* (IX 6.15, "uncle")

There are several conclusions to be drawn from this. One concerns the practice of etymological dictionaries. If Isidore and the other Visigothic scholars are writing their own language, and that is the language of the *vulgar*, then the earliest attestation of a word in vernacular can be attributed to these writers. Corominas' etymological dictionary (1954-57) occasionally mentions Isidore (e.g. in his article on *tío*), but eventually it can be the only sensible course to admit that words consistently used in the vernacular were consistently used in the vernacular, rather than pretending that they suddenly spring to life at exactly the same time as the spelling reforms that make them intelligible, or at least recognizable, to a modern Spaniard. (Some studies, such as that on *jerigonza* by Moralejo 1978, have in fact recently used Visigothic evidence.) A second conclusion might be a mild caveat against trusting the reconstructionist techniques too implicitly; they are at best partial. It can be seen that some of the vocabulary attributed by Isidore to the *vulgar* did not survive until the thirteenth century; therefore it looks as if that section of the vocabulary which can be reconstructed for seventh-century vernacular, on the basis of its earlier imperial origin and of its existence in an evolved form in the thirteenth century, is only a part of the whole. It may thus be that the reconstructed very old Spanish phonology as envisaged by R.A. Hall, and others, is only partial; phonetic features can arise and then disappear, changes can begin and then reverse without trace, never to be reflected

on paper. The reconstructed phonology is too neat and tidy to be realistic: it is at best skeletal.

Rustici in Isidore means "country people": "Rusticus dictus quod rus operetur, id est terram", X 239. Isidore mentions "rustic" nomenclature for wine terms (XVII 5.7 and 9), sticks (XX 13.2 and V 27.16), trees (XVII 7.66) and plants (XVII 11.9). The word regularly has a non-linguistic application, as in the titles of book XVII, *De rebus rusticis* (On Agriculture), section XV.12, *De aedificiis rusticis*, and section XX 14, *De instrumentis rusticis*. *Rustica* seems to have become a semi-technical linguistic term for "uneducated speech" (as contrasted with the new Church Latin) two centuries later and several hundred miles to the north-east, in the Carolingian scholarly circle, but this fact is no more relevant to the study of Isidore than is the specialized use of a word by Goethe to the study of a word with the same etymological root in Shakespeare. (*Hortulani* are similarly invoked for the terminology of irrigation in XX 15.3.) The post-Carolingian use of *rusticus* to mean "non-Latinate" unreformed natural vernacular can be seen, for example, in the famous mention of *rustica romana lingua* in 813 (see Chapter 3). The semantic ground is prepared for this development by such previous comments as those of Gregory of Tours (*MGH Scriptorum* I 1.31.14):

Philosophantem rhetorem intellegunt pauci, loquentem rusticum multi

(in his preface to the *Historia Francorum* — referring to his own style); in this the *rusticum* is still literally a countryman, but his talk is the object of interest, or the comment in Marculf's preface to his formulary (*MGH Legum* V 37.15):

Iuxta simplicitate et rusticitate meae naturae . . . ita scripsi which refers to writing. Here the usage is metaphorical, "unpolishedness", as in Quintilian's contrast of *rusticitas* with *urbanitas* (*Inst. Or.* VII 3.17), for a real *rusticus* would not write at all. With *rusticus*, a change of meaning in time and place can be seen, from "rustic-boorish" to "Romance", and it is over-simple to talk of the meaning of *rusticus* in Medieval Latin, tout court.

Vulgar, on the other hand, means "everyone at the time and place concerned". It is all the more worth avoiding the pejorative and contrastive implications of the term "Vulgar Latin" when the word *vulgar* itself had no such implications.

Julian of Toledo

Isidore died in 636. During the rest of the century the educational systems and habits that he had helped to spread became institutionalized, and the educational level of at least the ecclesiastical elite remained high. There is a considerable corpus of linguistically competent metric verse, church councils, liturgical texts and historical material that survives. It used to be thought, on no evidence (see Collins 1977), that the intellectual level of the community was already in considerable decline at the time of the Moslem invasion of 711; political instability can be accompanied by intellectual advance, however, and the atmosphere of knowledge and intellectual life established in the early seventh century continues with sufficient impetus and vigour for Toledo, Córdoba, Catalonia, the Rioja and the Asturian court to carry forward a recognizably Visigothic cultural tradition in the decades following the Moslem invasion. In addition, Visigothic texts and scholars are of considerable influence in the eighth-century cultural expansion of both the British Isles and France.

Of Isidore's Spanish successors, only Julian of Toledo seems to have been involved in the composition of an *Ars Grammatica* (c.585; *Maestre Yenes* 1973; not in Hilgarth 1976). Beeson stated that "in writing a school book on grammar, metric and rhetoric, Julian is simply following in the footsteps of the compiler of the *Origines*" (1924: 50). This is inaccurate. In the first place, it seems likely that the written compilation was made from lecture-notes by a pupil or pupils of Julian, although *Maestre Yenes* graciously permits us to call it Julian's *Ars* even so (1973: XVIII n.23); more importantly, this work is quite unlike the *Origines*. It is in the catechetical question-and-answer mode of the traditional teaching manual, lacking the philosophical and cultural air adopted by Isidore; although Julian takes material from Isidore, its form is based on Donatus' *Ars Minor* and on his African commentators. They are operating in different contexts. Isidore was working, initially at least, in an age when schools were not common. Either he, or possibly his contemporary Juan de Biciaro, trained Braulio de Zaragoza, whom Díaz y Díaz called "the best informed man in Visigothic Spain" (1976a: 32). Braulio in turn was able to impart sufficient knowledge of quantitative metrics to Eugenio de Toledo, the archbishop whose poems are the most technically competent of the time; and Eugenio taught Julian. Julian was thus learning and then teaching within a semi-institutionalized atmosphere that had established a continuity over several decades. By his time there were several educational centres functioning in the peninsula, usually in the form of an eminent scholar

accompanied by several students. Isidore's *Origines* would not have been at all wieldy as a practical textbook for linguistic instruction. Fontaine (1959: 192-94) suggested that Pompeius' commentary on Donatus was the one Isidore himself learnt from, but even if that is so, there seems to have been a lack of a practical textbook felt by Julian in his own language classes at Toledo's episcopal school. For example, *Maestre's* study of his chapter on punctuation (1973: III-LV) led her to see it as the result of Julian's concern for "live and direct teaching."

Julian's technique seems to have been normally to take the text of Donatus, with commentaries, and the remarks of Isidore if applicable, and then add something of his own, such as examples from Christian poetry (Beeson 1924: 53-55). The only sections not based on Donatus are the metrical chapters at the end, based largely on Mallius Theodorus, and the immediately preceding sections concerning the length of final syllables, which may descend from Marius Victorinus (*Maestre* 1973: XXXII) and are designed to aid the proper organization of the metric verse. At the least this shows that people could not be expected to know the length difference without being told. The rationale of the distinction is not presented in phonetic terms at all: e.g.

Verbum quod a terminatur, longum est, aut breve? In omnibus
modis, numeris, personis, temporibus, coniugationibus, si a fuerit
terminatus longus est, ut *ama* (II V 1)

In the main body of the grammar, Julian shows no interest in instilling a "correct" pronunciation into the pupils at the Toledo school; he makes even less mention of phonetics than Pompeius did. Here, for example, are the opening comments on vowels:

Quot sunt litterae vocales? quinque.
Quae? a, e, i, o, u.
Semper vocales sunt? a, e et o semper vocales sunt, i et u varias
habent conexiones.

Quomodo? quia modo vocales sunt, modo transeunt in consonantium potestatem, modo inter se geminantur, modo cum aliis vocalibus iunguntur, modo mediae sunt, modo u inter q et aliquam vocalem posita nec vocalis, nec consonans habetur, modo u per digrammon adscribitur quando sibi ipsa praepositur, modo i inter duas vocales posita in unam partem orationis pro diabus consonantibus accipitur.

(Cp. Donatus, Kell IV 367.12-20; *Sergius*, Kell IV 520.31 - 522.12; Pompeius, Kell V 102.19 - 103.6.)

This is a description of the uses of the written letters i and u. *Geminantur*, for example, refers not to phonetic gemination but to such consonant plus vowel sequences as *deicit*, *revulsam*. The sections that follow also prescribe traditional orthographic practice. There is, in

short, no evidence that Julian is concerned to train people to utter "Latin" rather than vernacular pronunciation, nor that he did that himself.

Julian is well educated. Other scholars of his age are also well educated. Valerio del Bierzo, for example, is a competent and prolific scholar; Aherne (1949: 34) made it clear that the stylistic achievements of his work are "evidence of no little skill in their author, and argue a rather elaborate education" (cp. the discussion of Visigothic prose style in Gerwin 1946). Valerio shares with his contemporaries a delight in prose rhyme and rhythmic cursus. Julian regarded rhythmic verse as undignified: "rhythmis uti, quod plebeis est solitum ex toto refugiat" (see Bischoff 1959); in his grammar, Julian declared: "Quid est rhythmus? Verborum modulata compositio, non metrica ratione sed in numero ad iudicium aurium examinata, ut puta velut sunt carmina vulgarium poetarum ... rhythmus modulatio sine ratione" (II XX 4.5; further discussed in chapter 4; Bischoff 1959: 254). The written skills in poetry and prose were taught and learnt, as we can see from the surviving grammars and commentaries. But the grammars and commentaries used in Visigothic Spain have nothing to say about learned and erudite pronunciation. There is no support here for any belief that the Visigothic scholars used a Latinate non-vernacular pronunciation, and good reason to presume that they did not.

Bede and the Insular Tradition

Julian's grammar has not survived in any Spanish manuscript. It is still thought probable that Aldhelm used Julian in a letter to the King of Northumbria (c.700); this is plausible, since there is evidence that Aldhelm was aware of Visigothic scholarship (Winterbottom 1977). Bede (673-735) uses Julian as a source in his *De Schematibus et Tropis* (Beeson 1924: 56; Maestre Yenes 1973: C); both Isidore and Julian seem to have been taken on board in the British Isles, and of both authors the earliest manuscripts are in an insular hand. In general, Visigothic scholarship was part of the intellectual equipment of the Insular scholars (Hilgarth 1958, 1962).

The Irish and Anglo-Saxon monks wanted and needed works that could help them learn and teach the use of the normal working language of their church. They were working in quite a different context from that of Donatus, his commentators, Cassiodorus, Isidore or Julian; these latter could assume that their audience were native speakers of some variety of Late Latin/Early Romance, but Aldhelm, Bede, Boniface and Tatwine (etc.) were hoping to train native speakers

of some variety of Germanic. Latin-Romance did not last into the sixth century in the islands as a spoken language, and subsequently everything had to be taught from scratch.

There are Irish Latin texts that pre-date the Gothic connection, but they strike Latinists as being very peculiar (e.g. Riche 1962: 357-59; Löfstedt 1965). The glee with which the Irish scholars fell on grammatical works can be explained by the previously precarious domination that they had had of this language that was so very different from their own. Hilgarth established that the route from Spain to Northumbria proceeded via Ireland, as early as the seventh century (1962: 174), carrying Biblical commentaries, liturgical books and grammatical works.

As Blair pointed out (1970: 244), "When Bede made his first approach to Latin, he did so as to a wholly foreign language, unlike his contemporaries in the schools of the more southerly countries where Latin still remained the living spoken tongue, and perhaps it was this distinction which led not only Bede himself, when he came to be the teacher, but also other English scholars, to compile their own grammatical treatises" (cp. Riche 1962: 436-37; Roger 1905: chaps. 8-11). No existing textbook had been elaborated with non-native speakers in mind. Pronunciation was not given an important rôle in the Anglo-Saxon treatises either, however; Tatwine's *Ars* helps the foreigner with its practical inclusion of much vocabulary, but has nothing on pronunciation (De Marco 1968). Even so, there can be no doubt that the learning of Latin in eighth-century England did involve reading it aloud, or reciting from memory, even if it did not involve spontaneous conversations in the language. The Bible and the works of the Holy Fathers were read aloud regularly in monastic communities. Ceolfrith, Bede's teacher, recited the Psalter twice every day (Blair 1970: 255).

Ceolfrith had been to Rome. Benedict Biscop went to Rome between 671 and 684, returning with some manuscripts, including a Vulgate. There was an English community in Rome in the eighth century, and at least one Italian came to England to teach church music. This seems to have involved the cantor singing from a text and the others learning it by heart. The evidence is so thin that we cannot tell whether the pronunciation of the sung words was simple eighth-century Roman Italian or not, but in church song, as we have seen, the music had (and has) a distorting effect on phonetics. In general, though, it could be that the pronunciation taught to at least some of the Saxons was of an Italianate tinge. This cannot be taken to be the norm, however. Alcuin was happy, for example, to continue the native insular

tradition of composing alliterative verse involving alliteration of *ce-* and *ce-* (e.g. *care* and *cero*; Norberg 1958: 52). Bede alliterated *celis* with *coritas*. There is the alliterative Irish chant which includes the phrase *et celis celi celi culmina*. The Irish never associated written Latin *i* and *e*, nor *o* and *u*, as was normal practice in Romance areas. Norberg also mentions the insular ability to alliterate *f* and *v* (1958: 51), as in *famine* and *versus*, *verba* and *fudit*; no Romance speech has ever merged these two sounds in initial position. This confusion is a feature of German Latin for some time, and has one echo in Alcuin's *De Orthographia*:

Vel si coniunctio est, per v; si humorem significat, per f; si idolum, per b scribendum est. (PL CI 918C)

If there was some Italianate influence in Anglo-Saxon Latin pronunciation, it did not get very ingrained. In general it seems that in order to teach Latin pronunciation to Anglo-Saxon students, a rule of thumb was based, *faute de mieux*, on some kind of correspondence between written letters and subsequent spoken sounds. No native language is ever learnt in this manner; in the Romance communities everyone learnt automatically to speak their local vernacular when young, and then learnt to spell the old-fashioned 'correct' way several years later, if at all. The reverse applied in England, where the pronounced form was taught as a function of the already learnt traditional spelling, regardless of the habits of faraway native speakers. (Occasionally people learn English in this manner nowadays, pronouncing *would* as [woud], *your* as [jowr], etc. They are quite unintelligible, even if fluent.) As a consequence, although Isidore pronounced *haud* and *aut*, *at* and *ad*, etc., as homonyms, Bede probably did not.

