

## Introduction

Latin and the Romance languages occupy a vast space along at least three dimensions: geographical, temporal, and social. Once the language of a small town on the Tiber River in Latium, Latin was carried far afield with the expansion of Roman power. The Empire reached its greatest extent under the reign of the emperor Trajan (98–117 CE), at which point it included modern-day Britain, Portugal, Spain, France, Italy, Switzerland, Austria, and the Balkan peninsula, as well as immense territories in the Eastern Mediterranean and beyond, making it by far the largest single state the Western world had ever known. Even those distances are dwarfed by the extent of Western European colonial expansion in the 1500s and 1600s, which brought Spanish to most of Latin America and the Caribbean, Portuguese to Brazil, French to Canada, and all three to their many outposts around the world, where they engendered some robust creoles. On the time dimension, the colloquial speech that underlies the Romance languages was already a constant presence during the seven centuries that saw Rome grow from village to empire – and then their history still has twenty centuries to go. Their uses in society extend to every level and facet of activity from treasured world literature to instant messaging.

A truly panoramic account of Romance linguistic history would find few readers and probably no writers. The scope has to be limited somehow. Our decision, which may disappoint some readers, is to cover five languages: French, Italian, Portuguese, Romanian, and Spanish. Two criteria converge to justify the choice: these are the five that are national languages, and they have the greatest number of speakers, ranking as follows in terms of first-language speakers (Lewis 2009):

<i>Language</i>	<i>Native speakers (millions)</i>	<i>Locations with most speakers (millions)</i>	
Spanish	328.5	Latin America	249.5
		Spain	28.2

## 2 Romance Languages

Portuguese	178.0	United States	28.1
		Brazil	163.0
		Portugal	10.0
French	67.8	France	53.2
		Canada	6.7
		Belgium	4.0
		Switzerland	1.5
Italian	61.7	Italy	55.0
		Argentina	1.5
Romanian	23.4	Romania	19.7
		Moldavia	2.7

In addition, Spanish is spoken as a second language by 60 million people (mostly in the United States and Brazil) and French by 50 million (mostly in Africa).

Venturing into the rest of the Romance languages and trying to count them, we encounter at once the word *dialect* used in contrast to *language*. Historically, our five languages are dialects of Latin, but we have no trouble calling them languages. The difference between *language* and *dialect*, however, is not always clear. The very word *dialect* is a trapdoor leading into a labyrinth of terminological confusion and clashing convictions (among linguists) as well as vested interests and aspirations to prestige (among the public at large). Consider this: one reputable source classifies Lombard, usually deemed a Gallo-Italian dialect, as both a language and a “group of dialects, some of which may be separate languages.” Is there any reliable criterion that can distinguish *dialect* from *language*?

Even a seemingly simple extralinguistic criterion such as number of speakers can be problematic. Sicilian has over twice as many speakers as Sardinian, yet Sicilian is usually considered a dialect and Sardinian a language. And how does one count the speakers without having first established the boundaries of the variety to be evaluated? Should Occitan (as a cover term for Southern French) include speakers of Franco-Provençal and Gascon, which some consider separate languages?

A more nuanced criterion might rest on the idea that languages are the object of attention, elaboration, and social implementation, measured not only by number of speakers, but also by range of uses, size of lexicon, codification (dictionaries, grammars, academies), official status (e.g. in education), a continuous literary tradition, and use in mass media. But this constellation of measures does not yield a binary categorization into language versus dialect. Rather, it points to a continuum along which any variety – be it conventionally called language or dialect – is situated. Varieties conventionally called languages may rank at either end of the continuum, and so can varieties conventionally called dialects. Piedmontese, nominally a dialect, has many accoutrements of a language, while

Cambridge University Press

978-0-521-88915-5 - Romance Languages: A Historical Introduction

Ti Alkire and Carol Rosen

Excerpt

[More information](#)

## 3 Introduction

Sardinian, said to be a language, has fewer. Ultimately, the terms *language* and *dialect* describe the same thing (a linguistic system), and the choice between the two is largely a matter of attitudes and ideology.

Let's rely on conventional wisdom, generated and endorsed by linguists, though in different versions. There are allegedly ten Romance languages, our top five plus these:

<i>Language</i>	<i>Native speakers (millions)</i>	<i>Locations with most speakers (millions)</i>	
Catalan	11.5	Spain	11.2
Occitan	2.0	France	1.9
Sardinian	1.0	Italy	1.0
Rheto-Romance	0.86	Italy, Switzerland	0.85
Dalmatian	0		

Other incipient Romance languages – the Latin of Britain and of North Africa – died in infancy, in the 400s and 700s respectively, and have no names.

