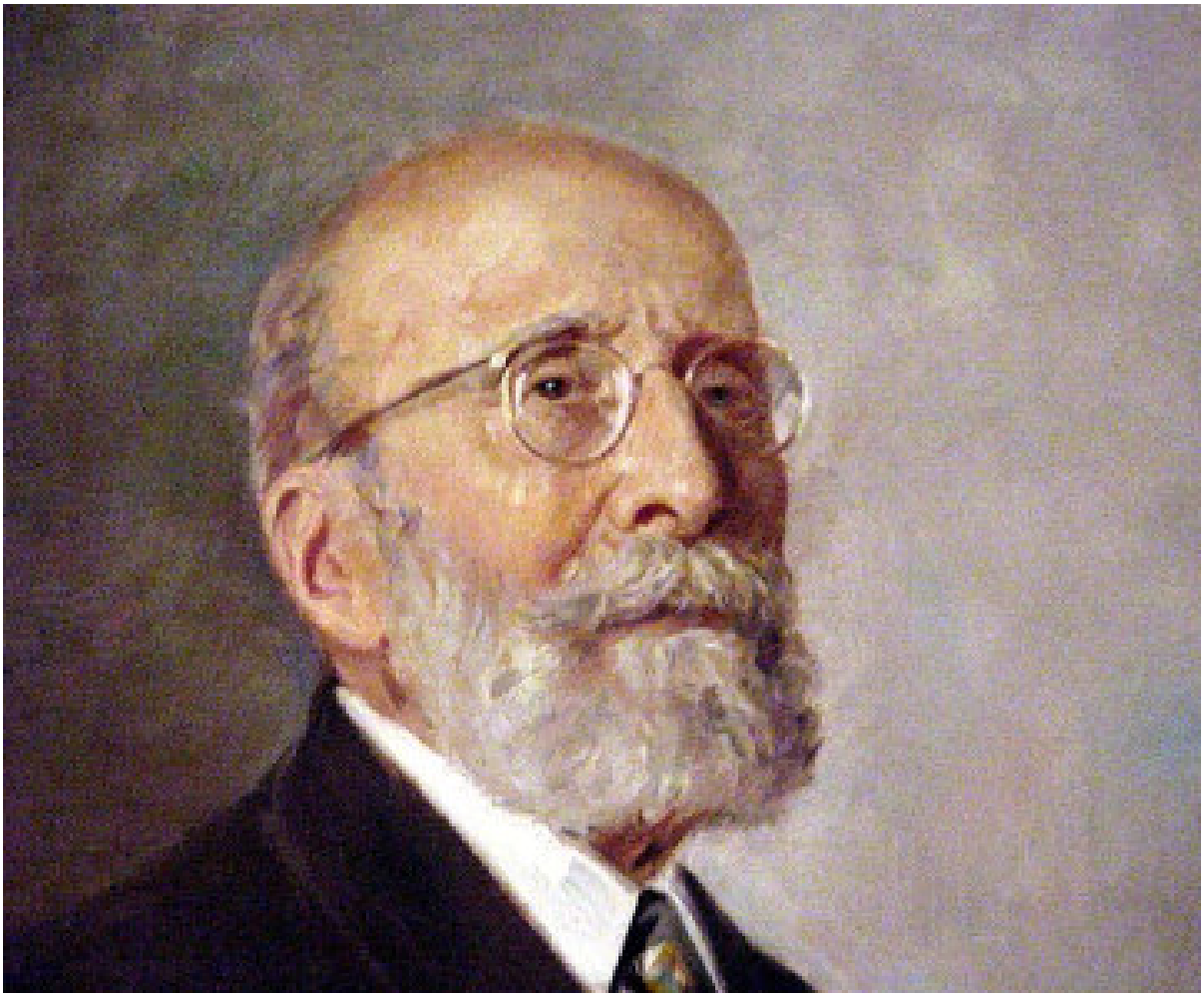
El sueño de mi juventud: Fonología automatizada

陳朝陽
Chen Zhaoyang
zc23@illinois.edu

November 10, 2022

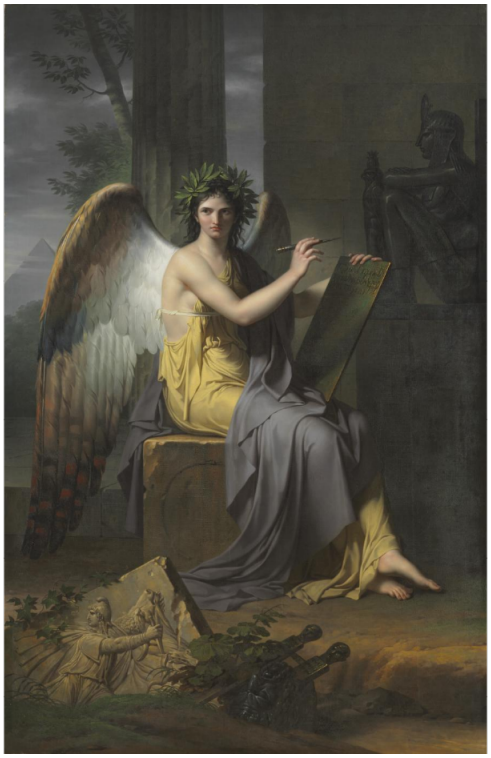


Ramón Menédez Pidal
1869 – 1968

Invocation of Clio[1]

I call to Clio, great of knowledge,
daughter of Zeus and wise Mnemosyne,
goddess who knows much of ancient days,
of what has been, and thus of what may be.
Great muse of history, you look to the past
and understand its import and its might;
you hold in your heart the lessons of time,
the wisdom that has passed through the world.
Through your power do we hear the words
of those long gone, do we receive their counsel.
Clio, chronicler of the good and the ill,
I pray to you, goddess, I honor your gifts.

[1] Underflow 2018.



Contents

Chapter 1

Introduccíon: España en nuestros corazones

Eternal vigilance is the price of knowledge.

George Santayana

1.1 *La lengua española*

1.2 Seasoned Techniques

枯れた技術の水平思考[1]

横井軍平

This paper mainly uses *seasoned* techniques in various fields – meaning that the fundamental works this project builds on have been around for quite some time.

1.2.1 Rule-based Phonology

The most fundamental technique used here, rule-based phonology, itself a term coined by its aftecomers, has been around as its *modern* form since at least late '60s.
At the beginning of the 20th century, phonology, and historical linguisitcs, started to be formalized. After many years of formalization, around and after the publication of *SPE*, the practice of phonology had long acquired a formal language, as we can see from the theoretic elegance and explanatory power of the phonology that *SPE* pioneered.

[1] "Lateral Thinking of Withered Technology."

1.2.2 Romance Historical Linguistics

The subject of study of this paper, Spanish, whose history belongs to the greater field of Romance historical linguistics, is one of the most ripened fruit of European historical linguistics since its birth. Countless literature are still being produced in this field. A couple of comprehensive monographs on the history of the Spanish language had been produced in English since the later half of the 20^{th} century e.g. Penny 2002 and Lloyd 1987, among many others. This work is not possible without their scholarship.

1.2.3 Typed Programming Languages

The last piece of the *puzle* ['puθ.le] is a *typed programming language*. What does it have to do with typed programming languages? Well, the answer to this question is one of the two themes of this paper, other than giving a concise account of the phonological history of the Spanish language; I want to show that, phonology, which is a formal system, can be simulated by a computer. And this system can be quite naturally implemented in a so-called *high-level programming language*.

1.3 The Implementation of a Phonology

In this paper we implement a kind of rule-based phonology that is capable of carrying out most of the historical sound changes from Latin to Modern Spanish. The underlying system and *Phonology* (with a capital "P") is not language-specific.
We have observed that rule-based phonology primarily has two components: the *statics*, in which rules are defined for constructing members of the surface representations in the phonology; and the *dynamics*, in which rules are set for defining the transformations between the aforementioned surface representations. In the lingo of computation, phonology, just like computation, revolves around *data* (segments, syllables etc.) and *computation* (sound changes).

Part I

Fonología automatizada

Chapter 2

The Little *Metaphoner*

ET COGNOSCETIS VERITATEM
ET VERITAS LIBERABIT VOS

IOANNES VIII•XXXII

Let us talk about the phonological notations and *pseudocode* conventions used in this paper. The notations used in this paper are not to obfuscate but to illuminate our thoughts; this is the standard that we hold ourselves to. Though one may still ask: why invent yet another notation/language for phonology? I would answer by quoting Geurts and Meertens 1978 through Gibbons 2020: p. 14:

> Suppose a textbook has to be written for an advanced course in algorithmics [or phonology]. Which vehicle should be chosen to express the algorithms [or sound changes]? Clearly, one has the freedom to construct a new language, not only without the restaint of efficiency considerations, but without any considerations of implementability whatsoever.

The pseudocode used in this paper (dubbed *Metaphono* in homage to the work of Hartman 2003b) is a *functional* vernacular (inspired by the spirit, less the essence, of Bird-Meertens formalism[1]), which is quite close to the syntax and semantics of real world programming languages like Haskell (Marlow 2010) and Standard ML (Milner et al. 1997). Here I would quote Geurts and Meertens 1978: p. 57 again to illustrate my understanding of *Metaphono*:

> This pidgin ALGOL [in our case, the ML-Haskell creole[2]] is a language. It is not really a programming, nor a natural language, but it has characteristics from both. It is not steady, but evolving. How it will evolve we cannot know.