The only text on pronunciation which circulated in Britain in the eighth century was the following passage from Martinus Capella (III 261; Dick 1969: 95.14–96.16), which was copied and reprinted independently of the *De Nuptiis* itself, and attributed to various authors (Lew 1982):

Namque A sub hiatu oris congruo solo spiritu memoramus.
B labris per spiritus imperium reclusis edicimus.
C molaribus super linguae extrema appulis exprimitur.
D apulus linguae circa superiores dentes inasculitur.
E spiritus facit lingua paululum pressore.
F dentes labrum inferius deprimentes.
G spiritus cum palato.
H contractis paululum faucibus ventus exhalat.
I spiritus prope dentes pressis.
K faucibus palatoque formatur.
L lingua palatoque dulcescit.

M labris imprimitur.
N lingua canticus appulsa collidit.
O rotundi oris spiritu comparatur.
P labris spiritus erumpit.
Q appulsa palati ore restricto.
R spiritum lingua crispante corrigitur.
S sibilum facit dentibus verberatis.
T appulsa linguae dentibusque impulsis excutitur.
V ore restricto labrisque prominulis exhibetur.
X quicquid C atque S formam exhibat.
Y appressis labris spiritu procedit.
Z vero iacito Appius Claudius delectatur, quod dentes morui, dum exprimitur, imitatur.

This list could have been used as a teaching aid for novice Latins; something similar probably underlies Alcuin's insistence as eventually propounded for Carolingian students. Without accompanying demonstrations these prescriptions would have been insufficient to specify the exact sound.

Bede's own *De Orthographia* (Jones and King 1975: 7.57), unlike earlier works of the same name, deals with much more than spelling. It used to be regarded as a very early work (e.g. by Blair 1970: 248–50). But as its most recent editor points out (Jones 1975: X–XI) it is in form a random collection of pedagogical notes filed under the letters of the alphabet, derived from very many sources, and almost certainly, in its surviving form, the fruit of many years' practical experience. Some of the specifications are indeed orthographical; e.g.

Aeger animo, aegrotus corpore; utrumque per *ae* diphthongon scribendum (9.45–46)

Grammatical comments are more numerous, e.g.

Psallo perfectum facit psalli (45.936)

Psallo perfectum facit psalli

There are many semantic distinctions made on the lines of Isidore's *Differentiae* (which is one of the sources):

Caeruleus naturae color est, caeruleus fingitur; ita alterum est, alterum fit (16.229–30)

The comment in Isidore's *Differentiae* I 154 (PL LXXXIII 26):

Inter *deportare*, *comportare* et *exportare*. *Deportare* est aliquid afferre, *comportare* in unum locum conferre, *exportare* tollere,

is rephrased as

Adportare est aliquid adferre; comportare in unum locum conferre; deportare, deponere; exportare, tollere. (12.126–27)

As these examples show, Bede is not merely the passive compiler of other people's apertus. He includes and rephrases anything that experience suggests might be useful in Latin classes, being prepared to

rephrase, and in particular abbreviate, when he sees fit. Even so, pronunciation does not seem to appear to Bede to be a matter of much significance. The discussions of Romance homonyms such as *ad* and *af*, carefully distinguished in both the *Differentiae* and the *Origines*, leave Bede cold; *af* is mentioned only in the context of a semantic distinction from *in* (10.74.78). *Origines* I 27 has little interest for Bede, despite being entitled *De orthographia*: only one comment seems to be taken verbatim, from I 27.19:

Pene, quod est coniunctio, per e; poena, quod est supplicium, per oe
(44.912-13)

Recte dicere is not applied to the manner of pronunciation:

Alvus vitrorum recte dicitur; uterus mulierum; venter in utroque sexu
(12.136-37)

Indeed *dicere* can specifically mean "write":

Caseus masculini generis est, sed Pomponius neutraliter dixit, caseum molle
(15.200-01)

Sometimes, however, *dicere* does mean "say":

Sollenne cum dicis sive scribis, m sequenti syllabae connectis. Sonnum similitur.
(49.1056-57)

Jones notes no source for this: perhaps it came from the Italian connection, since church usage may have been the only context in which syllable division was relevant. Similarly:

Maestas, cum scribis aut dicis, s secundae syllabae complicari debet. Sic in similibus
(35.685-86)

The source is: "Si maestus scribis, *stas* in diductione vocis esse debet, non *tas*" (Caper; Keil VII 96.11). Latin was read and recited, almost certainly on a rough basis of one sound per already written letter, but phonetic detail does not worry Bede.

Bede also composed a *De Arte Metrica* (Jones and King 1975: 81-141). King said of this (74) that "only now are we beginning to realize how skilfully and intelligently Bede shaped the exhausted heritage of the past to meet the new requirements of the early medieval monastic community." The implication that the earlier Spanish writers were not meeting the needs of their own community is extraordinary, but the recognition of Bede's practicality is refreshing. Palmer (1959) commented that the *De Arte Metrica* is a practical classroom tool, a critical synthesis. This down-to-earth approach made his a standard course book all over Europe, becoming, for example, the normal metrical manual at Ripoll in Catalonia (Nicolau 1920: 3). Bede's *De Orthographia* is, similarly, practically pedagogical; and the language taught was not in this case any variety of the contemporary vernacular of the recipient community, but an artificial text-based construction.

Bede taught Egbert of York, who taught Aelbert, who taught Alcuin. A tradition of education became established in eighth-century Britain, and unlike that of the seventh-century Spaniards, hampered by the Moslem invasions, the Anglo-Saxon tradition expanded and grew into the foundation of much of later educational practice. Alcuin of York was taught an Anglo-Saxon tradition of Latin pronunciation, divorced from any Romance community, but regarded as normal and acceptable in York. In the late eighth century Alcuin seems to have succeeded in conversing in Italy, where the vernacular was not hopelessly unlike his own usage, but it is quite understandable that the pronunciation he met in Northern France should have taken him aback. The French of Northern France had evolved considerably more; Alcuin's practice, based on a unit of sound for each unit of spelling, was quite unlike the vernacular of the Romance-speakers in the Carolingian realms. Here for the first time, the artificial and, in fact, archaic pronunciation of the Anglo-Saxons did coexist with vernacular in the same place.

Alcuin had grown up in a community in which this artificial "Latin" and the vernacular were self-evidently different languages. Anglo-Saxon was obviously not the written Church language of Latin. In previous Romance communities, however, the modern postulation of a similar distinction can only be a mirage and an anachronism. There is no need to postulate that both existed in Romance areas before Alcuin's arrival at the Carolingian court; since the evidence is consistent with the absence of such a distinction, it seems rational to regard the postulation of a "Latin" coexisting with Romance in Romance communities as a simple misconception of now outdated scholarship. With Alcuin's arrival in France, the stage is set for the imposition of the Anglo-Saxon style of reading aloud on the European ecclesiastical education systems; in short, for the invention of "Medieval Latin."

Mohrmann nearly realized this: "Thus for the first time, among the Irish and Anglo-Saxons, that symbiosis is found of Latin as the language of the Church, of education and of higher culture, with the vernacular as the language of everyday life ... this Irish-Anglo-Saxon culture, which was the first to assign to Latin a place alongside the vernacular ..." (1961c: 166). If she had only followed this remark up in the context of pre-Carolingian philology, or if only Romance philologists took more interest in the nature of their primary document sources (Late Latin), much misconceived argument could have been avoided. But now we can at last admit that in earlier centuries the grammarians, scholars, lawyers, poets and bishops spoke their own vernacular. As one would expect.

3

CAROLINGIAN FRANCE:
THE INVENTION OF MEDIEVAL LATIN

A great deal has been written about the Carolingian "Renaissance". Much of this scholarship has concentrated on the revival of interest in Classical culture, but this was merely one by-product of a wider reconstruction; in McKitterick's words (1977: xx), "Carolingian rule meant a consolidation, a reform, and a positive attempt at the re-shaping of society within a Christian framework ... the whole of society was to be taught and a Christian society created." This did not imply universal education, but it did necessitate a regenerated educational system for members of the Church. In the end, the re-shaping of European society was to take a few more centuries, but part of the foundation for the intellectual life of the later Middle Ages rests on the establishment of professional education for clerics in the early 800s; and that education rested on the international standard language we now know as "Medieval Latin", learnt for and used in official contexts by the Carolingian Church.

Standard Delivery of the Liturgy

Liturgical standardization seemed a necessity. By the eighth century there had arisen considerable variety in liturgical practice throughout Europe. In Merovingian Gaul, where Christianity had in any event a weak hold, there were the "Gallican" rites. When Pepin began the attempt to revive the intellectual life of his realm, he introduced features of the Roman rites, intended in time to become the normal usage. The extension of Papal practice into Gaul became a part of Charlemagne's policy for the reorganization of his church in the late eighth century, and in 787 the Roman rite was decreed to be the standard (McKitterick 1977: chap. 4). The introduction of Roman

elements had in the event led to less uniformity, as bishops in different areas chose to mix Gallican and Roman features in different ways, so a need for a standard usage seemed as pressing as ever. The problem was not immediately solved by the decree of 787, mainly because the *Hadrianum*, the supposedly authoritative book sent to Charlemagne by the Pope in 781, was deficient in a number of ways. The task of producing acceptable liturgical texts, which could become the official versions, was entrusted to Alcuin, the English church had a good reputation for preserving well the Roman traditions, so Anglo-Saxon scholars seemed suitable for this purpose (Bullough 1973b).

The whole Bible needed a standard edition. This need was particularly pressing for the Gospels and the Psalms, much used in liturgy (Ganshof 1947). This in turn required correct spelling; so at the same time as Alcuin was preparing his Vulgate, he compiled a *De Orthographia*. In the performance of the liturgy, the standard text, however well spelled, would not produce the required uniformity of excellence over all the Empire without some kind of decreed method for the manner of reading the texts aloud. The decreed method for reading aloud was not based on any Old French vernacular, but on the method Alcuin had learnt himself; each letter on the page had to be given a sound, and that sound was specified (as in lists such as that from Marianus Capella, printed above pp. 100-01). This was how Alcuin had learnt to recite Latin, on the basis of the written forms, and this was how his clergy had to read aloud in Church. In the absence of a phonetic script the standard pronunciation could not be unambiguously described on paper, but it is possible to reconstruct most of the details of the original instructions as follows. When reading aloud in Church it is probable that the written letters listed in the left-hand column were to be read aloud as the sounds listed in the right-hand column:

Written letter	Sound
a	to be read aloud as [a]
ae	[e]
b	[b]
c	[k]
d	[d]
e	[e]
f	[f]
g	[g]
i	[i]
k	[k] in some specified circumstances
l	[l] in other specified circumstances
m	[m] when found

Written letter to be read aloud as Sound

n	to be read aloud as	[n]
o		[o]
p		[p]
q		[k]
r		[r]
s		[s]
t		[t]
u		[u]

[u] in some specified circumstances
[v] in other specified circumstances
[w] after a or q

x was presumably [ks] (as in Martianus Capella III 261), although Alcuin's *De Orthographia* does not mention it. The above prescriptions were easy enough for Alcuin to prescribe, but the letter *h* was awkward; it looks as if *h* might have been prescribed [h], and *th*, *rh*, *ph* and *ch* might even have been [th], [rh], [ph], [kh] "ut plurius sonent". This list did not solve every problem that arose – in particular, concerning which syllable in polysyllabic words ought to receive the stress – but it has provided a solid basis for the pronunciation of Latin for the last 1200 years (as well as for phonetic script, as can be seen from the symbols above).

Norberg detailed some of the linguistic consequences of this reform, most pithily in his *Manual* (1968: 50–53). For example, the prescription that all words spelt with *t* are now pronounced with [t], and all words spelt with *e* are now pronounced with [e], means that rhyme of these written vowels no longer occurs in verse; previously such syllables often could rhyme, as originally short [i] and long [ē] had merged into the one vernacular sound. Similarly, words written with *u* and *o* cease to be found rhyming. Conversely, no distinction is now made in the new Latin between those words spelt with an *e* which was originally short and those with an *e* which was originally long; all have in the new Latin [e], so all could rhyme in later Latin, whereas in previous years vernacular had distinguished the two. Similarly, all words in *o* can now rhyme. These prescriptions also established a distinction between plosive *b* [b] and fricative *v* [v] that was often not there in the vernacular. These particular reforms seem to have been acceptable in recitation, since [i], [e], [o], [u], [v] and [b] existed anyway in the Romance inventory. Words borrowed later from this innovating pronunciation into French testify to the success of the reform; e.g. the [i] in *digne* from *DIGNUS*, the [u] in *étude* from *STUDIUM*, the [-v-] in *habile* from *HABILIS*. The practice of the Anglo-Saxon scholars was successful in these respects, and indeed in the general prescription that every written letter (or digraph such as *ph*)

should be given some sound (cp. Allen 1970: 106; Beaulieux 1927). This requirement inspired many long-silent letters into new productive life.

In other respects, the strict prescriptions seem not to have been generally followed. [k] may have been required by Alcuin before [e] and [i] – or it may not, since [u] before a front vowel could well have been an allophone of Anglo-Saxon /k/ as it was of Italian – but a palatalized [tse], [tse] or [tse], [tse] seems to have become usual for written *ce*, *ci*; the same may well have applied to -*ri* before a vowel (e.g. *rationem* [tasionem]) and to *ci* and *gi* ([dʒ] rather than [s]). The prescriptions concerning *h* were in any event odd, and the aspirate seems often to have been ignored. (Hence the common spellings of *nichi* and *michi* with *ch* to specify the presence of [h].) Nor was it clear to all how to produce geminate consonants. But on the whole, the introduction of the new reading pronunciation stuck, and ever since then the Latin pronunciation of a word has been a function of its traditional spelling (cp. Lüdtke 1978: 441).

It is worth pausing here to stress how different this prescribed method of production was from contemporary French. For example, *VERIDIARIUM* now had six syllables; in Old French *vergier* it had two. *FERIT*, *DIRECTUM*, *COGNITUM*, *LACET*, etc., became in the new system unrecognizable as vernacular *fieri*, *dreit*, *cointe*, *gistr*, etc. At a stroke, much of the vocabulary had become unintelligible to the uninitiated. Were such a reform to be introduced into Modern English or France, words generally pronounced as [ɔɪndʒ], [nait], [wɪmɪn], [stɛɪʃən], [pe] and [swet] would be unrecognizable in the new system as [ɔɪnʒ], [knaɪt] or [knɪxt], [wɒmən], [stæɪʃən], [paɪks] or [paɪx], and [səʊlaɪtɪŋ]. There has turned up suddenly the strict distinction between Latin and vernacular pronunciation that scholars have liked to think existed throughout the Middle Ages.