Catalan is spoken in Catalonia, the northeastern coastal region of Spain, once a powerful kingdom whose influence peaked in the 1300s. Outside Spain, Catalan-speaking enclaves exist in the Eastern Pyrenees and in Sardinia. Ramon Llull, the most renowned and prolific figure in medieval Catalan literature, fostered its linguistic divergence from the Occitan of Southern France. Once associated with Aragon to the west, Catalonia remained on its own with the union of Aragon and Castile (1479). Over the centuries, Catalan declined in importance as Castilian rose. Championed by nineteenth-century authors and standardized for writing in the twentieth century, Catalan became an official language (1931–1939), lost that status during Franco's attempt to stifle it, regained it in 1978, and is today the language of a lively regional culture.

Occitan is one name for the language of southern France, but also denotes a particular variety. Others prefer Provençal, but this term also refers to a smaller variety. Here we use Occitan as a cover term that includes Gascon (in the southwest), North Occitan (Limousin, Auvergnat, Alpine Provençal) in the central south, Franco-Provençal (in the east), and Middle Occitan (Languedocian, Provençal) in the far south. Old Provençal was the language of a flourishing and enormously influential literature in the Middle Ages.

Sardinia was among the first territories the Romans wrested from their rival, Carthage (238 BCE). Isolated and weak in resources, the island lived under a long succession of ruling powers, the most influential being Catalonia and Spain. Sardinian is best known for retaining certain features of Latin elsewhere lost. Of the four varieties on the island, Logodurese and Campidanese are genuinely Sardinian, while Sassarese and Gallurese are considered to be closer to Tuscan Italian. Sardinian documents appear from 1080 on, and a literary tradition emerges in the Middle Ages, but there is no standard today.

Rheto-Romance is a group of dialects or languages (opinions differ) spoken in the Swiss and Italian Alps (hence one alternative name, Alpine Romance)

Cambridge University Press

978-0-521-88915-5 - Romance Languages: A Historical Introduction

Ti Alkire and Carol Rosen

Excerpt

[More information](#)

#### 4 Romance Languages

stretching west to east from the Swiss canton of Graubünden (Grisons) to the Friuli-Venezia Giulia region in the northeastern corner of Italy. One group, Rhetic, includes the varieties spoken in Switzerland, often called Romansh. In 1938 it acquired the status of a national language in Switzerland. The other two groups are spoken in Italy: Dolomitic (known as Ladin < LATINU) in several valleys of the Dolomites, and Friulian in Friuli-Venezia Giulia.

Dalmatia is the Roman name for the eastern coast of the Adriatic, from Trieste in the north to Dubrovnik (Ragusa) in the south, today part of Slovenia and Croatia. Dalmatian, now extinct, was a relic of what must have been a wider range of dialects from the Latin of the Roman provinces there. Dalmatian is attested from 1325, and was submerged first in the south by Croatian. The northern variety, Vegliote, lived on until 1898.

This book, with some regret, ignores these five languages of the second tier. We first survey the historical phonology and morphology of Italian, Spanish, and French (Chapters 1–8). Portuguese has its own chapter (Chapter 9), as does Romanian (Chapter 10). The closing chapters, on the lexicon (Chapter 11) and the medieval emergence of the Romance languages (Chapter 12), deal with all five.

# 1 The evolution of stressed vowels

The languages of the Romance family are descended from Latin, but what kind of Latin? Just as Modern English exists in many varieties and registers, so also Latin came to be a socially complex language, extending over a vast territory and serving the needs of diverse speech communities. Among the educated, a codified literary Latin existed, enshrined in the classics and in treatises on grammar. But Latin also lived on as an evolving spoken language among the far-flung populations of the Roman Empire. The basic vocabulary of the Romance languages bears the imprint of a casual, spoken style of Latin, always open to change. In conservative social contexts, in the domain of religion and high culture, the frozen classical language remained an influential presence in the minds of the literate few, and became in later centuries a source of new layers of vocabulary.

## 1.1 Syllables and word stress in Latin

In this overview of the sound changes leading from Latin to the major Romance languages, we begin with the stressed vowels. Since the first requisite is knowing how to identify the stressed vowel, this lesson explains vowel quantity, syllable weight, and the rule that assigns stress in Latin.