In any regards, programming language theory is not the focus of this paper, whenever formalism becomes a burden in the clarification of our ideas, we take the liberty to explain them in a natural language.
The socratic style used in this chapter follows that of the book *The Little LISPer* by Friedman and Felleisen 1989.

[1] Bird and Meertens 1987; Bird and Moor 1997; Gibbons 2020.
[2] Maybe I should just call this language *Vulgar ML*.

2.1 *The Sound Patterns*

SEG ⊢ feat

What is [θ]? [θ] is a segment.
What is FRICATIVE? It is a *feature.*
Is θ a fricative? Yes, it is.
θ ⊢ fricative What does this mean?
It means that [θ] is a fricative. Is [θ] also some other things?
Yes:

θ ⊢ non-sibilant fricative
⊢ dental
⊢ voiceless

I feel like I have seen things of this sort before.
The ⊢ or ⊣ symbol can just be read as "is" or "exhibits the feature ..."; this is in line with the *SPE*[1] style notation $C_{[+feat]}$.
Thank you, Noam and Morris.

[1] Chomsky and Halle 1968.

[feat$_1$ → feat$_2$] SEG

What's the difference between [s] and [z]? [z] is voiced; while [s] is not.
Do they differ in anything else? No.
[voiceless → voiced] s ⇒ z What does this mean?
Change the *feature* of VOICING from *voiceless* to *voiced.* I see.

This notation has its roots in λ−calculi and other formal languages alike, this particular form is adopted from Pierce 2002. The semantics of this notation is simple: rewrite **feat**$_1$ to **feat**$_2$ within SEG.

2.2 The Vernacular

Types
val : τ

What is [ʃ]? [ʃ] is a consonant.
Is [ʃ] a vowel? No, it is not.
ʃ : Consonant What does this mean?
It means that [ʃ] is a consonant. What is an [a]?
a : Vowel So [a] is a vowel.

Functions
$\lambda : \tau \to \tau$

Functions are defined as such.

$$\begin{aligned} \text{raise} &: \text{Vowel} \to \text{Vowel} \\ \text{a} &\Rightarrow \text{e} \end{aligned}$$

Not so surprisingly, giving raise an [a] would yield an [e].

$$\text{raise}\ (\text{a} : \text{Vowel}) \Rightarrow \text{e} : \text{Vowel}$$

Pattern Matching
$\lambda\ x \Rightarrow x$

Consider this sound change for some hypothetical language that has {s z n m n̥ m̥}.

$$\begin{aligned} \text{occlude} &: \text{Consonant} \to \text{Consonant} \\ \text{C} \vdash \text{fricative} &\Rightarrow [\text{fricative} \to \text{stop}]\ \text{C} \\ \text{C} \vdash \text{nasal} &\Rightarrow \textbf{match}\ \text{voice}\ \text{C} \begin{cases} \text{voiced} & \Rightarrow [\text{nasal} \to \text{stop}]\ \text{C} \\ \text{voiceless} & \Rightarrow [\text{nasal} \to \text{stop};\ \text{unaspirated} \to \text{aspirated}]\ \text{C} \end{cases} \end{aligned}$$

$$\begin{aligned} \text{occlude s} &\Rightarrow \text{t} \\ \text{occlude z} &\Rightarrow \text{d} \\ \text{occlude n} &\Rightarrow \text{d} \\ \text{occlude m} &\Rightarrow \text{b} \\ \text{occlude n̥} &\Rightarrow \text{t}^{\text{h}} \\ \text{occlude m̥} &\Rightarrow \text{p}^{\text{h}} \end{aligned}$$

$$\begin{aligned} &\lambda \ : \ \tau \to \tau \\ &\text{pattern}_0 \Rightarrow \text{exp}_0 \\ &\vdots \\ &\text{pattern}_\omega \Rightarrow \text{exp}_\omega \end{aligned}$$

$$\textbf{match exp} \begin{cases} \text{pattern}_0 & \Rightarrow \text{exp}'_0 \\ & \vdots \\ \text{pattern}_\omega & \Rightarrow \text{exp}'_\omega \end{cases}$$

Valence
$\lambda : \tau \to \tau \to \tau$

Higher-Order Functions
$(\lambda : (\tau \to \tau) \to \tau \to \tau)\ (\lambda' : \tau \to \tau)$

Production Rules

Common Types

Functor

Chapter 3

In the Beginning was the *Word*: Segments, Syllables, and Lexemes

> En el principio existía la Palabra y la Palabra estaba con Dios, y la Palabra era Dios.
>
> Juan 1:1

3.1 Representation of Segments

3.1.1 Vocalic

```
vowel ::= height × centrality × duration

height ::= low | low-mid | mid | high-mid | high

centrality ::= front | central | back

duration ::= short | long
```

	Front	Central	Back
High	i		u
High-Mid	ɪ		ʊ
Mid	e		o
Low-Mid	ɛ		ɔ
Low		a	

3.1.2 Consonantal

```
consonant ::= manner × place × voice
```

```
manner ::= nasal
| stop
| fricative
| non-siblant fricative
| affricate
| approximant
| tap
| trill
| lateral

place ::= bilabial
| labiodental
| dental
| alveolar
| palatal
| velar | labiovelar
| glottal

place ::= voiced | voiceless
```

3.2 Representation of Syllables and Lexemes

3.2.1 Syllabic Structure

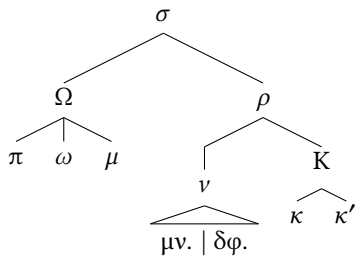
```
syllable ::= onset rhyme

onset ::= zero-onset
| monoonset of consonant
| dionset of consonant_1 consonant_2
| trionset of consonant_1 consonant_2 consonant_3

rhyme ::= nucleus coda

nucleus ::= zero-nucleus | monophthong of vowel | diphthong of vowel_1 vowel_2

coda ::= zero-coda
| monocoda of consonant
| dicoda of consonant_1 consonant_2
```



Under this schema, the utterance-initial onset clusters would like this in Modern Spanish[1]:

ω	*μ*
p β t d k$^{(w)}$ g$^{(w)}$ f θ s x ʝ ʧ m n ɲ l ʎ ɾ	∅
p b t d k g f	ɾ
p b (t) k g f	l

And the utterance-final coda clusters[2]:

κ	*κ′*
p t d k f θ s j n m l ɾ	∅
p t k f n l ɾ	s
(n)	(θ)

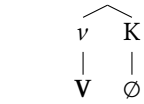
[1]Hualde 2005: p. 74.
[2]Hualde 2005: p. 75.

3.2.2 Syllable Weight and Stress

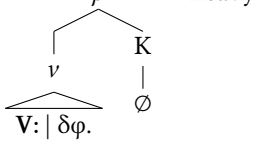
Before we can discuss the stress pattern in Latin, it is necessary that we define a function to decide the weight of a syllable in Latin.

$$\text{weight} \;:\; \text{Syllable} \to \text{Weight}$$

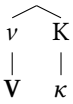
$\sigma \vdash \rho \Rightarrow$ light



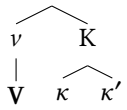
$\sigma \vdash \rho \Rightarrow$ heavy



$\sigma \vdash \rho \Rightarrow$ heavy



$\sigma \vdash \rho \Rightarrow$ heavy



To have a concrete idea, see the following examples.

$$\begin{aligned}
\text{weight } \text{MI}_{\dashv \text{SIMILIS}} &\Rightarrow \text{light} \\
\text{weight } \text{SŌ}_{\dashv \text{PERSŌNA}} &\Rightarrow \text{heavy} \\
\text{weight } \text{MEN}_{\dashv \text{FUNDĀMENTUM}} &\Rightarrow \text{heavy} \\
\text{weight } \text{DOC}_{\dashv \text{PARADOXUS}} &\Rightarrow \text{heavy}
\end{aligned}$$

After we are able to decide syllable weight, defining the Latin PENULTIMATE stress rule becomes natural.