Charlemagne's court was probably bilingual, being more or less on the Romance-Germanic border, although many of its most powerful members came from further east. Some of the German Christians would have met Insular missionaries, and many had already been taught the Anglo-Saxon pronunciation of Latin before encountering vernacular French. The difference between this Latin and that French would have been as noticeable to them as it was to Alcuin, and Alcuin's prescriptions probably coincided in most respects with what native German-speakers tended to do anyway. Nor was Alcuin the only scholar involved in the educational reforms; some of the works once attributed to Alcuin have recently been reallocated to St Benedict of Aniane and Theodulf of Orléans, who were scholars in the Visigothic

Spanish tradition. It is undoubtedly true that the Carolingian Renaissance owed a great deal to the Spanish Renaissance of the 600s, and indeed also to the Italian pair Peter the Deacon (of Pisa) and Paul the Deacon (of Monte Cassino); Peter was for many years the Court authority on grammar, and Peter and Paul each wrote an *Ars Grammatica* in the Donatus tradition. It is almost certain that Peter introduced Priscian's *Institutionum Grammaticarum Libri XVIII* to the Court from Italy; Priscian taught Latin in Constantinople in the sixth century, and his work was known in Italy at least at Vivarium, Monte Cassino, and Benevento. Eventually, Priscian became the standard advanced textbook as Donatus continued to be the beginners'; Alcuin studied Priscian avidly in France¹ (mainly, it seems, for information about Greek) (O'Donnell 1976). Yet despite the activities of other scholars of the time, it does seem that Alcuin was the *primum mobile* of the newly decreed pronunciation of Latin, particularly during his time at Tours after leaving the Court in 796; and in this respect a key work, for all that it has been neglected since c.830, is Alcuin's *De Orthographia* (PL CI 901-20; Kell VII 295-312; Marshall 1952).

Much of Alcuin's *De Orthographia* is derived from Bede's work of that name; other details are taken from Cassiodorus, Isidore and Priscian (books 1, 5, 7, 8 and 14). This debt to Priscian helps confirm the date of 796-800 usually given to the *De Orthographia*, for Alcuin is known to have been studying Priscian at that time and there is no reason to suppose that he had read the *Institutiones* in detail earlier. Assuming this rough dating to be correct, it thus coincides with the time when Alcuin was based at St Martin of Tours, working on his Vulgate (Garshof 1947; Fischer 1957); it also coincides with the circular Alcuin drew up for Charlemagne under the heading of *De Literis Colendis* ("On the necessity of studying *literae*"). It is thus reasonable to see all these works as manifestations of a desire for linguistic correctness in church.

Alcuin's *De Orthographia* may have been given that title as a result of the general later impression that it was merely a summary of Bede's work of the same name. Jones, for example, said it was "little more than a digest of Bede's" (Jones and King 1975: 2). As I have shown exhaustively elsewhere (Wright 1981), this is not true; there are many entries not in Alcuin that are in Bede, but also vice versa. For example, under the letter A, Bede lists c.68 words and Alcuin c.71, but

only 20 of the entries coincide, and the statistics are similar throughout. Those chosen from Bede are usually the comments that are relevant to spelling and pronunciation, and even then they are not often transcribed verbatim. This is the key to the constructive use Alcuin made of Bede and of his other scholarly sources; he was particularly concerned with authoritative comments made by established scholars on matters of spelling and pronunciation, and not particularly concerned with anything else. So although Alcuin does not seem to have given the title *De Orthographia* to his work himself, it is at least more accurate than it is as applied to Bede's work. Unfortunately (with the honourable exception of Roger 1905: 343-49) modern scholars have been led by this title not to notice that it also deals with pronunciation. Originally, instead of such a pity title, Alcuin headed his work with an elegiac couplet:

Me legat antiquas vult qui profertre loquelas
me qui non sequitur vult sene lege loqui.

(Vienna Nationalbibliothek 795; printed by Migone, with *sene* corrected to *sine*, PL CI 902D.) The use of *legat*, *profertre*, *loquelas* and *loqui* makes it unambiguously clear that at least one of the aims of the work is to have its clients reading aloud in the correct old manner. This heading caught on, being often used in later manuscripts of Priscian².

Alcuin's concern for pronunciation in his *De Orthographia* formed the subject of Wright (1981), which will not be recapitulated here. The evidence includes: the fact that Alcuin chooses many remarks from his sources relevant to pronunciation, and most of those not selected are not relevant; the fact that he adds information relevant to pronunciation to many of those chosen remarks; a concern for euphonia in a narrowly phonetic sense; prescriptions as to whether written vowel pairs are or are not diphthongs; the implication that the pronunciation is determined by the spelling; the implication that variation in spelling implies a variation in pronunciation; the use of *dicere* to refer specifically to phonetic matters; the similar use of *profertre*, and *legitur*; the assumption that the letter *h* has some consequential effect on the sound, even though Alcuin is not sure what. These interests are new. They are not the concerns of Cassiodorus, Isidore, Bede, or any other earlier grammarian. Alcuin was having to prescribe the "correct" manner of speaking for a clientele already used to reading aloud the same words in a vernacular manner; this was something new, that had to be taught and learnt.

1. There is no reason to suppose that the "Priscian" mentioned in Alcuin's poem on York was the complete *Institutiones*.

2. Marshall (1952) printed a variant: *Me legat antiquas cupiat qui recte loquelas/me spernent loquatur mox sine lege peritum*.

Vienna 795 is an intriguing manuscript, probably compiled in 799, and recently published in a facsimile edition (Unterkircher 1969). Of its 205 folios, folios 21r-150v concern Biblical exegesis, folios 184r-191v concern the geography of Rome, and folios 1v-4v, 150v-183v, 192r-197v contain 20 letters of Alcuin's (including eight to Arno of Salzburg). The folios from 5r to 20v form two separate *quaterniones*, which may have had originally a separate existence before the collation of the MS. Folios 5r-18v contain the *De Orthographia*, headed by the couplet quoted above; folio 19 contains the Greek alphabet, 20r the Runic alphabet and 20v the Gothic. The Greek alphabet is headed by the phrase "FORMAE LITTERARUM SECUNDUM GRECOS"; it is written mostly but not exclusively in capitals, with the equivalent Latin letter on the left (except for O, presumably omitted by mistake), the name of the Greek letter above, the number it represents in Greek mathematics and the equivalent Roman numeral to the right. In addition, there are three number symbols (for 6, 90 and 900) with no alphabetical use, inserted in arithmetical position. It is interesting to note that Latin *e* is said to be the equivalent of both epsilon and eta (spelt *hita*); that *z* is the equivalent of zeta; that theta, spelt *thita*, is said to correspond to *t et h*; that kappa, spelt *capra*, is said to correspond to *c et q*; that xi causes graphical confusion, but is unhesitatingly for *x*; that tau is spelt *thau*; that upsilon, called *eu vel ui*, is for *y*; that phi, spelt *fi*, is for *f*, even though alpha is spelt *alpha*; that chi, spelt *hi*, is for *h*. This might be taken to help fix what Alcuin and his associates understood by *h* in various positions, although the state of 19r does not inspire total confidence in the competence of the scribe.

19v is even more fascinating. The left hand column contains all the Greek consonants except rho and chi, each with a line to itself in which it is followed successively by *ai*, *et*, *α* and *oy* (sic). Above each syllabic group is a Latin alphabetical equivalent, successively with *e*, *i*, *y*, *u*. Thus, for example, 19 runs as follows:

ne	ni	ny	nu
Nai	Nei	Noi	Noy

This is presumably interpretable as being the instructions for the correct pronunciation of Greek vowel pairs. At the top right corner of this left hand half of 19v appears this list: *ay au, ey eu, ai e, H i, ei i, oy u, oi y, ω o*; which is also likely to be prescribing the pronunciation (on the right) of the Greek letters (on the left), even though the left entries are in Latin writing (other than H and ω). What *y* represents in speech

3. The one exception (*fi* instead of *fi*, over *wei*) is presumably a scribal error.

is not at all clear; the *eu vel ui* comment suggests that it could be the front rounded high vowel [y] (as in French *tu*); it is not mentioned in the *De Orthographia*. But the general point, that folio 19 is in part designed for the correct oral reproduction of Greek script, fits the suggestion that these folios are collectively a manual for reading aloud as well as for writing. (The right hand column of 19v deals with compound Greek numbers.) The entries for the Runic and Gothic alphabets (on folio 20) have similar Latin equivalents added above them; the thorn (þ), spelt *ðorn*, is said to represent *ð*, the *wyn* to represent *u*, for example. The experimental interest shown here in letter-sound correspondences, allied to the contemporary spread of written Germanic (one of the six marginal glosses in the *De Orthographia* is Germanic: *uoneni*, for *perpes*, 11v), could have proved a useful first step for the elaboration later in the century of a new orthography designed to represent Old French pronunciation.

The novelty of Alcuin's concerns at this time may be referred to (with the words *novas* and *incipiamus*) in a letter Alcuin wrote to Charlemagne in March 798:

... quid faciemus de litteris syllabis etiam et verbis, quibus uti nobis necesse est cotidie, nisi novas grammaticae artis regulas excogitare incipiamus?

Alcuin definitely wished the Court to participate in the reform; in a later letter to Charlemagne (799), he comments:

Ego itaque licet parum proficiens cum Turonica cotidie pugno rusticitate. Vestra vero auctoritas palatinos erudiat pueros, ut elegantissime proferant quicquid vestri sensus lucidissima dicant eloquentia.

The "rusticity" of Tours suggests that before Alcuin came they used French rather than the Anglo-Saxon reformed Latin even in this respected cultural centre; the use of *proferant* suggests that pronunciation is uppermost in his mind at this time. The word *rusticitas* is here in the middle of its semantic change from "lack of culture" to "non-Latinity", having both meanings at once. The letter which Alcuin wrote to Arno in 798, when Arno was bishop of Salzburg, starts its final paragraph as follows:

Nunc velim te properare in patriam et ordinare puerorum lectiones, quis grammaticum discat, quis epistolas et parvos libellos legat, quis sanctam scripturam (sic) sobria mente haurire dignus sit. Tu vero, sancte pater, evangelicis maxime studios lectionibus et canonicis sanctorum scripturarum inservire eruditionibus...

This letter is included immediately prior to the *De Orthographia* in Vienna 795, and may have been what inspired the composition of the

manuscript in 799. Several letters from these years show Alcuin's continual concern with the reform (the word *corrigere* turns up often).⁴

Quintilian declared that grammar was the *recte loquendi scientia*. Isidore agreed: "Grammatica est scientia recte loquendi, et origo et fundamentum liberalium litterarum" (I 5.1). Neither of them seem to have been thinking of pronunciation; the precepts are aimed at the written mode. Alcuin picked up this traditional phrase but, as often, altered it significantly to suit his own context: "Grammatica est litteralis scientia, et est custos recte loquendi et scribendi" (Alcuin's *Grammatica*: PL CI 857D). The old phrase used *loquendi* for language in general; Alcuin adds *scribendi*, apparently with the specific intention of distinguishing correct talk from correct writing. The two together, intimately connected in Alcuin's mind, form a single science of *litterae*; the *litteralis scientia*, "knowledge of how to read and write properly", becomes from this time on a fundamental requirement of all educated men, and is based on *litterae* in its simple meaning of "written letters and the sounds with which to read them aloud".⁵

Standard Education

Alcuin was a serious scholar with a demanding project to carry out from 796 to 800. He would not have wasted time in a mere copying exercise, as is suggested in the usual dismissive approach to his *De Orthographia*; it is, instead, part of his interest in the propagation of correct Latin literacy, *Litterae*. This concern is also attested in the circular known as *De Litteris Colendis*.

In 789 there had been issued an edict known as the *Admonitio Generalis*, concerning the need for education (*MGH Legum* II i 52-62). At some date in the late 790s this was given a supplementary section entitled *De Litteris Colendis*, "On the need to learn *litterae*" (Wallach 1959: 202-04). Wallach showed that Alcuin had an important hand in this; it is roughly contemporary to the *De Orthographia*, and stresses a requirement that had not been spelled out in the same detail in 789 but was much in Alcuin's mind after 796. It begins as follows:

... nos una cum fidelibus nostris consideravimus utile esse ut episcopis et monasteriis, nobis Christo propitio ad gubernandum commissis, praeter regularis vitae ordinem atque sanctae religionis

4. Alcuin's letters are printed in *MGH Epistolae* IV and in PL C. These are *MGH* 145.232.33 - 233.2; 172.285.21-23; 161.260.13-16.

5. In verses ascribed to Theodulf of Orléans and put in the mouth of *Grammaticus*, Grammar is seen as a medicine for incorrect *loquelas*, *oris vitia*: "Pulchra medela fito balbis dum recto loquelas/Oris enim vitis sum medicina potens" (PL CV 334D).

conversationem etiam in litterarum meditationibus eis qui, coramante Domino, discere possunt, secundum uniuscuiusque capacitatem docendi studium debeant impendere, qualiter, sicut regulantur norma honestatem morum, ita quoque docendi et discendi instantia ordinet et ornent seriem verborum ut, qui Deo placere appetunt recte vivendo, et etiam placere non neglegant recte loquendo ...

"We, in company with our *fideles*, have thought it desirable that bishops and monasteries, entrusted to us to be governed by the Grace of Christ, should emphasize, in addition to the observance of monastic discipline and the practice of holy religion, the study of *litterae* for those who by God's grace are able to learn, according to their ability to teach them; for as observance of the monastic rule preserves purity of behaviour, so in the same way this perseverance in teaching and learning should direct and embellish the series of words, in order that those who aim to please God by living correctly should in addition not fail to please him by speaking correctly ..."

(translation based on Laisner 1957: 196-97)

The training in the correct life had formed the subject of the original *Admonitio Generalis*; the concern for *litterae*, "to embellish the series of words" already written, is the new prescription here. The intention is to have the clerics, in the traditional phrase, *recte loquendo* (later: *quamobrem hortamur vos litterarum studia ... non negligere*). In practice, the main requirement of the system advocated, and extensively put into effect over the following decades, was that the priesthood should be able to read aloud the corrected texts entrusted to them by the specialist scribes. This is normally a desire for reading aloud only (see Riche 1973: IV chap. 2); the acquisition of the ability to write is an achievement of a higher order to which the average cleric would not need to aspire. Reading aloud included chanting; hence the common connection in texts of succeeding centuries between *legere* and *cantare*, with masters of both teaching in schools of both reading and singing. (Later in the *De Litteris Colendis* there are mentioned *scolasticos bene loquendo ... in legendo seu cantando*.) The performance of the Church service involved the oral reproduction of authorized texts. Even the sermons were specified in advance; priests were not expected to make up their own. So, at first, the practical requirement involved in the learning of *litterae* was merely the ability to read texts aloud *in situ* by correctly allotting a sound to every single previously written letter (or digraph) in accordance with the system promulgated by Alcuin and his colleagues.