### 1.1.1 Why word stress matters

Stress position in a Latin word (**etymon**) has a crucial effect on its Romance outcomes (**reflexes**). For example, consider these Latin infinitives and their reflexes:

<i>Latin</i>		<i>Italian</i>	<i>Spanish</i>	<i>French</i>
DEBĒRE	‘owe’	doverē	deber	devoīr
HABĒRE	‘have’	avere	haber	avoir
VĒNDERE	‘sell’	vendere	vender	vendre
PRENDERE	‘take’	prendere	prender	prendre
PERDERE	‘lose’	perdere	perder	perdre

## 6 Romance Languages

The normal Latin spelling shows no difference in stress, but in fact these infinitives reflect two different classes: **arhizotonic** (stressed off the root, e.g. DEBERE and HABERE) and **rhizotonic** (stressed on the root, e.g. VENDERE, PRENDERE, PERDERE).

Italian preserves the original contrast in stress pattern. Spanish has neutralized the contrast in favor of arhizotonic infinitives. French, however, shows the drastic consequences of the difference in stress. The stressed vowel in the arhizotonic infinitives gives <oi> /wa/, but what happens to that same vowel in the rhizotonic infinitives?

DEBERE	HABERE	VENDERE	PRENDERE	PERDERE
devo <u>ir</u>	avo <u>ir</u>	ve <u>nd</u> re	pre <u>nd</u> re	pe <u>rd</u> re

Having seen the import of stress, how do we determine where it fell in a Latin word? Good news: Latin has an easy, exceptionless stress rule known as the **Penultimate Rule**. To understand it (an absolute prerequisite for all of our subsequent work) you need to consider three concepts: vowel quantity, syllable boundaries, and syllable weight.

### 1.1.2 Latin vowels

Latin has a typologically common vowel system: five vowels arranged in the usual triangle, with front-back contrast neutralized for the low vowel /a/.

i ĭ	ū ŭ
e ĕ	o ō
ā ā	

The only complication is quantity – each of these five vowels can be either long or short, a phonemic contrast seen in such **minimal pairs** as:

ĔST ‘is’	VĔNIT ‘comes’
ĒST ‘eats’	VĒNIT ‘came’ [3rd sg.]

The vowel inventory also includes three diphthongs, written AE /aj/, OE /oj/, and AU /aw/. They count automatically as long vowels. The term **diphthong** means that the two elements (vowel + glide) belong to the same syllable. Caution! AE, OE, and AU are the only diphthongs in Latin.

### 1.1.3 Dividing syllables in Latin

More good news: English speakers understand most syllable boundaries intuitively. Stated in technical terms, the rule is: maximize the onset. This means: put as much material as possible into the later syllable, consistent with the requirement that its onset must be acceptable as the beginning of a word. Example: IM.PRIMA.TUR, IN.FER.NU (because MPR, NF, and RN are not possible onsets, and MP. R is wrong because the onset is not maximized).

## 7 The evolution of stressed vowels

**PRACTICE**

Indicate syllable boundaries in these Latin words.

LACRIMA      PARTE      CANTU      PETRA      HERBA      OPTIMU  
DECEMBER      ETERNU      MAXIMU (X stands for /ks/)

Sometimes you encounter two adjacent vowels (vowels in **hiatus**). Don't confuse these with diphthongs. Recall from § 1.1.2 that Latin has only three diphthongs: AE, OE, and AU. In other sequences written with two vowel graphemes <V V>, the two vowels belong to separate syllables. Example: SU.A.VE 'gentle', PA.LE.A 'straw'.

**PRACTICE**

Indicate where the syllable boundaries fall in these Latin words.

FEBRUARIUS      HODIE      PIETATE      VITREU      FILIA

In addition to long vowels, Latin has long (**geminate**) consonants formed by adding an extra timing unit to the closure. Even though in pronunciation there are not really two consonant releases, you obtain the correct results by placing the syllable boundary between the two written consonants (GUT.TA, VIL.LA, etc.).

The one unexpected fact about Latin syllabification concerns /s/ + consonant occurring word-medially. The syllable boundary in this case runs between /s/ and the consonant, which seems contrary to the principle of maximizing the onset. But even though /s/ + consonant can begin a word, it cannot begin a medial syllable. For example: SPA.TA, but CRIS.PA not \*CRI.SPA. This fact is inferred from clear linguistic evidence (§ 1.2.5). Interestingly, in some Romance languages /s/ + consonant proved to be an impossible onset even word-initially, as seen in Spanish *espada* 'sword' < SPATA and *escuela* 'school' < SCHOLA (§ 2.1.1).