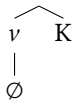
$$\begin{aligned}
\text{assign stress} \;:\; & [\text{Syllable}] \to \text{Lexeme} \\
[\sigma] \Rightarrow\; & \textbf{let } \sigma' = [\text{unstressed} \to \text{stressed}]\ \sigma \\
& \textbf{in } [\sigma'] \\
[\sigma_1 :: \sigma_2] \Rightarrow\; & \textbf{let } \sigma' = [\text{unstressed} \to \text{stressed}]\ \sigma_1 \\
& \textbf{in } [\sigma' :: \sigma_2]
\end{aligned}$$

$$[\sigma^*] \mathbin{+\!\!+} [\sigma_1 :: \sigma_2 :: \sigma_3] \Rightarrow \textbf{match } \text{weight } \sigma_2 \begin{cases} \text{heavy} & \Rightarrow \textbf{let } \sigma' = [\text{unstressed} \to \text{stressed}]\ \sigma_2 \\ & \quad \textbf{in } [\sigma^*] \mathbin{+\!\!+} [\sigma_1 :: \sigma' :: \sigma_3] \\ \text{light} & \Rightarrow \textbf{let } \sigma' = [\text{unstressed} \to \text{stressed}]\ \sigma_1 \\ & \quad \textbf{in } [\sigma^*] \mathbin{+\!\!+} [\sigma' :: \sigma_2 :: \sigma_3] \end{cases}$$

3.2.3 Syllabification

syllable empty? : Syllable → Bool

$\sigma \vdash \quad \rho \quad \Rightarrow$ True



$\_ \Rightarrow$ False

Chapter 4

Tilting against Windmills: Sound Changes and Phonotactics

星翳燈幻露泡夢電雲[1]

金剛經 §32

4.1 Strata of Sound Change

Wir müssen wissen.
Wir werden wissen.

David Hilbert

4.1.1 Feature and Segmental Level

In sections *The Sound Patterns* and The Vernacular we introduced some notations that rewrites features of a segment:

$$[\text{voiceless} \rightarrow \text{voiced}]\ \text{s} \Rightarrow \text{z}$$

[1] A shooting star, a clouding of the sight, a lamp, an illusion, a drop of dew, a bubble, a dream, a lightning's flash, a thunder cloud. Diamond Sutra §32

and segments themselves:

$$\begin{aligned} \text{devoicing} &: \text{Consonant} \to \text{Consonant} \\ \text{s} &\Rightarrow \text{z} \end{aligned}$$

$$\begin{aligned} \text{raise} &: \text{Vowel} \to \text{Vowel} \\ \text{a} &\Rightarrow \text{e} \\ \text{o} &\Rightarrow \text{u} \end{aligned}$$

These two schema constitutes the **feature** and **segmental** level of our language that defines sound changes. These two kinds of formulations by and large overlapps with each other. This expression

$$[\text{voiceless} \to \text{voiced}]\ \text{s} \Rightarrow \text{z}$$

is equivalent to the devoicing function defined above. Thus we would use these two notations interchangeably depending on what our emphasis is in the respective sound changes. But in cases where a consonant is changed to a vowel and vice versa, one has to rewrite the segments directly:

$$\begin{aligned} \text{vocalization} &: \text{Consonant} \to \text{Vowel} \\ \text{ɬ} &\Rightarrow \text{u} \end{aligned}$$

$$\begin{aligned} \text{fortition} &: \text{Vowel} \to \text{Consonant} \\ \text{u} &\Rightarrow \text{ʊ} \end{aligned}$$

4.1.2 Syllabic Level

Elision is among the most common sound changes found in languages. But to formulate it can be slightly tricky:

$$\begin{aligned} \text{elision} &: \text{Consonant} \to ? \\ \text{m} &\Rightarrow \emptyset \end{aligned}$$

We find ourselves unable to type this function: $\emptyset$ is not a consonant, it is not even a **segment**. Same difficulty is found when we try to define epenthesis:

$$\begin{aligned} \text{epenthesis} &: ? \to \text{Vowel} \\ \emptyset &\Rightarrow \text{ə} \end{aligned}$$

This time we have to even accept a $\emptyset$ as an input – only that there is no segmental $\emptyset$ in an utterance. Sound changes of this sort clearly operates in a different level than the ones we talked above.

An implementation of the elison of [m] in the coda position would look like this on the syllabic level:

$$\begin{aligned} &\text{elision} \;:\; \text{Syllable} \to \text{Syllable} \\ &\sigma \;\vdash\; \rho \;\Rightarrow\; \begin{array}{l}\textbf{let}\ \rho' = [\mathrm{m} \to \emptyset^{2}]\ \rho \\ \textbf{in}\ [\rho \to \rho']\ \sigma\end{array} \end{aligned}$$

ρ: ν (K), K → κ → m

Here is how it would be applied to a Latin word such as LUCEM:

$$\begin{aligned} &\lambda \;:\; \text{Lexeme} \to \text{Lexeme} \\ &[\sigma :: \sigma'] \Rightarrow [\sigma :: \text{elision}\ \sigma'] \end{aligned}$$

$$\lambda\ \textsc{lucem} \Rightarrow \textsc{luce}$$

4.1.3 Metrical Level

No sound change can be applied below the level of phonological word. In fact, we shall give *sound change* a type synonym as such:

$$\text{sound change}_{\tau} \coloneqq \text{Lexeme} \to \text{Lexeme}$$

$$[\ \bar{\sigma}\ \sigma\ \sigma\ ] \Rightarrow [\ \bar{\sigma}\ \sigma\ ]$$

$\bar{\sigma}$: $\bar{\text{O}}$ CU LU → $\bar{\text{O}}$ CLU

4.2 Representation of Sound Changes

4.2.1 Sound Change Patterns

4.2.1.1 Elision

Elision of some word-final consonant is a common sound change that shows up again and again in historical linguistics.
To see a couple examples, we construct the elision of a syllable final [m] as such:

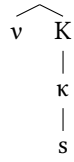
$$\begin{aligned} &\text{elision}_{\mathrm{m}} \;:\; \text{Syllable} \to \text{Syllable} \\ &\sigma \;\vdash\; \rho \;\Rightarrow\; \begin{array}{l}\textbf{let}\ \rho' = [\mathrm{m} \to \emptyset]\ \rho \\ \textbf{in}\ [\rho \to \rho']\ \sigma\end{array} \end{aligned}$$

ρ: ν (K), K → κ → m

[2] This ∅ is `zero-coda`, cf. Syllabic Structure.

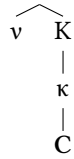
similarly, the elision of syllable final [s]:

$$\begin{array}{rcl} & \text{elision}_{\text{s}} & : \ \text{Syllable} \to \text{Syllable} \\ \sigma \ \vdash \ \rho & \Rightarrow & \textbf{let } \rho' = [\text{s} \to \emptyset] \ \rho \\ & & \textbf{in } [\rho \to \rho'] \ \sigma \end{array}$$



One may observe that these two sound changes differ only by the what kind of consonant they elide. As we can see from the following template that elides an arbitrary consonant C (although this is not a grammatical function):

$$\begin{array}{rcl} & \text{elision}_{\text{C}} & : \ \text{Syllable} \to \text{Syllable} \\ \sigma \ \vdash \ \rho & \Rightarrow & \textbf{let } \rho' = [\text{C} \to \emptyset] \ \rho \\ & & \textbf{in } [\rho \to \rho'] \ \sigma \end{array}$$



A grammatical function would have to take another argument, which is the segment that is to be elided:

$$\begin{array}{rcl} \text{elide}_{\text{cons}} & : & \text{Consonant} \to \text{Syllable} \to \text{Syllable} \\ \boxed{\text{cons}} \ \boxed{\sigma} \Rightarrow \textbf{match } \sigma \parallel \sigma \vdash [_{\rho} \ \nu \ [_{\text{K}} \ [_{\kappa} \ \text{cons} \ ] \ ] \ ]^{3} \Rightarrow & & \textbf{let } \rho' = [\text{cons} \to \emptyset] \ \rho \\ & & \textbf{in } [\rho \to \rho'] \ \sigma \end{array}$$

One can see that $\text{elide}_{\text{cons}}$ is able to construct arbitrary elision sound changes that elides any consonant at the end of a syllable:

$$\begin{array}{rcl} \text{elide}_{\text{cons}} \ \text{m} \equiv \text{elide}_{\text{m}} & : & \text{Syllable} \to \text{Syllable} \\ \text{elide}_{\text{cons}} \ \text{s} \equiv \text{elide}_{\text{s}} & : & \text{Syllable} \to \text{Syllable} \end{array}$$

But it still has one last problem: $\text{elide}_{\text{cons}}$ only constructs sound changes that rewrites a **syllable**, not a phonological word. Thus it is not really useful since not all words in the lexicon are monosyllabic words; we need to construct sound changes that elide some **word-final** consonant instead of only syllable final consonant. It turns out that we are actually not too far away from this goal, only two functions away:

$$\begin{aligned} \text{elide}_{\sigma} &: (\text{Syllable} \to \text{Syllable}) \to \text{Lexeme} \to \text{Lexeme} \\ \boxed{f} \; \boxed{[\sigma^*] \mathbin{+\!\!+} [\sigma']} &\Rightarrow [\sigma^*] \mathbin{+\!\!+} [f\ \sigma'] \end{aligned}$$

$$\begin{aligned} \text{elide}_{\lambda} &: \text{Consonant} \to (\text{Lexeme} \to \text{Lexeme}) \\ \text{cons} \Rightarrow &\ \textbf{let}\ f^{4} = \text{elide}_{\text{cons}}\ \text{cons} \\ &\ \textbf{in}\ \text{elide}_{\sigma}\ f \end{aligned}$$

Let us break it down bit by bit. First, recall that the **let**-expression is equivalent to substitution:

$$[\![\ \textbf{let}\ \text{exp} = \text{exp}'\ \textbf{in}\ \text{exp}\ ]\!] \coloneqq [\text{exp} \to \text{exp}']\ \text{exp} \Rightarrow \text{exp}'$$