What Alcuin understood by the word *littera* is partly based on

Donatus and Priscian. They had explicitly associated individual *litterae* with particular noises; Alcuin repeats, "*litterae est pars minima vocis articulatae*" (PL CI 855A) "... syllabas in litteras dividimus". The etymology Alcuin gives to the word, however, is adapted from that offered by Isidore: "*Littera est quasi legitura, quia legitibus iter praebent*" (ibid.). "Letters are so called because they show the way to the readers", as indeed they do in the Carolingian system (used ever since). This is not quite what Isidore meant, however:

usus litterarum repertus est propter memoriam rerum. nam ne oblivione fugiant, litteris alligantur. in tanta enim rerum varietate nec disci audiendo poterant omnia, nec memoria contineri. litterae autem dictae quasi legiturae, quod iter legitibus praebent, vel quod in legendo iterantur. (*Origines* I 3.2-3)

Isidore's phrasing here is based on Sergius; for them the value of letters (apparently as a collective noun: i.e. = "writing" in general) is as an aid to remembering what might otherwise be forgotten. Alcuin has either misunderstood Isidore, or, more probably, consciously adapted Isidore's words for his own purposes, since he seems to mean that individual letters help the reader to read correctly. As usual, the fact that an eminent scholar used words previously used by another is no indication at all that the two meant the same thing by those words; often, a scholar with something new to say tried to disguise the novelty and pretend to respectability by using authoritative turns of speech to his own ends.

The use of such *litterae* depends closely on both the legibility and the consistency of the written texts to be read. Not only should the words be spelt correctly, but each individual letter had to be immediately recognizable in order to aid the reader. Pre-Carolingian scripts were of at least a dozen varieties; several were cursive, and not all were easy to read even at the time (see Lesne 1938: 379-89; Reynolds and Wilson 1968: 80-82). In addition, manuscript punctuation was eccentric when it existed at all, with words sometimes not divided from each other, or split arbitrarily. Any cursive script would have been awkward for a reader who had to see the separate letters clearly in order to *recte loqui*; it is probably for this reason that the so-called Caroline Minuscule script was adopted, at least for the elaboration of important manuscripts designed to be read in church. In this script each letter was quite small, but separately written as an independent unit, with only a few, well-known, contractions. This script was, it appears, used first in a developed form at Corbie in the 770s; at that time it was one among many. Alcuin seems to have been instrumental

in having it widely adopted as a standard practice. His scriptorium at Tours became an important centre for both scribal training and the minuscule script, and the Bible that Alcuin had presented to Charlemagne at his imperial coronation (Christmas, 800) is also an important landmark in the establishment of the Tours script as the official model. By the middle of the ninth century, this script was in general use, for important texts, over all the empire. There was no necessity to use it for other contexts. Bischoff has, for example (according to Bullough 1973b: 585), identified Alcuin's hand in some marginalia in insular style dated to Tours, 797-8. Multiple copies could be made, once the system was established; a *lector* could read letters sound-perfect to several scribes at once, who could then reconvert to letters with professional accuracy (Riché 1973: 246-49). Punctuation begins to be more systematically introduced, both as a semantic aid to reduce possible ambiguity, and as a sophisticated series of phonetic hints for the performers (Parkes 1978; Hubert 1972); Alcuin mentioned the desirability of punctuation in a letter to Charlemagne of 799 (MGH 172.285.16-20). Further evidence of a desire for linguistic correctness at this time can be seen, for example, in the rewriting of the Benedictine Rule (Mohrmann 1952), or in the avoidance of non-Classical elements in the official *Annales Regni Francorum* after 796, and the elimination in c.814 of non-Classical vocabulary from the annals for 741-95, originally "written by someone untouched by the Carolingian revival" (Adams 1977b: 258). These dates of 796-800, when Alcuin was at Tours, keep recurring as a time of concern for linguistic accuracy.

The evidence of the *De Litteris Colendis* implies that the new system of *litterae* was intended to be established in centres outside Alcuin's own personal influence as well as at Tours; the extent to which other institutions took notice of the prescriptions is indeterminable. Perhaps it was only Alcuin's immediate circle who took it seriously at first. One of that circle was Bishop Theodulf of Orléans (see McKitterick 1977: 52-57). Theodulf carried out the prescriptions of the *Admonitio Generalis* energetically. In his *Capitula ad Presbyteros Pirochiae Suave* (Canon 20) he instructed his clergy to teach *litterae* free of charge to the children of the faithful (PL CV 196), "ut scholas ipsi habeant in quibus fidelium parvulos gratis erudiant". Theodulf declared that *grammatica* is the root of all education, as in a picture of a tree where "huius grammatica ingens in radice sedebat" (PL⁹ CV 333B; *Carmine* IV ii 3). This archdiocese, at least, was doing its duty.

One of Alcuin's pupils was Rabanus Maurus. Rabanus wrote *De Clericorum Institutione* — it is not clear when — of which Book III

Chapter 18 begins as follows:

De arte grammatica, et specibus eius.

Prima ergo liberalium artium est grammatica, secunda rhetorica, tertia dialectica, quarta arithmetica, quinta geometria, sexta musica, septima astronomia; grammatica enim a litteris nomen accepit, sicut vocabuli illius derivatus sonus ostendit. Diffinitio autem eius talis est: Grammatica est scientia interpretandi poetas atque historicos, et recte scribendi loquendique ratio. Haec et origo et fundamentum est artium liberalium. Hanc itaque scholam Dominicam legere convenit, quia scientia recte loquendi et scribendi ratio in ipsa consistit. Quomodo quis vim vocis articulae seu litterarum et syllabarum potestatem cognoscit, si non prius per eam id didicit? Aut quomodo pedum, accentuum et positurarum discretionem scit, si non per hanc disciplinam eius scientiam ante percepit? Aut quomodo partium orationis iura, schematum decorem, troporum virtutem, etymologiarum rationem, et orthographiae rectitudinem novit, si non grammaticam artem ante sibi notam fecit? Inculpabiliter enim, imo laudabiliter hanc artem discit, quisquis in ea non inanem pugnam verborum facere diligit, sed rectae locutionis scientiam et scribendi peritiam habere appetit.

(*PL CVII 395*)

"Grammar is the discipline of interpreting poets and historians, and also the rationale of correct writing and speaking; the means of correct speaking and writing. Who would know the value of speech sounds or letters and syllables if they had not first learnt it in *grammatica*?" (After this Rabanus moves on to *tropi*.) Speaking *recte* might in Rabanus' mind have become a technical term, with the meaning it had in Alcuin's *Grammatica* and the *De Literis Colendis*; Rabanus declares that those who have not studied this kind of *grammatica* are unable to talk correctly.⁶

Institutionally, there can be no doubt that schools were eventually established in the Alcuinian tradition. In a few centres, at least, the Carolingian reforms left a lasting mark. The basic ability to read Latin letters as the reformers wished continues. "The educational machine which Charlemagne and Alcuin had set in motion, working through the monastic and Cathedral schools, had sufficient momentum to keep going until a new age could take over..." (Reynolds and Wilson 1968: 91; cf. Gibson 1975; Bulough 1970). Over the next three centuries, some monastic and episcopal schools suffered deprivations from Moslems, Vikings, Magyars or organizational apathy, but Émile

Lesne's enormous volume (1940) on *Les écoles de la fin du VIII^e siècle à la fin du XI^e* is sufficient to reassure us that "la tradition des institutions scolaires ecclésiastiques a survécu par sa force propre" (443). Lesne also confirms here that all schools in these centuries were dependent on the church (414ff.); that there is an inherent connection between the teaching of *grammatica* and *scholae cantorum* (e.g. 415, 469, 574); and that *grammatica* continued to be the basis of all other disciplines (e.g. 587-93). Grundmann (1958) showed at length how *litteratus* continued to mean "Latinate" rather than "literate" throughout the Middle Ages. There can, in short, be little doubt that the reformed *litterae* remained the ideal; that in at least a few intellectual centres the coexistence of vernacular Romance and Latinate pronunciation is an established fact after the ninth century.

Riché (1973: 254-61) has described what actually occurred in these schools (also Lesne 1940: 572-77; Bousard 1972). The alphabet was an early lesson. It was taught so that the pupils could read aloud; not so that they could write, nor so that they could understand a text, but so that they could recite. "Savoir lire, c'est savoir le latin" (Riché 1973: 257). The initial grammatical primer was Donatus's *Ars Minor*. The teaching was almost all oral in nature, either by recitation of passages or mnemonics or through established question and answer routines; for trainee priests or monks this included the techniques of church recitation. The official standard was thus inevitably encountered by any literate person. Writing was usually only learnt by specialist scribes at a later stage. There were thus many who had learnt to read only, or had taken a short cut by learning by heart selected passages for recitation parrot-fashion, with no ability to write, nor compose speech independently in the artificial non-vernacular manner. (There had survived a comparable tradition of reciting Greek without knowing what it meant.)

Education was mainly, but not exclusively, for members of the church. If laymen wished to learn, they could do so, but only in the clerical schools. Nithard, for example, aristocrat, warrior and grandson of Charlemagne, learnt enough to compose his history on behalf of Charles the Bald. But he was exceptional; the training of priests, monks and choirs had been the initial point of the reform, and so it was to continue. The teaching of "Latin" pronunciation was essentially confined to the church schools; in Norberg's words, "L'histoire du latin médiéval est, en d'autres termes, l'histoire de l'école médiévale" (1977: 63). The teaching of correct written Latin was essentially intended to ensure accuracy in the oral reproduction of the written texts of the

6. Rabanus declared in a letter of 814 giving advice on how to read aloud, that "H non littera, sed nota aspirationis esse" (*MGH Epistolae* V 383.33); under any modern definition, it was a letter, but *litterae* here are specifically those letters from which sounds are proffered.

gospel, psalter and missal:

Et ut scolae legentium puerorum fiant. Psalmos, notas, cantus, compotum, grammaticam, per singula monasteria vel episcopia et libros catholicos bene emendate; quia saepe, dum bene aliqui Deum rogare cupiunt, sed per inemendatos libros male rogant. Et pueros vestros non sinite eos vel legendum vel scribendum corrumpere; et si opus est evangelium, psalterium et missale scribere, perfectae aetatis homines scribant cum omni diligentia.

(*Admonitio Generalis* Canon 72)

If there is a risk of God not understanding the supplications made to him, correct spelling and recitation of the texts becomes essential?

Sermons

Once the reformed pronunciation was introduced, the services stopped being intelligible to the congregation. They were no longer the collective celebrations of all the community; the congregation were uncomprehending spectators. The texts to be used were fixed. Their manner of performance was fixed. At first, this may have seemed admirable; when Alcuin died in 804 he may not have realized a worrying practical consequence of the new celebration. This correct but uncomprehended performance was incapable of achieving another of the essential aims of the church: the requirement to preach. Every church instruction of this time stresses how important it was for the clergy to preach. The clergy of the late eighth century, however, could not be trusted to make up their own sermons; if it was essential to preach, it seemed even more essential that the contents of their preaching should be certified in advance. Specific previously written texts were allotted to particular occasions, and priests were forbidden from using their own words:

... ut presbyteros quos mittitis per parrochias vestras ad regendum et ad praedicandum per ecclesias populum Deo servientem, ut recte et honeste praedicent: et non sinatis nova vel non canonica aliquos ex suo sensu et non secundum scripturas sacras fingere et praedicare populo. (*Admonitio Generalis* Canon 82)

The set texts were the homilies of the *patres*, the Fathers of the Church; Saints Gregory, Augustine, Ambrose, Jerome and soon Caesarius of Arles. Of these the most important collection was the 40 *Homilies on the Gospel* of Gregory. While at Tours, Alcuin had collected together

authorized homilies into two large manuscript volumes (Lesne 1938: 147-48). Rabanus Maurus collected many of his own (*PL* CX 13-468; McKitterick 1977: 97-104). One of the headings found in manuscripts of Alcuin's *De Orthographia* points out that those who pronounce in the prescribed manner will talk like the *Patres*, who had written these homilies. Accordingly, at Tours at least, the homilies must at first have been given the "Latin" treatment; they were the last and longest of the group of set passages fixed for recitation during the service. The purpose of the homilies was to improve those who heard them, but that could hardly be achieved if the congregation did not understand them. The reforms had created a problem. At this early time, when it may well only have been the diocese of Tours that was wholeheartedly carrying out Alcuin's prescriptions, it may similarly have been only a problem at Tours.

In 813 the Church held five regional councils, at Arles, Chalon, Mainz, Rheims and Tours, designed to assess the state of the Church at the time and set guidelines for the future establishment of a Christian community (for a full account, see McKitterick 1977: chap. 1). The bishoprics based on these five towns cover most of the Empire. Charlemagne did not attend any of them, but he gave them his blessing. The five councils had similar long agendas; they each produced many recommendations (*Canones*), and a composite list of their resolutions was produced (printed in *MGH Legum* III ii 1). Many practical questions were raised, discussed and decided, in a markedly business-like spirit. The fourteenth item in the composite list of decrees concerns the general need to make preaching more intelligible under the heading of *De officio praedicationis, ut iuxta quod intellegere vulgus possit assidue (sic) fiat*. This wording seems to have been taken from the 25th Canon of the Council of Mainz, ... *qui verbum Dei praedicet iuxta quod intellegere vulgus possit*. Here, perhaps for the first time, is at last attested the use of *vulgus* to mean "laymen" in specific contrast to the Latinate; a distinction which, as we have seen, was not made before. At Rheims, the same idea was expressed with *omnes*, in Canon 15: *Ut episcopi sermones et omelias sanctorum patrum, prout omnes intellegere possent, secundum proprietatem linguae praedicare studeant*. These requirements are most easily understood to refer to the manner of pronunciation, since the words are fixed and invariant; it is likely that in practice some substitution of vocabulary, or even glossing, could have taken place, but this does not seem to be the intention of these decrees.

The Council of Tours, however, went into this question in more detail. Alcuin was no longer alive, but active followers of his such as

7. Hahler (1957: 225) grasps this point, but has apparently not heard of the Romance languages; he says that before the reform the texts were "gibberish". In fact it would be the new recitation that sounded incomprehensible to the congregation.