**1.1.4 The Penultimate Rule**

This rule assigning word stress in Latin is stated in terms of syllable weight. Once you have identified the boundaries of a syllable, you have to determine whether it's heavy or light.

Definition: A syllable is **heavy** if it consists of two timing units. Otherwise stated, it contains a long vowel or ends in a consonant. Syllables meeting neither criterion are **light**.

**PRACTICE**

In these words you can't identify all the heavy syllables unless you know which vowels are long, but you can identify some on sight. Which ones?

MULTU      NOCTE      UNDECIM      CAELU      AURORA      AUDIO

Cambridge University Press

978-0-521-88915-5 - Romance Languages: A Historical Introduction

Ti Alkire and Carol Rosen

Excerpt

[More information](#)

## 8 Romance Languages

The Penultimate Rule stated below is so named because it looks at the next-to-last (**penultimate**) syllable. No other syllable is relevant. The rule says: if the penultimate syllable is heavy, stress it, otherwise stress the preceding syllable (the **antepenult**).

$$\begin{array}{cc} \sigma \underline{\sigma} \sigma \# & \underline{\sigma} \sigma \sigma \# \\ \text{H} & \text{L} \end{array}$$

The Penultimate Rule establishes a regular relationship between syllable weight and stress position. Given Latin vowel quantity, we can deduce stress position. Vice versa, if we know the stress position (from Romance reflexes), we can deduce the weight of the penultimate syllable in Latin.

**PRACTICE**

(Latin → Romance) In the following words Latin vowel quantity is indicated when long. Syllabify and show which syllable is stressed.

DIFFICILE	MULIER	MULIĒRE
INIMĪCU	AQUILA*	EXEMPLU
FORMĪCA	DIRECTU	ROTUNDU

\*QU counts as a consonant, /kʷ/.

**PRACTICE**

(Romance → Latin) Here vowel quantity is not shown. Try to infer the stress position from any Romance reflexes you may know, and from that, compute the weight of the penultimate syllable, and the vowel quantity when possible.

AMICA	‘friend’	OPERA	‘work’	NUMERU	‘number’
LACTUCA	‘lettuce’	CAMERA	‘room’	CATENA	‘chain’
INSULA	‘island’	FARINA	‘flour’	DOMINA	‘lady’
ANIMA	‘soul’	CONFLICTU	‘conflict’		

In words containing only two syllables, the penultimate is stressed unconditionally, revealing nothing about vowel quantity.

### 1.2 Stressed vowels: the (almost) pan-Romance seven-vowel system

The most consequential of all the pre-Romance sound changes was the refashioning of the stressed vowel system. The Latin system has five vowels and a





Latin stressed vowels and their outcomes in Italian, Spanish, and French

CRŪDU 'raw'  
 IŪRAT 'swears'  
 SCŪTU 'shield'  
 ACŪTU 'sharp'  
 MATŪRU 'ripe'  
 SECŪRA 'safe'

GULA	'throat'
IUVENE	'young'
AUGUSTU	'August'
CURRIT	'runs'
BUCCA	'mouth'
FURCA	'fork'
SAPŌRE	'taste'
MELIŌRE	'better'
FLŌRE	'flower'
HORA	'hour'
SOLU	'alone'
CŌRTE	'court'

PEDE 'foot'  
 \*MELE 'honey'  
 TENET 'holds'  
 HERBA 'grass'  
 TERRA 'earth'  
 PERDIT 'loses'  
 NOVU 'new'  
 FOCU 'fire'  
 OVU 'egg'  
 HOSPITE 'host'  
 MORTE 'death'  
 PORCU 'pig'

SALE 'salt'  
 MARE 'sea'  
 CAPUT 'head'  
 CAMPU 'field'  
 FLAMMA 'flame'  
 ARBORE 'tree'

AMĪCA 'friend'  
 DĪCIT 'says'  
 NĪDU 'nest'  
 SPĪNA 'thorn'  
 FORMĪCA 'ant'  
 SCRĪPTA 'written'

PILU	'hair'
VIDET	'sees'
BIBIT	'drinks'
CRISPA	'curly'
LITTERA	'letter'
VIRIDE	'green'
DĒBET	'owes'
PARIĒTE	'wall'
HABĒRE	'have'
*MĒSE	'month'
VĒLA	'sail'
TĒRNU	'triad'