Thus, the function elide_{λ}, whose meaning may seem to be obscured at first glance, is a shorthand for this expression, using the $[\text{exp} \to \text{exp}']$ notation:

$$\text{elide}_{\lambda} \coloneqq [\text{f} \to \text{elide}_{\text{cons}}\ \text{cons}]\ \text{elide}_{\sigma}\ f$$

or, it can also look like this:

$$\text{elide}_{\lambda}\ \text{cons} \Rightarrow \text{elide}_{\sigma}(\text{elide}_{\text{cons}}\ \text{cons})$$

A concrete example of (a) constructing a sound change that elides word-final [t] and (b) applying the sound change to a word in one expression:

$$\text{elide}_{\lambda}\ \text{t}\ \textsc{bibit}^{5} \Rightarrow \textsc{bibi}$$

To see a more rigorous picture, I have provided the following proof tree for this expression:

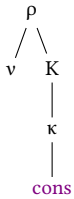
$$\dfrac{\dfrac{\text{elide}_{\lambda} : \text{Consonant} \to \text{Lexeme} \to \text{Lexeme} \qquad \text{t} : \text{Consonant}}{\text{elide}_{\lambda}\ \text{t} : \text{Lexeme} \to \text{Lexeme}} \qquad \textsc{bibit} : \text{Lexeme}}{(\text{elide}_{\lambda}\ \text{t})\ \textsc{bibit} : \text{Lexeme}}$$

The proof tree above is not the whole picture, here I provide a complete proof tree for the expression

$$\text{elide}_{\lambda}\ \text{cons} \Rightarrow \text{elide}_{\sigma}(\text{elide}_{\text{cons}}\ \text{cons})$$

after expansion:

[3]This tree would look like this in full expansion:



[4] f : Syallble → Syllable.
[5]Internally looks like [bí :: bit].

$$\frac{\mathrm{elide}_{\sigma} : (\mathrm{Syllable} \to \mathrm{Syllable}) \to \mathrm{Lexeme} \to \mathrm{Lexeme} \qquad \dfrac{\mathrm{elide}_{\mathrm{cons}} : \mathrm{Consonant} \to \mathrm{Syllable} \to \mathrm{Syllable} \qquad t : \mathrm{Consonant}}{\mathrm{elide}_{\mathrm{cons}}\ t : \mathrm{Syllable} \to \mathrm{Syllable}}}{\mathrm{elide}_{\sigma}\ (\mathrm{elide}_{\mathrm{cons}}\ t) : \mathrm{Lexeme} \to \mathrm{Lexeme}}$$

4.2.1.2 Vowel Shifts

4.2.1.3 Intervocalic Lenition

4.2.1.4 Apocope

4.2.1.5 Syncope

4.2.2 Context of Sound Changes

4.2.3 Application of Sound Changes

4.2.3.1 Composing Sound Changes

luz : Lexme

luz ≔ LŪCEM

≫ LUCEM ◁ Loss of Vowel Quantity

≫ LUCE ◁ Elision of [m]

≫ luʧe ◁ Palatalization and Affrication of Velars

≫ lutse ◁ Fronting of Palatal Affricates

≫ luʤe ◁ Lenition I

≫ luʤ ◁ Apocope

≫ luz̪ ◁ Deaffrication of Dental Affricates

≫ lus̪ ◁ Devoicing of Fricatives

≫ luθ ◁ Desibilation of Dental Fricatives

4.2.4 Admissibility of Sound Changes

4.3 Representation of Phonotactics

4.4 Rule Ordering

> Die ganzen Zahlen hat der liebe Gott
> gemachet, alles andere ist Menschenwerk.
>
> Leopold Kronecker

Recall that the natural numbers look like this: $\mathbb{N} \coloneqq 0 \sqsubset 1 \sqsubset 2 \sqsubset 3 \sqsubset \cdots \sqsubset \omega$.

no.	
0	Deaffrication of Dental Affricates
1	Devoicing of Fricatives
2	Desibilation of Dental Fricatives

Part II

Los elementos de Fonología histórica española

Chapter 5

Misc. Remarks and Comments

5.1 Sources for Lat.-Rom.-Es. Gloss.

The majority of the words in examples in this paper comes from Alkire and Rosen 2010, Lloyd 1987, and Penny 2002.

5.2 Abbreviations for Languages

lingua.	abbrv.
LATINA	LAT.
Proto-Romance	PrRom.
Western Romance	WRom.
Germanic	Gm.
Old Spanish	OSp.
Modern Spanish	Es.
Portugese	Port.
Catalan	Cat.
French	Fr.
Italian	It.

5.3 Organization of Each Chapter

In the beginning of each chapter, we give our reconstruction of the **Segmental Inventory** of the language in focus. In the first chapter on Proto-Romance, we also list the reconstructed Latin segments. After the listing of the inventory, the rest of each chapter is dedicated to the sound changes from the previous stage of the language to its current stage. Regrettably, not all sound changes in the history of the Spanish language are listed in this paper – not even close.

5.4 Criteria for Selecting Sound Changes

Only a relatively small portion of all the sound changes in the history of Spanish were listed in this paper, and the criteria for selecting each sound change should be justified. Our main aims behind the selection of each sound change is (a) to cover the widest variety of sound changes possible, (b) to implement sound changes that operates on the syllabic and metrical level, whose representation in causal phonological notations always seemed awkward, and (c) to identify sound changes that form chain shifts or ordinally related and implement them in a way that preserves their rule interactions.

5.5 The Final Result

At the very end of this part, we want to roughly derive a very conservative Northern-Central Peninsular dialect of Castilian.


VERITAS
BIBLIA SACRA
CÆLESTI
PELAGIVS
IVLIANVS

Chapter 6

Proto-Romance

FINIS ORIGINE PENDENT

MARCVS MANILIVS

Proto-Romance, among other terms like *Vulgar Latin* and *Popular Latin*, denotes a reconstructed language that evolved into the surviving Romance languages.

6.1 Segmental Inventory

6.1.1 Vocalic

Proto-Romance Monophthongs

	Front	Central	Back
High	i		u
Mid	e		o
Low-Mid	ɛ		ɔ
Low		a	

6.1.2 Consonantal

Proto-Romance Consonants

	Bilabial	Dental	Alveolar	Palatal	Velar	Labiovelar
Nasal	m		n	ɲ		
Stop	p b	t d			k g	kʷ gʷ
Fricative	ɸ β			j		
"			s z			
Affricate		ʦ ʣ		ʧ ʤ		
Trill			r			
Lateral			l	ʎ		

6.2 Sound Changes

6.2.1 Vocalism

6.2.1.1 Loss of Vowel Quantity

	Front		Cent.		Back	
High	ĭ	ī			ŭ	ū
Mid	ĕ	ē			ŏ	ō
Low			Ă	Ā		

⇒

	Front	Central	Back
High	i		u
High-Mid	ɪ		ʊ
Mid	e		o
Low-Mid	ɛ		ɔ
Low		a	

Latin had a 5 vowel system with each vowel can be either long or short (Alkire and Rosen 2010: p. 6; Penny 2002: p. 44; Lloyd 1987: p. 70). Here are some minimal pairs showing the phonemic vowel length[1]:

HĪC	'here'	HIC	'this'
LĪBER	'free'	LIBER	'book'
LĒVIS	'smooth'	LEVIS	'light in weight'
VĒNIT	'he came'	VENIT	'he comes'
MĀLUM	'apple'	MALUM	'evil, misfortune'
ŌS	'mouth'	OS	'bone'
PŌPULUS	'white poplar'	POPULUS	'people'

The distinction in vocalic quantity eventually collapsed (Lloyd 1987: p. 112):

> [I]n the fifth century St. Augustine remarked that in Africa people could not distinguish between long and short vowels. We may assume that he was referring to literary Latin since most of our evidence would indicate that spoken Latin had discarded quantity as a phonological feature long before then.

[1] Penny 2002: p. 45.

The long vowels [i:], [u:], [e:], and [o:] did not change in quality, as it can be reflected in these Spanish reflexes[2]:

LATINA	Español
VĪTA	vida
VICĪNA	vecina
FARĪNA	harina
LŪNA	luna
DŪRA	dura
MŪRU	muro
HŌRA	hora
CŌRTE	corte
DĒBET	debe
TĒRNU	terno

The long [a:] merges with the short [a]; while the short vowels [i], [u], [e] and [o] would experience lowering (Alkire and Rosen 2010: p. 13). As an interesting sidenote: in Sardinian, long and short vowels merge without changes in quality,yielding a five-vowel system (Lloyd 1987: p. 112).