Theodulf were there. Canon 2 of the Council suggests that bishops should learn to recite by heart the gospels, the epistles of St Paul, the works of the *Patres* as far as possible, and *ceteri libri canonici*. Canon 4 suggests that they have a duty both to preach and to set a good example. (Theodulf elaborates this theme in the instructions to his own clerics of c.814: *PL CV* 191-224.) The next 12 and last 34 Canons are on other topics, but Canon 17 returns to the question. This Canon begins with an emphatic introduction stressing that the council members are unanimous; this introduction appears in no other Canon, and its inclusion seems to suggest that the matter in question had been the subject of considerable argument previously in the church at large:

Videtur est unanimiter nostrae, ut quilibet episcopus habeat omelias continentes necessarias ammonitiones, quibus subiecti erudiantur, id est de fide catholica, prout capere possint, de perpetua retributione bonorum et aeterna damnatione malorum, de resurrectione quoque futura et ultimo iudicio et quibus operibus possit promereri beata quibusve excludi, et ut easdem omelias quisque aperte transferre student in rusticam Romanam linguam aut Theotiscam, quo facilius cuncti possint intelligere quae dicuntur.

This canon specifies not only the content and purpose of the essential homilies, but the manner in which they are delivered; the preacher should be sure to switch into *rustica Romana lingua* or German (*Theotiscam*), so that the congregation can follow. Here the word *rustica* appears in specific contrast to the Latin of the rest of the Church services.⁸ There can be no doubt that this instruction was aimed at exactly the same people who were simultaneously being nagged to use the new "Latin" elsewhere; the novelty of this problem explains why no council in preceding centuries had even mentioned the subject (Riché 1962: 537). It was a new problem, possibly confined to Tours.

I have argued at length (Wright 1981: 355-58) that *transferre* in the usage of the time does not mean "translate" but "transfer", in linguistic contexts it means "to use metaphorically", as in St Augustine's *translatio verbo* and in Donatus himself. Rabanus uses it in the same way in his *De Clericorum Institutione*, chapters VIII, X and XIII. Caelius Sedulius, in the second preface to his combined verse and prose work (*Carmen* and *Opus Paschale*: *PL XIX* 547A-B) in the late fourth

century, said that he composed the verse *Carmen* first and then transferred it to prose (*in rhetoricum me transferre sermonem*), the later operation being called *translatio* (*dicentque nonnulli fidem translationis esse corruptam*). This is unambiguously not a transfer from one language to another, but from one style to another within the same language. There is, in short, no justification for assuming that *transferre* in this 17th Canon necessarily means "translate". Since it appears that the wording of the text of the authorized homilies was not supposed to be altered, "translation" is probably the wrong interpretation in Romance communities such as the Tours bishopric. In Germanic communities, translation would be necessary, so transferring to *Theotisca* would indeed involve different words and grammar; in Romance communities, the old method, used anyway before 796, of reading texts aloud in vernacular pronunciation, was here being reestablished. The distinction may not be being made here between two separate languages, Latin and Romance, but between two methods of reading aloud written texts, the *litteralis* and the *rustica*.

Unfortunately, this picture has been confused by modern philologists. The last part of Canon 17 has been reprinted many times in their handbooks, often without any explanation at all of the context. It has been claimed to be conclusive proof of the theory that Latin and Romance had coexisted as separate distinct spoken norms for centuries; that the natural speech of the educated had hardly changed for five hundred years or more, and was thus so far removed from popular Romance that translation was now necessary (e.g. Hall 1974: 105; Rickard 1974: 27, etc.). Some of the philological comments seem to imply that the vernacular was something new, or "emergent". This is impossible to take seriously. Every community has its own vernacular, by definition. There is no possible doubt that vernacular existed through the preceding centuries; what does seem to be new in the early ninth century is the Latinizing alternative. There is no reason to suppose that Latin and Romance were thought of as completely different languages for at least another two centuries. The decision of philologists to interpret *transferre* as "translate" is merely a consequence of their inbuilt assumptions, and can no longer seriously be used to support them.

Tours was the only centre to decree the switch in reading styles in 813. By the middle of the century, most of Alcuin's precepts had become standard practice; the Roman liturgy and Caroline minuscule were normal usage, Priscian and Donatus were mainstays of an ambitious educational system firmly based on *litterae*, and the existence in many

8. This use of *rustica* may have been Tours' practice already. There is a famous gloss by Berno of Tours, of the early ninth century, *rustice dicitur rehi*, which suggests a) that *rustice* refers to non-Latin pronunciation, and b) that it indeed represented a fricative [v] (or possibly [d]). It was made famous by Savage (1928).

areas of Latinizing pronunciation alongside the Romance was accepted. The Council of Mainz in 847 reproduced the 17th Canon of the Council of Tours whole in its second Canon, saying that although the old Canons ought to be prescribed reading for all priests anyway, the Council specifically thought it essential to repeat this one. In total, this Canon runs as follows:

De dogmate ecclesiastico, cum igitur omnia concilia canonum, qui recipiuntur, sint a sacerdotibus legenda et intelligenda et per ea sit eis vivendum et predicandum, necessarium duximus ut ea, quae ad fidem pertinent et ubi de extirpandis vitiis et plantandis virtutibus scribitur, hoc ab eis crebro legatur et bene intellegatur et in populo praedicetur, et quilibet episcopus habeat omelias continentes necessarias ammonitiones, quibus subiecti erudiantur, id est de fide catholica, prout capere possint, de perpetua retributione bonorum et aeterna damnatione malorum, de resurrectione quoque futura et ultimo iudicio et quibus operibus possint promereri beata quibusve excludi, et ut eandem omelias quisque aperte transferre studeat in rusticam Romanam linguam aut in Thioriscam quo facilius cuncti possint intellegere quae dicuntur.

Rabanus Maurus was bishop of Mainz at this time and keenly interested in effective preaching; the prescription of the Council of Tours seems by 847 to be accepted practice.

The Invention of "Romance" Writing

a) Oaths

By 847, however, we have already passed the most famous linguistic event of the century; the "Strasbourg Oaths" of 842. These were oaths taken by the Romance-speaking Charles the Bald and the German-speaking Louis the German to cement their agreement against their brother Lothair. They were included by Nithard in his *De Dissensionibus Filiorum Ludovici Pii*, and in recent years have been discussed in minute detail. Ewert (1935) put them into their context in a classic study. He concluded that although the two French and the two German sections are not translations of any surviving Latin passage, they are probably produced from a first draft in Latin; in Elcock's words, "It seems almost to be a Latin consciously vulgarized by adaptation to the most common features in the vernacular of the day, features which were widespread over Northern France" (1975: 351). Elcock may have included the word "almost" in order to prevent the need for the logical next step, that of wondering why the language was "vulgarized" at all, when Latin transcriptions of ordinary speech in historical works were normal practice. In fact, as Elcock points out, Nithard, cousin to the

Princes and chronicler of these events, may well have helped to draft the oaths himself. In any event, the written forms of the oaths were certainly prepared before the oaths were taken rather than being an earwitness account of how they actually sounded. Such oaths would be very carefully scrutinized in detail in advance; in this case, that these texts had to be fixed in advance for subsequent reading aloud becomes clear when we see the strange linguistic character of the sequence of events.

By 842 the official "Latinate" pronunciation of reading aloud was known to the literate. German-speakers of that time could easily have been much more fluent in reading aloud in Latin than they were in French vernacular. It is possible that Louis the German might not have known how to read aloud an oath in French vernacular pronunciation, had he held in his hand a version in traditional Latin orthography. In this situation, where it was essential that the Romance-speakers at Strasbourg should hear and understand exactly what Louis was saying, it is understandable that Nithard, or someone else in the Chancery of Charles the Bald, should have taken the peculiar step of attempting to transcribe what Louis was to say according to the sound-letter correspondences that had become current in the reformed Latin system, but with the sounds to which the written letters corresponded being the sounds of vernacular French rather than those of reformed Latin. Louis must have been able to read, or else the Romance text would be no more help than any other. The text of the Oaths could have been enough to prompt him into a passable rendition of contemporary French, if he produced the sounds he had already learnt to use for written letters.

In short, whereas the Romance Strasbourg Oaths are usually thought to be produced for the benefit of French-speakers inexperienced in Latin, I suggest that the evidence points to their being produced for the benefit of a Germanic-speaker who had learnt to read Latin but was not fluent in French.

The scene is solemn. Louis the German-speaker and Charles the Romance-speaker (the first native Romance-speaker in the dynasty) assembled with their respectively German-speaking and Romance-speaking troops. The Princes addressed their own troops in their own language, and the gist of what they said is presented by Nithard in ordinary Latin (transcriptions from Elcock 1975: 346-48)⁹:

9. Most early Romance texts have recently been collected together in Sampson (1980); the Oaths are no. 50.

Erigo xvi kal. marci Luthuicus et Karolus in civitate que olim Argentaria vocabatur, nunc autem Strazburg vulgo dicitur, conveniunt et sacramenta que subter notata sunt, Luthuicus romana, Karolus vero teudisca lingua, iuraverunt. Ac sic, ante sacramentum circumfusam plebem, alter teudisca, alter romana lingua, alloquuntur sunt. Luthuicus autem, quia maior natu, prior exorsus sic coepit:

'Quotiens Luthuicus me et hunc fratrem meum, post obitum patris nostri, insectando usque ad internecionem delere conatus sit noster. Cum autem nec fraternitas nec christianitas nec quodlibet ingenium, salva iusticia, ut pax inter nos esset, adjuvare posset, tandem coacti rem ad iudicium omnipotentis Dei delinimus, ut suo nutu quid cuique deberetur contenti essemus . . .'

Cumque Karolus haec eadem verba romana lingua perorasset . . .

It is worth noting that *vulgo* at the start here applies to German; Strasbourg is in the Germanic area.

The exact wording of these initial addresses is not of legal or other significance, so there is no need for Nithard to transcribe these verbatim in German and Romance, even if he could. Louis speaks in German, then Charles says *haec eadem verba* in Romance, which refers to the content of the words rather than the words themselves.

In the following oaths of alliance, Louis the German spoke the Romance oath and Charles spoke the German one. Both were thus using the tongue that was not their normal native language, and in Louis' case, at least, a tongue he may not have known well. The German oath was concocted according to what was becoming common practice in Eastern scriptoria. The Romance oath was more unusual, but it had precedents in all those pre-Carolingian legal documents designed to be in substance intelligible when read back to their depositors (cp. Chapter 2), and the usage of Merovingian notaries has been detected in it. (That this old notarial tradition was known in Carolingian scriptoria can be seen from the *Lex Saxe*, a sixth-century code that only survives to us in Carolingian script, with the original quasi-vernacular passages left uncorrected, presumably on purpose, to preserve intelligibility; Elcock 1975: 320-22.) Nithard's text presents them as follows (continuing the previous extract):

... Luthuicus, quoniam maior natu erat, prior haec deinde servaturum testatus est:

'Pro Deo amur et pro christian populo et nostro commun salvement, d'ist di in avant, in quant Deus savir et podir me dunat, si salvarai eo cist meon fradre Karlo, et in eynha et in cadiuna cosa, si cum om per dreit son fradre salvar diu, in o quid il mi altresi fazet, et ab Ludher nul plaid nunquam prindrai, qui,

meon vol, cist meon fradre Karle in dainno sit.'

Quod cum Luthuicus expleisset, Karolus teudisca lingua sic haec eadem verba testatus est:

'In Godes minna ind in thes christianes folches ind unser bedhero . . .'

The only surviving manuscript is of the late tenth century, so may not reproduce Nithard's version precisely. There are, for example, the following contractions in the manuscript (Paris BN, *FL* 9768) which may well have been spelt out fully in the original: *dō* for *deo*, *p* for *pro*, *xrian* for *christian*, *nō* for *nostro*, *cōmun* for *commun*, *dē* for *deus*, *sicū* for *sicem*, *p* for *per*, *nūquā* for *nunquam*. There is also a punctuation point (.), probably designed to indicate proper pausing, after *salvement*, *avant*, *dunat*, *eo*, *Karlo*, *adidha*, *com*, *dist*, *fazet* and at the end. To this extent the standard reproduction as in Elcock may be mildly confusing.

After that, the followers of each side (*utrumque populus*) pronounced an oath in their own language, which Nithard also reproduces in the two vernaculars:

Sacramentum autem quod utrumque populus, quique propria lingua, testatus est, romana lingua sic se habet:

'Si Luthuicus sacrament quod son fradre Karlo iurat conservat et Karolus, neos sendra, de suo part lo fraini, si lo retourner non l'ist pois, ne io ne neuls cui eo retourner int pois, in nulla ajudha contra Luthuicung nun li iv et.'

Teudisca autem lingua:

'Ota Karl then eid then er sinemo broodher Luthuicunge gesur . . .'

It is not clear why this second Romance oath was written in this way; not many of Charles' *populus* could have been likely to be able to read at all, let alone from a baffling orthography that they had not met before. The answer may simply be that once Nithard, or whoever it was, had worked out the first oath for Louis to read, they felt they were sufficiently pleased with what they had achieved to carry on in the same vein for the other Romance oath, even if strictly it was going to be of less use as Charles' *populus* knew French anyway.

The nature of the connection between these Oaths and contemporary Romance has given rise to a huge bibliography that will not be recapitulated here. Copious references appear in Ewert (1935) and in Sabatini (1968). Sabatini points to the close descent from Merovingian legal practice; he does not argue that this is in fact Merovingian Latin making a comeback (as Nelson 1966 does), nor that it is a "preocious" phonetic script representing in accurate detail some actual localizable Romance dialect, as most early philologists assumed (Castellani 1978

also does, but his approach has been uprooted by Hilty (1978), but the compromise view that although this is indeed an attempt to represent vernacular speech on paper, it is not based on an improbable synchronic phonemic or phonetic analysis so much as on a development of the previously used analogous techniques. Clearly, whoever drafted Louis' oath would prefer to help him rather than hinder him; something in a recognizable tradition might help more than a new system of phonemic transcription. Hilty's conclusion (1978: 141) is that there is not an exact correspondence here of vernacular sounds to letters; he exemplifies this anisomorphism with final [ə] (as in English *paper*), a sound not existing in the prescribed Latin sounds of the Carolingian system, which is represented with different letters in different words in the written version of the French oath.