Metaphono

$$\text{loss of quantity} \;:\; \text{Vowel} \rightarrow \text{Vowel} \qquad \text{(Loss of Quantity)}$$

$$\mathbf{V} \vdash \text{long} \Rightarrow [\text{long} \rightarrow \text{short}]\ \mathbf{V}$$

$$\mathbf{V} \vdash \text{short} \Rightarrow \textbf{match V} \begin{cases} \mathbf{V} \vdash \text{low} & \Rightarrow \mathbf{V} \\ \mathbf{V} \vdash \text{mid} & \Rightarrow [\text{mid} \rightarrow \text{low-mid}]\ \mathbf{V} \\ \mathbf{V} \vdash \text{high} & \Rightarrow [\text{high} \rightarrow \text{high-mid}]\ \mathbf{V} \end{cases}$$

6.2.1.2 Great Merger

	Front	Central	Back
High	i		u
High-Mid	ɪ		ʊ
Mid	e		o
Low-Mid	ɛ		ɔ
Low		a	

⇒

	Front	Central	Back
High	i		u
Mid	e		o
Low-Mid	ɛ		ɔ
Low		a	

Later on, the vowel system of Romance languages continues to evolve, reducing the 9 vowel system shown above to a 7 vowel system (Alkire and Rosen 2010: p. 13): the high-mid vowels [ɪ] and [ʊ] would merge with their mid counterparts [e] and [u][3]:

[2] Alkire and Rosen 2010: p. 12.
[3] Alkire and Rosen 2010: p. 14.

LATINA	PrRom	Español
GULA	[ʊ]	gola
CURRIT	[ʊ]	corre
MŬSCA	[ʊ]	mosca
BIBIT	[ɪ]	bebe
LITTERA	[ɪ]	letra
VĬCE	[ɪ]	vez

For the low-mid vowels [ɛ] and [ɔ], both of them diphthongizes in either open or closed syllables in Spanish (Alkire and Rosen 2010: p. 15-16), later on in Diphthongization I and Diphthongization II we show the diphthongization of these vowels in details.

great merger : Vowel → Vowel (Great Merger)

ɪ ⇒ e

ʊ ⇒ o

6.2.1.3 Merger in Atonic Vowels

	Front	Central	Back
High	i		u
Mid	e		o
Low-Mid	ɛ		ɔ
Low		a	

⇒

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

So far in both Loss of Vowel Quantity and Great Merger we mainly discussed the outcomes of **stressed** (*tonic*) Latin-Romance vowels. The atonic vowels showed greater degrees of reduction than their tonic counterparts (Lloyd 1987: p. 113):

> The result [of the vowel system reduction in atonic vowels] was the coalescence of these [high and mid] vowels into two phonemes only: /i, e:, e/ > /e/, and /u, o:, o/ > /o/. Thus for vowels in these [unstressed] syllables a five vowel system was established[.]

Some examples for the outcomes of atonic vowels[4]:

LAT.	Es.
HĪBERNU	iviern**o**
CIRCĀRE	c**e**rcar
VĒNĀTU	venad**o**

[4] Lloyd 1987: p. 113.

atonic merger[5] : Vowel → Vowel (Atonic Merger)

ɛ ⇒ e

ɔ ⇒ o

6.2.1.4 Monophthongization

OE ⇒ e

AU ⇒ o | a

AE ⇒ ɛ | e

There were 3 diphthongs in Classical Latin that underwent monophthongization in its evolution to Popular Latin/Proto-Romance: AE, OE, and AU. Although there were other diphthongs inherited from Indo-European (EI and OV) that had been merged with long vowels in the Old Latin period (Clackson 2011: p. 18):

> [F]rom the mid-second century BCE the digraphs EI and OV, which earlier spelled inherited diphthongs, spelled the long vowels /iː/ and /uː/ respectively, regardless of their etymological source, e.g. [V]EIVAM /wiːwam/ "living" (*CIL* I^2.1837), COVRAVERVNT /kuːraːweːrʊnt/ "oversaw" (*CIL* I^2.1806).

Here we focus on the 3 diphthongs that survived into Classical Latin.
The diphthong AE mainly reflects a short [ɛ][6]:

LAT.	Es.
CAECU	ciego
CAELU	cielo
QUAERO	quiero

In some cases it would give a long [eː][7]:

[5]This function applies after Loss of Quantity and Great Merger; but this is not **rule ordering**. The *order of evaluation* is only important in this case because these vowel mergers are implemented in this particular way for *convenience* in our system.
[6]Lloyd 1987: p. 105.
[7]Lloyd 1987: p. 105.

LAT.	Es.
CAESPITE	césped
FAECE	hez
FAENU	heno
PRAEDA	prea
SAEPE	sebe
SAEPTU	seto
SAETA	seda
TAEDA	tea

The diphthong OE would evolve into a long [e:][8] e.g. FOEDUS → Es. feo[9]:

PHOEBUS	PHEBUS
COENA	CĒNA
POENA	PENA

And AU gives a long vowel [o:][10]:

LAT.	Es.
AURU	oro
THESAURU	tesoro
PAUPERU	pobre
PAUCU	poco

monophthongization : Vowel → Vowel (Monophthongization)

OE ⇒ e

AU ⇒ o | a[11]

AE ⇒ ɛ | e[12]

[8] Lloyd 1987: p. 106.

[9] Alkire and Rosen 2010: p. 23.

[10] Alkire and Rosen 2010: p. 24.

[11] AU monophthongizes to [a] only when followed by a labialized velar e.g. AUGUSTO → AGUSTO → Es. *agosto*, AUSCULTĀRE → ASCULTĀRE → OSp. *ascuchar*, and AUGURIUM → AGURIUM → Es. *agüero* (Lloyd 1987: p. 107).

[12] AE giving [e:] is sporadic; sporadic changes like this in practice are implemented as another function (sound change); for simplicity's sake we list them like as if they are the same function.

6.2.2 Consonantism

Latin Consonants

	Bilabial	Dental	Alveolar	Palatal	Velar	Labiovelar	Glottal
Nasal	m		n				
Stop	p b	t d			k g	k^w g^w	
Fricative	ɸ						h
"			s z				
Approximant				j		w	
Trill			r				
Lateral			l				

6.2.2.1 Betacism I

$$w \Rightarrow \beta$$

The articulation of Latin [w] was strengthened to [β] in Proto-Romance[13]:

> [F]rom the first century A.D. onward, we find confusion of the letters V and B in many written documents, e.g., IUVENTE for IUBENTE (first century), IUBENTUTIS for IUVENTUTIS (155 A.D.), LIVERTUS for LIBERTUS (207 A.D.), etc[.]

In Italian and French, the Latin [w] is realized as a labiodental [v], while in Spanish it collides with [b][14] (for more details on Spanish *betacismo*, cf. Betacism II and Lenition II):

LAT.	Es.	Fr.	It.
VACCA	vaca [b]	vache	vacca
VINU	vino [b]	vin	vino

This sound change later on would make the original intervocalic [w] from Latin collide with the intervocalic [b], due to lenition (cf. Lenition I):

$$-b- \Rightarrow \beta$$

The merger of intervocalic [w] and [b] is attested in Spanish, French, and Italian[15]:

LAT.	Es.	Fr.	It.
CABALLU	caballo [β]	cheval	cavallo
BIBAT	beba [β]	boive	beva
DEBERE	deber [β]	devoir	devere
CAVERNA	caverna [β]	caverne	caverna
LAVARE	lavar [β]	laver	lavare
VIVUNT	viven [β]	vivent	vivono

Likewise, in Visigothic documents, the spelling exhibits the "[c]onfusion of *u* and *b* (*haueo* for *habeo*) or vice versa (*pabor* for *pauor*)"[16].

[13] Lloyd 1987: p. 132.
[14] Alkire and Rosen 2010: p. 32-33.
[15] Alkire and Rosen 2010: p. 33.
[16] Coulson and Babcock 2020: p. 159.

Diphthongs from [w]

Strengthening of [w] in Germanic Loanwords

Gm.	Es.
wisa	guisa
werra	guerra
warten	guardar
want	guante

(PrBsq. *euscala →) LAT. VASCONEM → Es. Gascón

Delabialization of Postconsonantal [w] Other than yods, dehiaticization also results in postconsonantal [w] that was subject to deletion[17]:

LAT.	Es.
BATTUŌ	bato
MORTUU	muerto
QUATTUOR	cuatro
FEBRUĀRIU	febrero
CONSUERE	coser

This [w] deletes after velars as well, even though labiovelars [kʷ] [gʷ] were part of the Latin-Romance inventory: ANTĪQUU → Es. *antigo*, QUŌMODŌ → Es. *como* (Lloyd 1987: p. 135). When the labiovelar is followed by an [a], the [w] can be preserved[18]:

LAT.	Es.
QUANDŌ	cuando
QUANTU	cuanto
QUĀLE	cual
AQUA	agua
AEQUĀLE	igual

6.2.2.2 Yod Fortition

j- (⇒ ʝ– ⇒ ɟ–) ⇒ ʤ–
-j- (⇒ –ʝ– ⇒ –ɟ–) ⇒ –ʤ–

LAT.	Es.	
IOCU	juego	[x]
IVIS	jueves	[x]
IUVENE	joven	[x]
IUDICE	juez	[x]

[17] Lloyd 1987: p. 134-135.
[18] Lloyd 1987: p. 135.

LAT.	Es.
MAIUS	mayo
MAIORE	mayor
CUIUS	cuyo

LAT.	Es.
IACET	yace
IAM	ya

Visigothic document witnesses "[t]he use of *g* instead of *j* [...] (*magor*) or *i* instead of *g* (*ienitor*)", cf. MAIOR, GENITOR (Coulson and Babcock 2020: p. 159).

6.2.2.3 Palatalization and Affrication of Dentals

ti ⇒ ʦ
di ⇒ ʤ

The voiceless dental [t] followed by a yod started to show signs of palatalization and affrication "[i]n the second century A.D. inscriptions begin to reveal this pronunciation: CRESCENTSIANUS for CRESCENTIANUS (A.D. 140), VINCENTZA for VINCENTIA, LAURENTZIO for LAURENTIO, MARSALIS for MARTIALIS (4th cent.)"[19]. There are also hints that for some speakers of Latin-Romance at that time TI- and CI- sounded quite close, as these two are sometimes confused in inscriptions: "MUNDICIEI for MUNDITIEI (A.D. 136), TERCIAE for TERTIAE (A.D. 179), FATIO for FACIO, NUNCIARE for NUNTIARE, DEFINICIA for DEFINITIO (A.D. 222-35)"[20].
As for the voiced dental [d] plus the yod, it collides with the palatalized voiced velar GI- → [ʤ][21] (cf. Palatalization and Affrication of Velars) i.e. LAT. DIURNATA → Es. jornada [x]. "Inscriptional evidence for the merger is found from the first century onward: AIUTOR for ADIUTOR, AIUTORIS for ADIUTRIX, IOSIMUS for ZOSIMUS (Gr. Ζώσιμος), ZIOMEDIS for DIOMEDIS, ZEBUS for DIEBUS, ZONISIUS for DIONISIUS (beginning of third century), MADIA for MAIAS (A.D. 364), ZABULUS for DIABOLUS[22]".
Sound changes of this sort is less regular in Castilian and other Western Romance languages than in Italian; there are many cases in which the reflex in Spanish is [j] instead of the expected [x] (← *ʒ), compare these Spanish words from Latin DI- and GI- with their outcomes in Portugese and Italian[23]:

LAT.	Es.	Port.	It.
HODIE	hoy [j]	hoje [ʒ]	oggi [ʤ]
RADIU[24]	rayo [j]	raio [j]	raggio [ʤ]
EXAGIU	ensayo [j]	ensaio [j]	saggio [ʤ]

[19] Lloyd 1987: p. 133.
[20] Lloyd 1987: p. 133.
[21] Alkire and Rosen 2010: p. 60.
[22] Lloyd 1987: p. 133.
[23] Alkire and Rosen 2010: p. 60.
[24] cf. Cat. *raig* [raʧ], Fr. *rai*, It. *raggio*; the reflex [j] seems to be quite widespread within Romània.

6.2.2.4 Palatalization and Affrication of Velars

ki | ke ⇒ ʧ
gi | ge ⇒ ʤ

The voiceless velar [k] palatalizes before front vowels:

LAT.	Es.
CIVITATE	ciudad [θ]
CINQUE	cinco [θ]
CENTU	ciento [θ]
VICINA	vecina [θ][25]
IACERE	yacer [θ]

The voiced velar [g] palatalized in the same context as [k]:

LAT.	Es.
GEORGIU	Jorge [x]
GENIU	genio [x]
GEMMA	yema [j]
GENERU	yerno [j]
GINGIVA	encía
GELARE	helar
LEGE	ley [j]
FUGIT	huye [j]

6.2.2.5 Palatalization of Sonorants

ll | li ⇒ ʎ
nn | ni | gn ⇒ ɲ

6.2.3 Structural Changes

6.2.3.1 Loss of Hiatus

V.V ⇒ VV

[25] cf. Judeo-Spanish *vizina* [z]

In Latin, other than the 3 diphthongs (AE, OE, and AU) mentioned above in Monophthongization, other sequences of vowels are heterosyllabic (Alkire and Rosen 2010: p. 57), while in Romance "[v]owels which had been in contact but had maintained their individual character by being pronounced in separate syllables began to be fused into a single syllable with the following vowel. The first result of this fusion was simply the pronunciation of /i/ in hiatus as a palatal semiconsonant [j] (a 'yod'). The next step was the change of /e/ in hiatus to a yod. This change may have occured fairly early. An inscription dated 125 B.C. has PARIAT for PAREAT (CIL 2, 592/10). Inscriptions from Pompeii also illustrate this same change: VALIA for VALEAT, ABIAT for HABEAT" (Lloyd 1987: p. 132). The resulting yods would palatalize the consonants through coarticulation, see the following examples[26]:

LAT.	Es.
PRETIU	precio
FACIA	faz
HODIE	hoy
FAGEU	haya
PALEA	paja
SENIOR	señor

Other than giving rise to a yod [j], dehiaticization may result in synizesis[27]:

LAT.	Es.
QUIĒTU	quedo
BATTUIT	bate
FUTUIT	jode
PARIĕTE	pared

or epenthesis[28]:

LAT.	Es.
GENUA	Genova
RUINA	ruina
IOHANNES	Juan

These yods from dehiaticization were one of the primary sources for palatalization (cf. Palatalization and Affrication of Dentals and Palatalization and Affrication of Velars). And some yods survived into Modern Spanish[29]:

LAT.	Es.
RUBEU	rubio
CEREU	cirio
(NERVUM →) NERVIU	nervio

Delabialization

[26] Alkire and Rosen 2010: p. 58.
[27] Alkire and Rosen 2010: p. 57.
[28] Alkire and Rosen 2010: p. 57-58.
[29] Lloyd 1987: p. 134.

6.2.3.2 Syncope

CV.C ⇒ CC

"There was extensive syncope of atonic vowels in words having four or more syllables in prehistoric Latin, and this process continued throughout the historical period" (Lloyd 1987: p. 113).
Here are some examples of syncope that were attested in Popular Latin[30]:

LAT.	Pop. Lat.
OCULU	OCLU
AURICULA	ORICLA
CALIDU	CALDU
ANGULUS	ANGLUS
SPECULUM	SPECLUM
VETULUS	VECLUS[31]
VIRIDIS	VIRDIS

And some that are reflected in Spanish[32]:

LAT.	Es.
LEPORE	liebre
I(N)SULA	isla
VIRIDE	verde

"The factor that seems to have had the greatest importance [in the application of syncope] was whether the consonants which came into contact after syncope formed normally occurring [consonant] groups in Latin" (Lloyd 1987: p. 114). Here are more examples of syncopes and their contexts[33]:

1. **posttonic**
   POSITUS → POSTUS → Es. *puesto*
   SUSPENDERE → SUSPENDRE
   TABULA → TABLA

2. **pretonic**
   MALEDĪXĪ → MALDIXI → Es. *maldije*

Occationally the syncopated and unsyncopated forms coexist in Spanish, usually the unsyncopated form being a *cultismo* e.g. parabola v. parabla.

6.2.3.3 Prothesis

$[_w$ s ⇒ $[_w$ is

[30] Alkire and Rosen 2010: p. 28-29; Lloyd 1987: p. 114.
[31] This is one of the words in which -TL- → -CL- (Alkire and Rosen 2010: p. 68).
[32] Alkire and Rosen 2010: p. 29.
[33] Lloyd 1987: p. 114.

This phenomenon is attested in the Visigothic documents[34]:

> [T]he addition of *i* at the beginning of a word that starts with an *s* followed by another consonant (*iscriptura*), or its suppression by hypercorrection (*sta* for *ista*).

LAT.	Es.
SPONSA	esposa
SPATA	espada
STUDIU	estudio
SPONGIA	esponja
SCUTU	escudo

6.2.3.4 Elision of Intervocalic [g]

$$\mathbf{V}.\mathrm{g}\mathbf{V} \Rightarrow \mathbf{VV}$$

LAT.	Es.
DIGITU	dedo
IAM MAGIS	jamás
MAGISTRU	maestro
SAGITTA	saeta
PAGENSE	país
CŌGITĀRE	cuidar

6.2.3.5 Elision of [m]

$$\mathrm{m}\ {}_{w}] \Rightarrow \emptyset\ {}_{w}]$$

The traces of this sound change go back to Old Latin (*CIL* I^2.9) (Clackson 2011: p. 17):

Old Lat.	Class. Lat.
OINO	UNUM
DVONORO	BONORUM
OPTUMO	OPTIMUM
VIRO	UIRORUM
SCIPIONE	SCIPIONEM

Also, "[i]n metrical texts, where the the final syllables of words ending in /**V**/ + <M> are treated in the same way as those in final /**V**/, being elided before a following initial /**V**/." (Clackson 2011: p. 87)

6.2.3.6 Deaspiration

$$\mathrm{h} \Rightarrow \emptyset$$

[34] Coulson and Babcock 2020: p. 159.

We have evidence that initial /h/ already weakens in Latin: "[i]n metrical texts final vowels are elided before words beginning with /h**V**/, exactly as before initial /**V**/." (Clackson 2011: p. 87)
An orthographic evidence for this sound change was the fact that the letter <h> sometime was used to indicate hiatus between vowels in inscriptions (Clackson 2011: p. 18): AHENVM[35] [ae:nʊm] (*CIL* I^2.581), AHENA[36] (*CIL* I^2.2093).

LAT.	Es.
HISPALIS	Sevilla

6.2.3.7 Nasal Spirant Law

n.s ⇒ s

Old Latin inscriptions reflect this change (Clackson 2011: p. 17): COSOL v. Class. Lat. CONSUL, CESOR v. Class. Lat. CENSOR (*CIL* I^2.8).