These Oaths show that by 842 the idea has arisen that on some occasions — for example, when some text has to be fixed in writing but nevertheless its reproduction has to be intelligible — there might be an advantage in a reformed spelling in which the vernacular sounds were the points of reference for fixing the orthography. This is not yet evidence of a general split into two languages; priests continued to read sermons written in the old way with vernacular phonetics, as was demanded five years later in the second Canon of the Council of Mainz. Yet there were others around who could read traditional spelling in Latin pronunciation, but not with the French that every Romance-speaker had used automatically in France before the reform. These would include, in particular, Germanic speakers. The origins here of vernacular spelling seem then to lie not in an attempt to record what the writer has heard others say, but to record what the writer wants others to say. We have entered now the strange stage in Romance culture where most works were written in uniform "Latin", but just a few others, designed for oral reproduction in an unreformed manner (e.g. oaths, songs, sermon notes), were written in a newer more experimental semiphonetic manner that could vary from place to place. The whole is still thought of as one language, but with two possible reading pronunciations and two possible methods of manuscript transcription. This is comparable to the coexistence in modern Britain of traditional orthography and phonetic script in the technical works of some linguists; or to the coexistence of the Initial Teaching Alphabet with traditional orthography in some British primary schools of the 1970s. The additional complicating factor in the Frankish domains was the reformed Latin pronunciation, which had produced as an eventual consequence a need for the adoption of a reformed way of writing vernacular.

b) The Sequence

One other invention of the early ninth century seems intimately linked to the reformed "Latin", in which each letter gave rise to a sound; the Sequence (see Ryle 1976; early sequences are printed in *AH* VII). The "classic" sequence that originated then took the form of words composed to fit preexisting melodic lines, specifically allotting one Latin syllable to each note of the music. Nearly all the early sequences were antiphonal, and the parallel isosyllabic structure of most of the surviving sequences appears to guarantee that this practice was normally followed even in the many cases where the precise music is not known to us. To begin with, other considerations such as stress or rhyme were non-existent, or at most secondary, and the melodies used were often the extended melismas on the Alleluia (to which the word *sequentia* was applied first). It may be that their original purpose was merely as a practical mnemonic aid to remembering the contours of these phrases. Since sequences were performed in church, the Latin reading pronunciation would have been the only one ever required. It is only this equivalence of one Latin syllable to each note that seems to be the innovation of the classic Carolingian sequence; antiphonal form predates it. Dronke (1965) has unfortunately confused the two issues in proposing that sequences existed long before 800. He has two arguments for this. Firstly, he decides that some of the early sequences are so sophisticated that the genre cannot have been of recent origin: "the prevalent view that 'postulates the beginnings of sequence composition around 830 at the earliest' is, even on purely stylistic grounds, impossible" (47). By "stylistic" Dronke seems to mean "guesswork" here. Dronke could well be right in his controversial view (50) that Alcuin did indeed compose a sequence attributed to him in an eleventh-century manuscript; even so, a date of c.800 for its invention should be early enough for such a simple concept to have been skilfully elaborated by 830. Dronke's second argument concerns the antiphonal form, which he declares to be reminiscent of the earlier Spanish *preces* of the Visigothic liturgy (as in Férotin 1912: XXXIX-XL). His argument is not convincing (as Szövérfy 1971: 69-70 shows), but even if he is right he is not postulating that the Carolingian equivalence of notes to Latin syllables existed in Spain. The *preces* are verses of a stress-patterned nature and possibly antiphonal in performance. Parallelism of literary structure is constantly present in Iberia from at least the third century until the twelfth and later (Rico 1975). The syllabic equivalence found in the Carolingian sequence, however, seems to have been invented at the same time as the official pronunciation of Latin, being another

aspect of the requirement that in church one written vowel implies one oral syllable. There is no reason to suppose that these sequences were ever sung in anything other than strict Latin pronunciation, or before Alcuin's reforms.

One of the main centres of scholarship in the late ninth century was the abbey of St Amand, near Valenciennes in North-Eastern France (Plattelle 1962). It had been eminent in pre-Carolingian times (Charles Martell's son was educated there), and it continued to be. Arno was abbot from 782 till 785, when he became bishop of Salzburg, and he seems to have returned at times afterwards. Manuscript Vienna 795 was compiled for Arno, and as well as Alcuin's *De Orthographia* it contains many letters to Arno from Alcuin. Alcuin himself visited St Amand in 798 and 799, apparently leaving behind fifteen metrical inscriptions for churches and altars (*MGH Poetiae* I 305-08). It is inconceivable that the active school at St Amand could have been run in the ninth century on any lines other than those laid down by the reformers. In 860, the scriptorium had such renown that Charles the Bald commissioned a sacramentary to be written there. Two of his children were educated at St Amand. Charles' son Carloman became abbot from c.867-70; his successor, Gozlin (870-84 or 86) had been imperial chancellor from 867. Gozlin's successor, Robert (until 922 or 23), was the brother of King Eudes (888-98). Milo, the chief scholar at St Amand (d.872), had been a pupil of Alcuin; Milo and his successor Huchald (840-930) were two of the most respected scholars of the late ninth and early tenth centuries. Milo's poetry includes a verse life of St Amand which in itself is evidence of expert reformed Latinity (*MGH Poetiae* III 557-676).

In 882 the Vikings ransacked the monastery; but the monks had already taken their valuables to St Germain des Prés, of which Gozlin was also abbot. They returned by 886 at the latest, and little serious disruption seems to have been felt. Their library was one of the largest of the age, containing educational and church manuscripts of all kinds, including at least seven ninth-century manuscripts of grammatical works, in particular Priscian and Alcuin's epitome of Priscian. Lesne (1938: 246) declared that St Amand inherited the aura held earlier in the century by St Martin de Tours; cp. also Plattelle (1962: 67-69); O'Donnell (1976); Boureny (1946-47); McKitterick (1977: 27, 124). It is, in short, certain that in the late ninth century it was the base of an enterprising, well-informed, linguistically sophisticated group of Latinate scholars and teachers. It is from here that the next example of written Old French comes.

There is a manuscript from St Amand, probably but not necessarily written there, now at Valenciennes (MS 143), which contains the Romance Sequence of Saint Eulalia. Eulalia of Mérida (Spain) was "a rather unpleasant young virago who had sought and easily obtained martyrdom at the age of thirteen" (Collins 1980: 196). Book III of Prudentius' *Peristephanon* had spread her fame. Relics said to be hers were brought to Harnon, near St Amand, in 878 (Tagliavini 1972: 486). The manuscript contains a Latin treatise of St Gregory of Nazianzus; at the end there is a Latin sequence on Eulalia, an Old French sequence on Eulalia, the German *Ludwigslied* on the battle between Louis III and the Vikings, and another Latin poem of fifteen couplets (see De Poerck 1963: 4). Writing in German was by this time established practice, nearly always for texts designed to be read aloud or sung in Church to German speakers (see McKitterick 1977: chap. 6; by "vernacular" she means "German").

There is a general truth worth remembering at this point; it is easier to write using a system that one has learned than it is to write in a system that one has not learned, even if the latter is closer to one's vernacular habits. Germanic speakers found writing in German considerably harder than the writing in Latin that they had learnt first (until a later time when German was the first form of writing taught to Germans). There are native Catalan-speakers who can only write letters in Spanish. Specialist English phoneticians find ordinary English spelling easier than phonetic script, even after lengthy practice; reflexes need to be deschooled and the ear trained in a novel technique.

The Romance Eulalia sequence must have been written by someone who could also write Latin. No one could write at all if he had not learnt to write Latin. The invention of the Romance orthography represents an awkward task; that it should have been achieved is a sign of an inventive nature and exceptional linguistic sophistication. It is a virtuoso achievement comparable to the invention of shorthand, or George Bernard Shaw's new alphabet. There is, accordingly, nothing incongruous whatsoever in finding the earliest piece of Romance "literature" in a centre of expert Latinity.

Unlike the Strasbourg Oaths, which are based partly on preceding notarial practices, the techniques used for transcribing the Eulalia sequence seem to have been worked out from scratch. A sequence had (probably) never been sung in anything other than reformed Latin before, with a note per vowel; whoever had invented the Romance words to set to this particular music had presumably done so on the basis of a vowel per note. It seems reasonable to assume that

the sound-letter correspondences of the reformed Latin were the starting-point for the new orthography; that the system outlined above for Latin pronunciation could be applied in reverse for the creation of a Romance orthography. The *Sequēce* is thus as good as evidence of Latin pronunciation as it is of Romance. For example, the fact that the [ts] of, e.g. *ciel* (6, 25), is spelt *c*, suggests that in Latin *ci* was pronounced [tsi] here now, whatever Alcuin had done; the fact that the [k] of, e.g. *chi* (6, 12), is spelt *ch*, suggests that in Latin *ch* was pronounced [k] here now, whether or not others elsewhere used [x] or [h]. It is hard to see why these spellings were chosen if this were not the case. For the first time we find diphthongization represented with consistency and even sophistication (e.g. *buona*, 1; *veintre*, 3; etc.); this is because, for the first time, the writer of a new written text specifically wants to have reproduced in performance the [w] in *buona* and the [j] in *veintre*. Outside church, this might have happened all the time, but in church *buona* would be read or sung as [bona]. Porter (1960) has shown the skill and care that has been expended on writing the words in such a way that their phonetic representation could not be in any doubt. *ci* then at St Amand must have been unambiguously [tsi], and *chi* [ki]; *ch* had a precedent in the spelling of words taken from Greek, but the use of *cz* for [ts] before *o* is adventurous (e.g. *czō*, 21). ([ts] is similarly *cz* in the *Ludwigslied*, so the German practice may have come first; the two texts are not, however, in the same hand. See Fought 1979; Penz 1973.) There is some redundancy in the reverse direction, such as *ko* (as well as *co*) for [ko], e.g. in *eskoitet* (5), but that presents no practical problem to performers. There is an attempt here to specify vernacular reproduction in performance; presumably it is designed to be read by skilled professionals used to reading Latin, or else they could not have read at all; probably it is designed for a choir at least some of whose members were German-speakers and thus uncertain or ignorant of French. Otherwise there seems little point in the exercise, since Latin spelling and Romance pronunciation are still prescribed for the homilies and the same could have been indicated here. (Perhaps some of the choir spoke different forms of Romance; any hesitation or inconsistency would stultify the effect, for a unison performance is required for every word.) As with the Strasbourg Oaths, the St Amand *Sequēce* (from near the linguistic frontier) seems to be specifying Romance sounds for at least some German-speakers. It is quite likely that the written version was concocted while the community of St Amand was at St Germain des Prés in or after 882; if so, we can speculate that one of the communities was transcribing its own

invention, a Romance *sequēce*, for the benefit of intrigued parties in the other.

Romance spelling is only needed because of the earlier reform of Latin pronunciation. For those trained in reading Latin correctly, a vernacular French oral performance cannot be unambiguously prescribed except by adapting the standard written letters to the non-Latinate phonetics of French words. *Sequēces* had not previously been performed in anything other than Latin; the novel idea of a Romance *sequēce* necessitated a spelling reform. The phonetics, grammar, morphology, syntax and vocabulary had been current in speech for ages. Only the spelling is new here. That the ninth-century scholars had the technical ability to create a new system for this purpose is attested by the strange fascination some of them felt for alphabets, as in the Vienna manuscript of Alcuin's *De Orthographia*. They were used for "short transliterations" (Bullough 1973a: 117). Germanic writing in Roman letters was developed by them, although, as Bischoff (1961: 213-14) pointed out, they had probably found previous examples of Gothic script in Ravenna. Even so, the fact that this method of Romance writing does not seem to have been widely used suggests that many of the choir were bemused enough by it not to perform the *sequēce* with great fluency.

The transcription designed for this *sequēce*, then, is specifically intended to prompt a desired oral performance. It is not a phonetic transcription of what was actually sung, any more than the Strasbourg Oaths indicate what was actually said. It is, accordingly, a mistake to assume that the text is an infallible guide to north-eastern French. It could well be in a formal style. Perceptible Latinisms such as the *l* in *palcella* (1), or the words *element* (15), *Christus* (27), *clementia* (29), may indicate that these are Latin borrowings, or ecclesiastical vernacular, or scribal confusion, or all three. The fact that some forms may not represent the contemporary French is no bar to the postulation that that was what they were aiming at; not everyone using an unfamiliar variety of phonetic script always gets everything right.

The Latin *sequēce* (transcribed from von Winterfeld 1901: 135) is similar to the French one (transcribed from Ewert 1933: 353-54):

1a	Cantica virginis Eulalie
1b	Concine suavisona cithara,
2a	Est operae quoniam pretium
2b	Clangere carmine martyrum:
3a	Tuam ego voce sequar melodiam
3b	Atque laudem imitabor Ambrosiam.

- 4a *Fidibus cane meos eximium;*
 4b *Vocibus ministrabo suffragum;*
 5a *Sic pietatem, sic humanum ingenium*
 5b *Fudisse fletum compellamus ingentium.*
 6a *Hanc puellam nam iuventutis sub tempore*
 6b *Nondum thoris maritalibus habilem*
 7a *Hostis aequi flammis ignis inpluit;*
 7b *Mox columbae evolutae obstituit;*
 8a *Spiritus hic erat Eulaliae*
 8b *Lacteo oculis celer innocuus.*
 9a *Nullis actis regi regum displicuit*
 9b *Ac idcirco stellis caeli se miscuit.*
 10a *Famulos flagitemus ut protegar,*
 10b *Qui sibi laeti pangunt armoniam;*
 11a *Devoto corde modis demus innocuos,*
 11b *Ut nobis pia deum nostrum conciliet*
 12a *Bius nobis ac adiuvet auxilium,*
 12b *Cuius sol et luna tremunt imperium,*
 13a *Nos quoque munder a criminationis*
 13b *Inserat et bona sideribus*
 14 *Stemate luminis aureoli deo famulantibus.*
- 5 *Buona pulcella fut Eulalia*
Bel auret corps, bellezour anima.
Voldrent la veindre li Deo infini,
Voldrent la faire diable servir.
Elle no'nt eskoitret les mals consellers,
Qu'elle Deo raneiet, chi maent sus en ciel,
Ne por or ned argent ne parameiz,
Por manates regiel ne preient;
 10 *Niule cose non la poutet onque pleier*
La polle sempre non amast lo Deo menestier.
E por o fut presentede Maximien,
Chi rex etet a cels dis soure pagiens.
Il li enoret, dont lei nonque chieft,
Qued elle fujet lo non christien.
 15 *Ell'ent aduret lo suon element;*
Melz sostenetret les empedementz
Qu'elle perdesse sa virginitet;
Por os furet morte a grand honestet.
 20 *Enz enl fou lo getterent com arde tost;*
Elle colpes non auret, por o nos coist.
A czo nos voldret concridre li rex pagiens;
Ad une spele li roveret tolr lo chieft.

- 25 *La domizelle celle kose non contredist:*
Voit lo seule lazsiere, si ruover Kristi.
In figure de colome volat a ciel.
Tuit oram que por nos degnet preter
Qued autisset de nos Christus merit
Post la mort et a lui nos last venir
Par souve clementia.