6.2.3.8 Syllabification of the [Cr] Cluster

Spanish have reflexes of some Latin words contains [Cr-] in the ultimate syllable which should have resulted in a stressed antepenultimate while the Spanish reflexes rather show a stressed penultimate[37]:

LAT.	Es.
ÁLACRE → ALÁCRE	alegre
CÁTHEDRA → CATHÉDRA	cadera
ÍNTEGRU → INTÉGRU	entero
TÉNEBRA → TENÉBRA	tinieblas
CÓLUBRA → COLÚBRA	culuebra
TÓNITRU → TONÍTRU	tronido

[35] cf. AĒNUM
[36] cf. AĒNA
[37] Lloyd 1987: p. 115.


DESVMMO B
ETYMO

Chapter 7

Western Romance

VENITE IGITVR DESCENDAMVS ET CONFVNDAMVS IBI LINGVAM EORVM VT NON AVDIAT VNVSQVISQVE VOCEM PROXIMI SVI

GENESIS XI•VII

7.1 Segmental Inventory

7.1.1 Vocalic

Western Romance Monophthongs

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

7.1.2 Consonantal

Western Romance Consonants

	Bilabial	Dental	Alveolar	Palatal	Velar	Labiovelar
Nasal	m	n		ɲ		
Stop	p b	t d			k g	k^w g^w
Fricative	ɸ β	ð			ɣ	
"			s z			
Affricate		ʦ ʣ		ʧ ʤ		
Trill		r				
Lateral		l		ʎ		

7.2 Sound Changes

7.2.1 Vocalism

7.2.1.1 Diphthongization I

$$\varepsilon \Rightarrow \text{ie}$$

The **stressed** low-mid vowel [ɛ] (for the origin of this vowel cf. Loss of Vowel Quantity and Great Merger) in open syllables underwent diphthongization in Western Romance (and Italian)[1]:

LAT.	Es.	Fr.	It.
PETRA	piedra	pierre	pietra
FELE	hiel	fiel	fiele
VENIT	viene	vient	viene
HERI	ayer	hier	ieri

In Spanish [ɛ] also diphthongizes in closed syllables (Diphthongization II).

$$\text{diphthongization}_1 \ : \ \text{Vowel} \rightarrow \text{Vowel}$$
$$\varepsilon \Rightarrow \text{ie}$$

7.2.2 Consonantism

7.2.3 Structural Changes

7.2.3.1 Degemination

$$\text{CC} \Rightarrow \text{C}$$

LAT.	Es.
OSSU	hueso
SUMMA	suma
APPELLAT	apela
LITTERA	letra
SICCU	seco

[1] Alkire and Rosen 2010: p. 15-16.

7.2.3.2 Lenition I

V.GV ⇒ V.ΓV
V.CV ⇒ V.GV

LAT.	Es.
SAPORE	sabor [β]
CAPUT	cabo [β]
COPERTU	cubierto [β]
VITA	vida [ð]
FATA	hada [ð]
CATENA	cadena [ð]
AMICA	amiga [ɣ]
SECURU	seguro [ɣ]
FOCU	fuego [ɣ]

LAT	Es.
CABALLU	caballo [β]
DEBERE	deber [β]
HABERE	haber [β]
CRUDU	crudo [ð]
PEDE	pie
AUGUSTU	agosto [ɣ]
LIGARE	ligar [ɣ]
PAGANU	pagano [ɣ]

LAT.	Es.
LEPRE	liebre [β]
CAPRA	cabra [β]
PETRA	piedra [ð]
PATRE	padre [ð]
HEMI-CRANIA	migraña [ɣ]

LAT.	Es.
SERPENTE	serpiente
ALPES	aples
RUMPERE	rumper
ORTICA	ortiga
MENTA	menta
ARCU	arco
FALCONE	halcón

The orthography of Visigothic texts shows signs of lenition and the confusion was caused by it (Coulson and Babcock 2020: p. 159):

> The use of *b* instead of *p* (*abtum*) or *p* for *b* (*puplicum*), *g* for *c* (*eglesia*[2]) or the reverse (*intecritate*), *k* for *c* (*kaput*), *t* for *d* (*aput*) or *d* for *t* (*sustinead*).

$$\mathbf{V}_{\sigma}][_{\sigma}[_{\Omega}\mathbf{C} \vdash \text{stop; voiced}][_{\rho}\mathbf{V}]] \Rightarrow \mathbf{V}_{\sigma}][_{\sigma}[_{\Omega}[\text{stop} \rightarrow \text{friactive}]\ \mathbf{C}][_{\rho}\mathbf{V}]]$$
$$\mathbf{V}_{\sigma}][_{\sigma}[_{\Omega}\mathbf{C} \vdash \text{stop; voiceless}][_{\rho}\mathbf{V}]] \Rightarrow \mathbf{V}_{\sigma}][_{\sigma}[_{\Omega}[\text{voiceless} \rightarrow \text{voiced}]\ \mathbf{C}][_{\rho}\mathbf{V}]]$$

[2] cf. Es. *iglesia*

D. AFONSO X ENPERADOR murió 1284.

Chapter 8

Old Spanish

> Ya lo vedes, que partirnos emos en vida, yo iré e vós fincaredes remanida.
>
> Cantar de mio Cid

8.1 Segmental Inventory

8.1.1 Vocalic

Old Spanish Monophthongs

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

8.1.2 Consonantal

Old Spanish Consonants

	Bilabial	Dental	Alveolar	Palatal	Velar	Labiovelar	Glottal
Nasal	m		n	ɲ			
Stop	p b	t d			k g	kʷ gʷ	
Fricative	ɸ β	ð		ʝ	ɣ		h
"			s̪ z̪	ʃ ʒ			
Affricate		ʦ ʣ		ʧ ʤ			
Trill		r					
Tap		ɾ					
Lateral		l		ʎ			

8.2 Sound Changes

8.2.1 Vocalism

8.2.1.1 Diphthongization II

ɛ ⇒ ie

ɔ ⇒ we

Spanish is unique among other Western Romance languages in that the diphthongization for the front low-mid vowel [ɛ] happens in both open and closed syllables, and Spanish diphthongizes also the back low-mid vowel [ɔ] in open and closed syllables[1]:

LAT.	Es.
HIBERNU	invierno
APERTA	abierta
SEPTE	siete
CERVU	ciervo
FERRU	hierro
FORTE	fuerte
PORTA	puerta
MORDIT	muerde
MORIT	muere
MOVET	mueve
POTET	puede

[1]Lloyd 1987: p. 122; Alkire and Rosen 2010: p. 15-16.

The two unique phonological processes of Spanish, (a) expanding the context of the diphthongization of [ɛ] to both open and closed syllables and (b) including [ɔ] as the target of diphthongization (also in open and closed syllables), separate it from Western and Eastern Iberian Romance languages; cf. the reflexes of these two Latin words in Portugese to Castilian[2]:

LAT.	Cast.	Port.
TERRA	tierra	terra
PORTA(M)	puerta	porta

The diphthongization of [ɛ] (cf. Diphthongization I) and [ɔ] had shown its traces very early in Latin inscriptions: NIEPOS for NEPOS ca. A.D. 120, PUOSUIT for POSUIT ca. A.D. 157[3]. Diphthongization continued to be attested in the era of the Visigoths: [V]ALIENTEM for VALENTEM, CURRIENTE for CURRENTE, and PARIENTIBUS for PARENTIBUS[4].

$$\text{diphthongization}_2 : \text{Vowel} \rightarrow \text{Vowel}$$
$$\varepsilon \Rightarrow \text{ie}$$
$$\text{ɔ} \Rightarrow \text{we}$$

8.2.2 Consonantism

8.2.2.1 Debuccalization of [ɸ]

$$\phi \text{V} \Rightarrow \text{hV}$$

Many Latin words with word-initial singleton onset F- now have zero onset in Modern Spanish (once was [h] and written as such <h> in Old Spanish), while in clusters [f] (← [ɸ]) is preserved[5]:

LAT.	Es.
FILU	hilo
FERIRE	herir
FERRO	hierro
FALCONE	halcón
FOCU	fuego
FORA	fuera
FONTE	fuente
FRONTE	frente
FLORE	flor
FLACCU	flaco

[2] Hualde 2010: p. 3.
[3] Lloyd 1987: p. 129-130.
[4] Lloyd 1987: p. 130.
[5] Alkire and Rosen 2010: p. 50.

Sometimes this sound change is attributed to the Romance-Basque bilingualism in the areas where Old Castilian was developed; recent discovery in Basque historical phonology has challenged this view (Manterola and Hualde 2021).