Von Winterfeld suggested that the Latin sequence, as printed by him, should be emended by adding an additional line 14a, *Aetheris in glorio caerulea*, leaving *Deo famulantibus* as line 15. Given this emendation, the two sequences have the same number of lines (29), with fourteen couplets and a finale; the same linking for assonance, which is at times full rhyme, within each couplet; more or less the same number of syllables in each line of a couplet in both poems, although a little emendation is required if we wish this to be exact; and indeed the same subject matter, although the one is not a simple translation of the other. It is a plausible hypothesis that they might have been intended for the same music, the one version in official language and the other intelligible. (Porezinsky 1965 postulated a Germanic model for the French sequence, but he seems to have been unaware of the Latin genre.)

This analysis implies the presence at St Amand of at least one virtuoso linguist. Was there such a person there at the time (882-900) that the manuscript poems were written? Yes!¹⁰ This was an age of virtuoso poets, skilled in the esoteric arts of acrostic, telestich, alphabetical verse and similar tours-de-force, but one of the most striking pieces of the age was Huchbald's *Ecloga de Calvis*, a poem of 146 hexameters in praise of baldness in which every word begins with the letter C. (*MGH Poetae* IV 267-71. It is the length that starts; the idea had occurred before, e.g., Valerio del Bierzo; Díaz y Díaz 1958.) As an example of intellectual joie-de-vivre, this is comparable to the extraordinary idea of writing down a musical piece in a phonetic script of one's own devising. And even if the system used in the Eulalia sequence was not Huchbald's idea, one of his associates must have worked hard at it. These things do not "emerge" unbidden in the way some phonetic changes can. (Boutemy 1949 offered a similar argument from virtuosity concerning illuminations.) The *Ecloga de Calvis* begins as follows:

Carmine convitii cernitis, carpere calvos
Comatus, cecinit: celebrantur carmine calvi

10. The *Luciwigstiede* was probably composed in 881, but not written into this manuscript until after Louis' death on 5th August 882. The sequence might be from any time between 878 and 900.

Conspicuo clari; carmen cognoscite cuncti.

Carmina, clariora, calvis caritate, Canenae:

Comere condigno conabor carmine calvos,

Contra cirrosi crines confundere colli.

Cantica concelerent callentes clara Canenae;

Collaudent calvos, concludant crimine cluras

Carpere conantes calvos crispante cactinno.

Conscendat caeli calvorum caue cacumen.

Confiteant cuncti concreto crine comati

Cerrito calvos calventes carmine cunctos.

Consona coniunctim cantentur cantica calvis.

It has been suggested before that Huchbald composed the Eulalia sequence (e.g. Contini 1966; Harvey 1945). He was a musician and a choirmaster; his musical treatise *Harmonica Institutio* advocated a number of practical innovations, and his biographer stated directly that he composed *cantilenae*. His hagiographic writing also shows an untypically enquiring approach (see Van der Vyver 1947: 61-62). It has also been suggested that he wrote the *Ludwigsted*; or all three. Harvey (1945: 19-20) dismisses the idea, but it is not implausible; they are all works in search of a skilful author, and must have been composed by someone. Even if they were the inventions of colleagues, it is inconceivable that Huchbald would not have known about the outlandish experiment in the Romance sequence. Huchbald was the most admired intellectual and teacher of the time, being invited, for example, to reorganize education at Rheims from 896 to 900 (Van der Vyver 1947: 69; Platelle 1962: 66; Lesne 1940: 276-81). Attributions of authorship have been made on far less evidence than this.

The foregoing observations are neither adventurous nor novel. Unfortunately, philologists have often been as unaware of the background to the Eulalia sequence as of that to the Council of Tours or the Strasbourg Oaths. Barnett, for example, decided (1961: 19, as regards Latin) that "it is hard to imagine that the oral tradition had yet, as it was to do in later centuries, given way to the re-established 'classical' spellings as the guide to pronunciation". Even the most cursory glance at its context shows that it is impossible to imagine anything else; yet Barnett claims to be examining the Latin background. In fact, Barnett's conclusion (and that of Lorenz 1963) supports the view that the Romance sequence must have been composed by someone skilled in the Latin sequence tradition: "the story it tells has its equivalent, point for point, in many Latin 'passions'. A striking number of the words used, when not part of the basic vocabulary, have common Latin equivalents in religious, and more particularly hagiographic, literature ... it is clear

that in content, structure and language, the French *Eulalia* stands very close to the Latin tradition" (Barnett 1961: 24-25). This is not surprising. Porter (1960: 588), however, apparently believed that before such experiments occurred Romance vernaculars did not exist: "I should hesitate to pretend to impose on his language, considering its then nascent state, the dialectal divisions which came into existence only in later centuries." Fought (1979: 846) pointed out, commenting on this remark, that "the vernacular itself was doubtless no more nascent than any other language"; vernaculars exist, by definition, whatever the quirks required for writing them. Porter's comment illustrates how linguists can be as naive about language as any historian, whereas historians tend not to be naive in addition about history.

The Survival of Latin Speech and Romance Writing past 1000 A.D.

Huchbald died in 930, at the age of 90. The idea of recording words intended subsequently to be spoken in vernacular survived him. By chance there has survived part of a so-called "bilingual" sermon on Jonah, used in the binding of Valenciennes MS 521 (from St Amand), and dated by De Poerck (1956, 1963) to between 937 and 952. Rather than bilingual, it is in appearance a monolingual text with varying orthography. If the whole was to be read aloud on the normal pattern of a sound per letter, the result would have been Latin pronunciation for the text (*Jonah* 4) and French for the commentary. Part of it runs as follows (transcribed from Sampson 1980: no.52):

Et egressus est Ionas de ciuitate et sedit contra orientem ciuitatis donec uideret quid accideret ciuitati. dunc co dicit cum Ionas propheta cel populum habuit pretiet e convers. et en cele ciuitate ... habuit demoret. si escit foers de la ciuitate. e si sist contra orientem ciuitatis. e si auardevet cum Deus parfereiet sa promesse. se Ninive destruite astreici u ne fereiet. Et preparauit Dominus ederam super caput Ione ut faceret ei umbram. labo- rauerat enim dunc Ionas propheta habebat mult laboret e mult penet a cel populum co dicit. e faciebat grant jholt. et eret mult las et preparauit Dominus en edre sore sen cheve qet umbre li fesist. e repauser si podist. Et letatus est Ionas super edera letitia magna. dunc fut Jonas mult letus co dicit. por qe Deus cel edre li donat a sun soueuv. et a sun repausement. Et precepit Dominus uerui ut percuteret ederam et exaruit. et parauit Deus uentum calidum super caput Ione et dixit: 'melius est mihi mori quam uiuere.' ... surde dunc co dicit si rogat Deus ad un verme. qe percussist cel edre sost qe cil sedebat eo ... cilge edre fu seche si vint grancesmes jholt la super caput Ione et dixit: 'melius est mihi mori quam uiuere.'

This piece of paper was ephemeral, and would not normally have been preserved for the next thousand years. Anything worth keeping would be written properly. Thanks to its chance survival we know both that the idea of recording vernacular sounds in a quasi-phonetic script did not die out in tenth-century St Amand, and that the prescription of the Councils of Tours and Mainz concerning vernacular speech for sermons was still, sometimes at least, observed. This sermon is not one of the authorized homilies, having been apparently composed for the occasion of a Viking attack, and the French shows syntactic divergence from the Latin as well; the giver of the sermon has no authorized text to follow other than the scriptural passage and feels free to elaborate on this theme (of deliverance from adversity). The discovery by Gysseling (1949:210) of linguistically comparable marginalia in an eleventh-century manuscript from Douai suggests that these techniques were taught and used during all the period between Huobald and the first extended texts in French vernacular; a developing tradition exists, hidden from us because such works were not intended for the eyes of posterity.

The continuation of the reformed Latin in the tenth century can be exemplified in the life of Abbo, abbot of Fleury (c.940-1004; Cousin 1954; Lesne 1938:549-58; 1940:191-95). Fleury, on the Loire, had had Theodulf of Orléans as abbot, and had been an active centre of the new culture in the ninth century. After a period of decline, it was reestablished under Odo of Cluny from 931 till his death in 942. Abbo seems to have been entrusted to Fleury in the late 940s at the age of seven, been a star pupil and then a teacher; as his biographer Flaimo (Aimoin) puts it, with endearing clumsiness:

Abbo... in Floriacensi monasterio scholae clericorum ecclesiae Sancti Petri obsequentium traditur litteris imbuendus, divina pro certo, ut credimus, id praedictae providentia, ut inde pueritia sumeret litterarum, ubi postmodum, plenissime fluentia doctrinae mentibus erat propinquatur sapientiam sitientium, redderetque illis, seu eorum posteris...

(PL CXXXIX 389A, D)

As usual, when he taught the others, reading and singing were the joint subject:

... imbuendis praeficitur scolasticis; quos ille, per aliquot annorum curricula, lectione simul et cantilena cum tanta eruditura...

(390B)

Litterae, doctrina, litteraria ars, lectio et cantilena were taught and learnt at Fleury in a manner admired by others; Abbo himself was obsessively expert at the *litterarum exercitium* and *studium lectionis* (390A). It is of interest, then, to look at his own comments on grammar as expressed in his *Quaestiones Grammaticales*, compiled for the benefit of English colleagues at Ramsey that he taught from 980 to 982.¹¹

The problems of where to put the stress, and of what was long and short for metric purposes, were perennial questions not catered for in the simple letter-sound correspondences of elementary *litterae*; these questions occupy sections 2-8. In section 9 the question arises of syllable-final plosives (p, c, t, b, g, d); and Abbo apparently embarks on neutralization theory, suggesting that the distinction lapses in 'actus pro agnus, scriptus pro scribitus, et attinet pro adinet' (528A). In section 10, the pronunciation of c and g is dealt with. When a, o or u follow, the consonants are velar ('sonant in faucibus'). With a following e or i, however, life is more complicated:

Cum vero c litteram sequuntur in eadem syllaba e vel i, tritium solet pronuntiar, et nunc quidem ut fere viceatur sonare g, maxime s praecedente, ut *suscipio, suscepit, suscepit*; nunc autem cum quodam sibilo, ac si s illi haereat, ut *cipis, cepit*; quod magis solet fieri ubi r profertur sono z in principio syllabae, ut *laetitia, iustitia*. Denique qui tertium modum addunt sono quae vel qui, eadem syllabas pronuntiant decemur, et *suscepit* pro *suscipio*, et *suscepit* pro *suscipit*, et *quis* pro *cipis*; quod quam involum constet, omnibus vera sapientibus liquet.

The 'strengthened' [g] attributed to *sce*, *sci* is surprising. The sibilant sound elsewhere is almost certainly [ts], here compared to the z in *laetitia*. The absurdity of the slip into [kw] is explained at some length, and in the course of this tirade Abbo mentions that a morpheme boundary is no bar to the velar + back vowel // sibilant + front vowel pattern, even in a verbal paradigm: 'siquidem *vinco, vinet, vince, vincam, mutato cum vocalibus sono dicimus, quemadmodum et *lego, legi, lege, legam**'. It is certain, therefore, that *ce, ci* were read [se], [tsi], and *ge, gi* as [dʒe], [dʒi], in even the best centres of tenth-century Latin. The scholars of Ramsey might well have raised the question of [k] and [g] here, as Alcuin had before; but the wheel has come full circle. Alcuin had the authority to tell Frenchmen to say [ke], but even

11. The edition in PL CXXXIX 521-34 is seriously misleading; it is essential to consult Bradley (1921-23) who, unlike Migne, was able to consult the sole MS (Vatican 596). The text reproduced here is as corrected by Bradley. A new edition is currently being prepared in Paris by Antje Quereau-Jalabert.

if they took notice then, they do not in Abbo's time, and now Abbo has the authority to tell Englishmen to say [tse].

The next problem (section 11) concerns *h*. "Semper absque ullo sono vocalibus praeposimus", at the start of a word, or between two vowels, it has no sound. After *c*, *t*, *p* it is reserved for Graecisms (nearly all of which the manuscript misspells but Migne prints in corrected form). The Greek letters *χ*, *θ*, *φ* have attached to them in the manuscript the names *chi*, *theta*, *phi* (sic), *fi*, which Migne's printed text ignores. In general, "illas proferimus Graecarum litterarum sonis". *Fi* for *φ* is as in Vienna MS 795, and causes no problem: *ph* means [f]. The other two are less simple. *χ* is *hi* in Vienna 795; *χ* is sometimes thought, says Abbo, to be *x* [ks], but (if the manuscript is correct) should be [k]. "Quapropter cum Graece scribitur *Χρονος*, ita proferatur ac si scriberetur *s. kerea*". Bradley suggests this is a bad copy of *cherea*. The *ch* digraph is only used before *e* and *i*, so there is need of clarity in representation of [ke], [ki]; the fact that the Eulalia Sequence has *chi* for [ki] suggests that [x] or [h] are not the standard reproduction of *ch*, and perhaps the manuscript *k* should be respected despite the general requirement to reproduce the Greek sounds. What the *s* is remains an enigma; short for *sono* perhaps. The immediately following comment is "... et παροχια quasi parraechia" (sic, in *PL*), which, if Abbo really wrote it like that, adds to the confusion. As Abbo said earlier, the only doubt about *h* arises "quando inter *c* et *e* vel *i* interponitur nota aspirationis, ut *Chereas* et *parrochia*"; unfortunately the scribe seems to have either transcribed a (hypothetical) *(c)herea* as *s. kerea* or a (hypothetical) *parrochia* as *parraechia*, and the confusion persists as to what Abbo originally meant. The best bet could be that Alcuin may have proposed [x] or [h] on the basis of his knowledge of Greek (and his English [h]), but that in France, where [x] and [h] were absent from vernacular, [k] became a common and understandable (if unauthorized) pattern, once *ce* abandoned [ke] for [tse].

θ is called *thita* in Vienna 795. In section 12 (which is gibberish in *PL* CXXXIX 529), as Bradley points out, Abbo says to his English audience "sed aspirationes bene vos angli pervidere potestis, qui pro *θ* frequentius *p* scribitis" (*effertis* is added here in the margin). The English are better than the French at [θ], having the *thorn* (þ) and its accompanying sound. This seems to confirm that Abbo is aware that *θ* ought to be [θ], but the spelling *teta* — if it is not a copyist's error — may indicate that not everyone could achieve it, [θ] not existing in Old French.

The next problem, in section 13, concerns the need in both speech and writing not to move the final consonant of one word onto

the initial syllable of the next: "tandem dicendum est quod vitando cavenda est collisio quae solet fieri vel pronuntiato vel scripto, ut *venit rex* pro eo quod est *venit rex*, et *par test* pro *pars est*, et *feli xes* pro *felix est*."¹²

This concluded the phonetic section. The nature of reduplicating perfects, second conjugation verbs, the length of *i* in verb-endings, Greek words in -on, the plural forms of *bos*, *zeugnas*, the case used with comparisons, and the implications of negation in a non-complementary predication in the Athanasian Creed, fill the remaining sections.

Given that the pronunciation problems raised by Abbo's English colleagues are (apart from questions of metric length and stress) confined to how to read *c*, *g*, *h*, the implication is that there was no great difference between the English pronunciation of Latin and that which Abbo taught at Fleury. Alcuin's reform of Continental Latin, on the English model, has largely stuck. And Fleury is the home of the manuscripts of three of the earliest Romance compositions, as we shall see below; this can no longer be thought to be a coincidence.

The other outstanding intellectual figure of these times is Gerbert of Rheims, who was slightly younger than Abbo and in touch with Fleury. Hucbald of St Amand and Remigius of Auxerre had in the late ninth century reestablished scholarly traditions at Rheims. Southern (1953: 13) dated the intellectual revival of the Central Middle Ages to Gerbert's arrival at Rheims in 972, but the tradition was there before him; Gerbert's real achievement lay in his large number of active students, who began to expand the number of those studying in the fashion set by Alcuin two centuries before.

The two surviving extended works written in French before 1000 come from Clermont-Ferrand MS 240 (formerly 189: De Poerck 1963, 1964). The main text of the manuscript is a large *Liber Glossarum*, wrongly attributed to the Visigoth Ansileubus, written in the ninth or tenth century. On some blank spaces, hands of the late tenth century recorded nine pieces of verse, the sixth and the ninth of which are in Romance; a *Passion* and the *Vie de Saint Léger*. Together these have 756 lines of verse. Until De Poerck attacked it angrily as a "hypothèse néfaste" (1963: 19-20), the dominant view was that these were the work of an unreliable Southern copyist transcribing Northern poems, and thus not evidence of any particular vernacular (Linskill

12. The *to* vel of *pronuntiato* vel and the *ni* of *venit* are supplied by Bradley to supplement a hole in the page.

1937; accepted by Sampson 1980: nos. 53-54). Lines 157-172 of the *Passion* are here reproduced from Sampson:

Sanct Pedre sols venjar lo vol:
estrais lo fer que al laz og,
si consegued u serv fellon,
la destre aurelia li excos.
Jesus li bons ben reu per mal,
l'aurelia ad serv semper saneç.
l'isade(n)s mans, cume iadron,
si l'ent menen a passiu.
Dunc lo en gupissen sei fedel,
cum el desanz diz lor aveia;
sanz Pedre sols seuguen lo vai,
quae sua fin veder voldrat.
Anna nomavent le Judeu
a cui Jesus furet menez,
dunc s'adunavent li felon,
veder annavent pres Jesum.

The suggestion that this is a Southern adaptation of a Northern text implies a greater degree of dialectal self-awareness than may be thought plausible; it may merely be that the spelling system used was based closely but not exactly on models from further north. This is quite possible. Hucbald himself travelled widely, and it is quite likely that the experiment at St Armand was known in other centres. St Armand's scriptorium also sent books to many other places (Piatelle 1962: 68-69). Hucbald was an associate of the scholars of Auxerre, and Gibson (1975: 8-9) has pointed out that their work established many of the texts available in the tenth century. "On the linguistic side of the curriculum, the school of Auxerre made a contribution that was clearly useful, and may prove to have been fundamental." Monasteries of central and southern France with contacts further north must at the least have been aware of the idea, and probably also of the technique, of recording vernacular in new orthography for vernacular reading; the Clermont poems have a sufficiently confident air to suggest that both their composer and scribe could even have been trained in the art of using vernacular script for notes or poetry. That users of the manuscript had linguistic interests can be seen in the fact that it is a glossary. The third of the Latin poems in the manuscript is a lament for the assassination of William Longsword, the second Duke of Normandy, in 942; as Robson (1955: 132) observed, "only a rich monastery with a strong literary tradition would have undertaken the transcription of this great folio copy of the glossary, and the added Latin and French poems show a considerable range of interests and contacts". De Poerck (1963)

argued that the *Saint Léger*, at least, might have come originally from Breuil, slightly north of Clermont, and more or less on the dividing line between Occitan and French; his view is that the language used could well represent the usage of some such transitional area. Robson shows that "standard written French existed therefore, in terms of flexion, vocabulary, and consonant-systems, from the tenth century" (1955: 165), implying that instead of representing any usage exactly, the system used here is "inter-regional", although scribes of different vernaculars felt at liberty to adapt it if they thought it necessary.¹³ It is not logically impossible for unconnected attempts at vernacular writing to arise independently, but Rickard's conclusion (1974: 42) is more probable: "behind the vagaries of southern scribes we can discern important features which confirm and are confirmed by those early texts which were not copied by southern scribes, and even more by later texts"; Robson concluded that "the few literary texts copied mainly in monastic centres owed their fixity and lack of local quality to a monastic discipline established in the great age of Cluniac reform" (1955: 135).

If there really is a standard written French even in the tenth century, as seems likely, the Cluniacs could easily be the standardizers. This would mean that the main centres of spoken liturgical Latin were also the main centres of written vernacular, which is a predictable consequence of the theory that the invention of Medieval Latin speech was the precipitating stimulus for the invention of Romance writing. The reformed Benedictines of Cluny were established in Burgundy in the early tenth century under Odo, its first abbot. Odo had probably been taught by Remigius of Auxerre, associate at Rheims of Hucbald of St Armand; it is unlikely that Odo was unaware of the practice of vernacular writing at St Armand (although, as Robson 1955 showed, the "standard" did not adopt the same conventions for representing vowels). The Cluniac order was obsessed with both liturgy and literacy (for their library, see Lesne 1938: 524-33). *Illiterati* were only given manual work. Spoken Latin was essential there; this could explain why a (semi-)standardized written vernacular was also taught there, if indeed it was.

Elcock (1961: 16-18) noticed that the idea of writing in a way other than the traditional Latin was suggested by the model of writing in German (or Anglo-Saxon) to the Carolingian scholars, and spread out geographically from there. This is broadly true. The Norman *Vie*

13. This essay is essentially a warning against the belief that spoken Old French was standardized.

de Saint Alexis is usually said to date from 1040-50 (but could well be later). The earliest surviving distinguishably Occitan texts probably postdate the Clermont poems slightly. The *Boecis*, 257 decasyllables on Boethius, is written by a scribe of the eleventh century, in an Orleans manuscript that originally came from Fleury in the tenth century (Lavaud and Machicot 1950); the *Chanson de Sainte Foi d'Agen* (593 octosyllables) is probably from the area around Narbonne, although the manuscript is from Fleury. The first stanza of the *Chanson* mentions Latin, and the second stanza apparently mentions French; the third stanza states that the author heard the tale in Latin from *clerczons* and *gramadis* (transcribed from Elcock 1975: 389-90):

I
Legir audi sotz eiss un pin
Del vell temps un libre Latin;
Tot l'escoltei tro a la fin.

Hanc non fo senz q'el nonl declin;
Parled del pair' al rei Licin
E del linnadg' al Maximin.
Cel meïrols saintz en tal traïn
Con fal venavels cervs matin:
A clusals menan et a fin;
Mortz los laissavan en sopin;
Jazon els camps cuma fradin;
Nolz sebelliron lur vizin.

Ozo fo prob del temps Constantin.

Cançonz audi q'ès bella 'n tresca.

Que fo de razo Espanesca;

Non fo de paraulla Grezesca

Ne de lengua Serrazinesca.

Doiz' e suaus es plus que bresca

E plus qe nulz pimentz q'om mesca;

Qui ben la diz a lei Francesca,

Cuig me qe sos granz pros l'en cresca

E q'en est segle l'en paresca.

Tota Basconn' et Aragonz

E l'encontrada delz Gasconz

Sabon quals es aquist cançonz

E ss' es ben vera 'sta razons

Eu l'audi legir a clerczons

Et a gramadis, a molt bons,

Si qon q monstral passions

En que ont lig estas leicrons

E si vos plaz est nostre sons,

Aisi cond guidal primers tons,

Eu la vos cantarei en dons.

Elcock (1961: 16) suggested, attractively, that *a lei Francesca*, "in the French manner", means "in the vernacular, rather than in

Latin"; not a distinction between vernaculars, but between all vernaculars and Latin. Zaai (1962), however, interpreted *Francesca* as "French" in the sense of "a French genre" rather than anything linguistic. The poem is the second of four in the manuscript, of which the other three are Latin; as usual, early Romance writing is found in the company of expert Latinity.

Arguing from the word *grezesca*, Burger has recently (1978) dated the *Chanson* to 1060-80. The date of the *Boecis* is indeterminable, so it is quite possible that the earliest surviving Occitan poem is the two-line refrain of a Latin *alba* in another manuscript from Fleury, which may be as early as 1000 (and thus contemporary with Abbo's abbacy), and is now in the Vatican (1462). Tagliavini (1972: 495-96) prints the text as follows:

Phebi claro nondum orto iulare,
Fert aurora lumen terris tenue:
Spiculator pigris clamat: "surgite".
Laiba par(i) umet mar atra sol
Poy pas abigil miraclat tenebras.
En incautos ostium insidie
Torpentesque gliscunt intercipere,
Quos suadet prece, clamat surgere
Laiba par(i) umet mar atra sol
Poy pas abigil miraclat tenebras.
Ab Arcturo disgregatur Aquilo,
Poli suos condunt astra radios,
Orienti tendit(ur) Septentrio,
Laiba par(i) umet mar atra sol
Poy pas abigil . . .

The manuscript includes the music¹⁴. Every Latin line is a hendecasyllable in the usual Latin pattern (given the emendation to *tenditur*), with a caesura after four syllables, and homoteleutic rhyme, apart from the *s* of *radios*. The Occitan refrain, of 9 and 11 syllables, represents a conscious attempt to specify vernacular performance. The result is a song with the verse unintelligible to laymen but the chorus not. If this is indeed the earliest surviving Occitan composition, we have here too the origins of the tradition in a combination of "Latin" and "Romance" in the same manuscript (as in the Oaths, the Eulalia sequence and the Jonah sermon).

The Occitan impulse to vernacular poetry in Italy and Catalonia is well known. In Catalonia, Provençal culture and Carolingian education

14. Tagliavini summarizes a variety of possible interpretations. Dronke fantasized another (1968: 170-72).

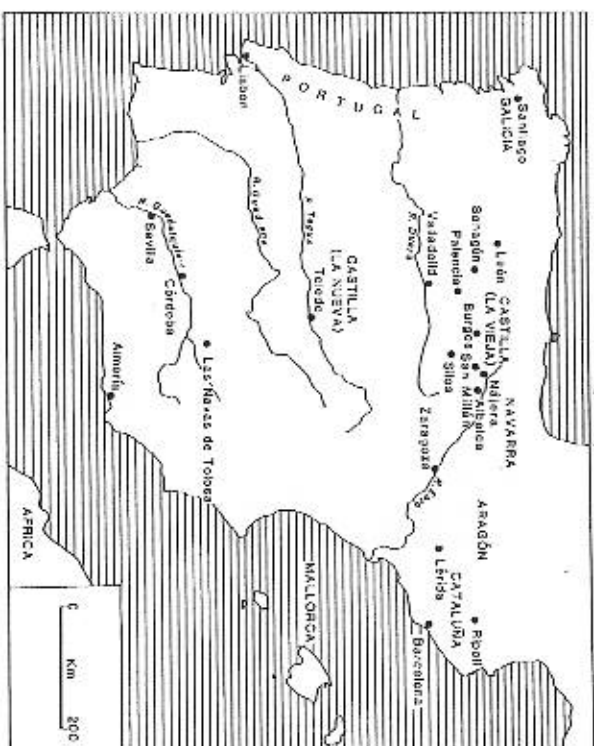
were already well established by the eleventh century. By 1100 or so, in France, the idea is spreading that writing in Romance might be an activity autonomous from writing in Latin. Spain, however, has been quite different. The advent of a distinction between Romance and Latin in Spain (outside Catalonia) does not seem to start until c.1100, or be accepted till the thirteenth century; the complicated history of Romance and Latin in Spain from 700 to 1250 forms the subject matter of the two remaining chapters.

Italy falls outside the brief of this study, but a few initial points can be made. In the first place, it tends to be thought that Northern Italy had many lay schools in the ninth and tenth centuries, with the result that education is not so intimately tied to the requirements of the church; that view was exploded by Bullough (1964), whose arguments have never been rejected. Bullough's study emphasises that *litterae*, reading and singing are all part of the same process in Northern Italy as they were elsewhere in the Carolingian area. In the tenth century, for example, Bullough has found no priest unable to write his name in documents from Novara and Parma, several references to the ecclesiastical establishment of *magistri grammaticae et cantorum*, and no evidence at all of lay schools. Norberg summarizes evidence for the perception of a difference between Latin and Italian vernacular in the tenth century (1968: 34); the coronation of Berengar I in 915 is said to have been marked by speeches *parto ore* and *parto voce*. Gunzo of Novara distinguishes them in his *Epistola ad Augustines* of 965: "Falso putavit sancti Galli monachus me remocum a scientia grammaticae artis, licet aliquando retarder usu nostre vulgaris lingue, que Latinitati vicina est" (Manitius 1958: 27); Pope Gregory V's epiphany (999) mentions his ability to speak *francica, vulgariter voce Latina*, presumably French, Italian and Latin. Apart from the so-called Veronese "riddle" of c.800, which is notoriously hard to interpret¹⁵, the earliest surviving intentionally vernacular Italian appears in four legal documents of the early 960s, where they are new versions of traditional sentences written elsewhere in normal Latin. Presumably the lawyers were experimenting to see if the reformed orthography aided vernacular reading back to, or by, the depositor (Sampson 1980: no.77). This experiment apparently failed, being dropped after 964. The evidence, in brief, suggests that Carolingian-influenced Northern Italy was in a similar position to France: the invention of Latin speech had led to experimentation in Romance writing.

15. *Se pareba bones, alba pretaia araba/albo versorio teneba, negro cemen servitaba* (Sampson 1980: no.76).

4

SPAIN (711-1050)



(Map drawn by A. G. Hodgkins)

The Iberian peninsula is heterogeneous. In the central Middle Ages it was even more heterogeneous than usual. The Moslem invasion of the Visigothic Kingdom began in 711, and until the mid-eleventh century Moslem rulers controlled the central and southern areas. The