8.2.2.2 Fronting of Palatal Affricates

$$\widehat{\text{tʃ}} \Rightarrow \text{ts}$$

The palatal affricate [t͡ʃ] coming from palatalized [k]'s fronts to merge with the dental affricate [ts], both written with a cedilla <ç> in Old Spanish[6]:

LAT.	OSp.
CAEPULLA	çebolla [ts]
CĒNA	çena [ts]
CENTU	çiento [ts]
CINQUE	çinco [ts]
MARTIU	marçço [ts]
TERTIĀRIU	terçero [ts]

8.2.3 Structural Changes

8.2.3.1 Metaphony

$$[_{\sigma}\text{V}] :: [_{\sigma}\text{V} \vdash \text{high}] \Rightarrow [_{\sigma}[\text{height+1}]\ \text{V}] :: [_{\sigma}\text{V}]$$

LAT.	Es.
MULTU	mucho
AUSCULTAT	escucha
LUCTA	lucha
LACTE	leche
FACTU	hecho
BASIU	beso
CASEU	queso
RENIŌNE	riñón
GENESTA	hiniesta
CAEMENTUM	cimiento
TENEBRAS	tinieblas
COCHLEĀRE	cuchara
COGNĀTU	cuñado
MULIERE	mujer

[6] Lloyd 1987: p. 258-259.

8.2.3.2 Apocope

$$\mathrm{V}_{w}] \Rightarrow \varnothing_{w}]$$

OSp.	Es.
primero	primer
tercero	tercer
santo	san
segundo	según

LAT.	Es.
PARIETE	pared
MERCEDE	merced
PĀNE	pan
MARE	mar
FIDĒLE	fiel
PĀCE	paz
CALCE	coz
FALCE	hoz
FASCE	haz
PISCE	pez

8.2.3.3 Palatalization of Clusters

-CT- ⇒ ʧ

LAT.	Es.
DICTU	dicho [ʧ]
STRICTU	estrecho [ʧ]
PECTU	pecho [ʧ]
TECTU	techo [ʧ]
NOCTE	noche [ʧ]
OCTO	ocho [ʧ]

8.2.3.4 Metathesis

Chapter 9

Modern Spanish

Todo está cumplido.

Juan 19:30

9.1 Segmental Inventory

9.1.1 Vocalic

Modern Spanish Monophthongs

	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

9.1.2 Consonantal

Modern Spanish Consonants

	Bilabial	Labiodental	Dental	Alveolar	Palatal	Velar	Labiovelar
Nasal	m			n	ɲ		
Stop	p b		t d			k g	k^w g^w
Fricative	β	f	θ ð		ʝ	ɣ	
"				s̪ z̪			
Affricate					ʧ		
Approximant							w
Trill				r			
Tap				ɾ			
Lateral				l	ʎ		

9.2 Sound Changes

9.2.1 Vocalism

9.2.2 Consonantism

9.2.2.1 Betacism II

$$[_w\ \beta \Rightarrow [_w\ b$$

As a continuation of the change Betacism I, word-initial [b] and [β] merged in Modern Spanish.

9.2.2.2 Labiodentalization

$$\phi \Rightarrow f$$

Modern Castilian has a labiodental production of the Latin <F>.

9.2.2.3 Deaffrication of [ʤ]

$$ʤ \Rightarrow ʒ$$

At some point after 1492, Castilian [ʤ] were deaffricated to [ʒ], mirroring the change in French. Judeo-Spanish retains [ʤ]: *gente* is [ʤente] (Hualde and Şaul 2011: p. 92).

9.2.2.4 Devoicing of Fricatives

$$\mathbf{V}.\underset{\sqcup}{\text{z}}\mathbf{V} \Rightarrow \underset{\sqcup}{\text{s}}$$
$$\mathbf{V}.\underset{\sqcup}{\text{z}}\mathbf{V} \Rightarrow \underset{\sqcup}{\text{s}}$$
$$\mathbf{V}.\text{ʒ}\mathbf{V} \Rightarrow \text{ʃ}$$

$$[_w\ \text{ʒ} \Rightarrow [_w\ \text{ʃ}$$

Alfonsino Spanish had sibilants represented as such orthographically (Lloyd 1987: p. 328):

	dental affricate	alveolo-apical	palatal
voiceless	ç [ʦ] (→ [$\underset{\sqcup}{\text{s}}$])	-ss- [$\underset{\sqcup}{\text{s}}$]	x [ʃ]
voiced	z [ʣ] (→ [$\underset{\sqcup}{\text{z}}$])	-s- [$\underset{\sqcup}{\text{z}}$]	j, ge, gi [ʒ] (← [ʤ])

As a part of the sibilant shifts from Old Spanish to Modern Spanish, the voicing of these sibilants would neutralize.

9.2.2.5 Deaffrication of Dental Affricates

$$\text{ʦ} \Rightarrow \underset{\sqcup}{\text{s}}$$
$$\text{ʣ} \Rightarrow \underset{\sqcup}{\text{z}}$$

"Medieval Jews in Spain preponderantly transcribed ç with the letter **samech** ס rather than the letter **tsade** צ from the thirteenth to the fifteenth centuries" (Lloyd 1987: p. 333).

9.2.2.6 Desibilation of Dental Fricatives

$$\underset{\sqcup}{\text{s}} \Rightarrow \theta$$

"We find the word *capuz* spelled *capud* in 1495 [...] [i]n 1547 there appear such spellings as *Trinidaz*, and *paternidaz* for *Trinidad* and *paternidad*" (Lloyd 1987: p. 334).

9.2.2.7 Retraction of Palatal Fricative

$$\text{ʃ} \Rightarrow \text{x}$$

The palatal fricative [ʃ] seemed to be in a pretty unstable state toward the end of the Old Spanish era, possibly influenced by the deaffrication of [ʤ] → [ʒ], these two sounds of different origins started to be

interpreted as a *voiceless* : *voiced* pair just like many other sounds in Spanish. In some documents we can see them being alternated (Lloyd 1987: p. 342):

[ʃ]	[ʒ]
celosía	celogía
frisol	frijol
iglesia	eclegia
quise	quije
resistir	registir
visitar	vigitar
residente	regidente
relisión	religión
colesio	colegio
disistir	digestir
mesor	mejor
parasismo	parajismo
sanguisuela	sangijuela

Eventually in the final result of the consonantal rearrangement from Old Spanish to Modern Spanish, the sibilant [ʃ] retracted to [x]. Curiously, this sound is identified with the [h] ← [ɸ] in some speeches (Lloyd 1987: p. 343):

hizo	gizo
harán	garan
Herrera	gerera
harta	garta

9.2.3 Structural Changes

9.2.3.1 Lenition II

V.b**V** ⇒ β
V.d**V** ⇒ ð
V.g**V** ⇒ ɣ

V.ʦ**V** ⇒ ʣ

At the dawn of Modern Spanish, the final stage of intervocalic lenition is completed. This sound change had greatly disturbed the Alfonsino orthography. Just take the case of the merger between [β] and [b] intervocalically (Hualde 2005: p. 279). At the time of Alfonso X, Old Spanish orthography has <v> for Latin <v>:

Lat.	OSp.
LAVĀRE	lavar [β]
VACCA	vaca [β]

and also for Latin <B> intervocalically:

Lat.	OSp.
DEBĒRE	dever [β]
CABALLU	cavallo [β]
HABĒRE	aver [β]
CANTĀBĀS	cantavas [β]

While some Old Spanish <b> is from Latin intervocalic <P>:

Lat.	OSp.
SAPĒRE	saber [b]
LUPU	lobo [b]

After this sound change and Betacism II had both been installed, the contrast between [β] and [b] collapsed, even at word-initial position. The pairs [g] : [ɣ] and [d] : [ð] must have undergone the same change; but without a letter like <v> to give us direct hints about the fricative stage, the history of lenition for these two pairs is more obscured than the pair [b] : [β] (although later on around the time of sound change Desibilation of Dental Fricatives, <z> [z̪] and <d> [ð] were sometimes confused).
Another participant of lenition is the [ʦ] from [ʧ] (← [kʲ]) and [tʲ], which irregularly undergoes lenition[1]:

Lat.	PrRom.	OSp.
RATIŌNE	[ʦ]	razón [ʣ]
PUTEU	[ʦ]	pozo [ʣ]
CAPITIA	[ʦ]	cabeça [ʦ]
PLATEA	[ʦ]	plaça [ʦ]
ACIĀRIU	[ʧ]	azero [ʣ]
CORTICEA	[ʧ]	corteza [ʣ]
CORĀCEA	[ʧ]	coraça [ʦ]
PELLĪCEA	[ʧ]	pelliça [ʦ]

[1] Lloyd 1987: p. 261.

Chapter 10

Glossarium Latinocastellanum

10.1 Nominal

10.2 Numeral

Bibliography

Phonology

Chomsky, Noam and Morris Halle (1968). *The Sound Pattern of English*. Harper & Row.

Kaplan, Ronald M. and Martin Kay (1994). "Regular Models of Phonological Rule Systems". In: *Computational Linguistics* 20.3, pp. 331–378.

Kondrak, Grzegorz (2002). *Algorithms for Language Reconstruction*. URL: https://webdocs.cs.ualberta.ca/~kondrak/papers/thesisnew.pdf (visited on 10/17/2022).

Hartman, Steven Lee (2003a). *Modeling Phonological Change*. URL: https://langev.com/pdf/harman03phonologicalChange.pdf (visited on 10/17/2022).

— (2003b). *Phono (Version 4.0): Software for Modeling Regular Historical Sound Change*. URL: https://langnhist.weebly.com/files/theme/ver40.pdf (visited on 10/18/2022).

Prince, Alan and Paul Smolensky (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell Publishing.

Bouchar-Côté, Alexandre et al. (2007). *A Probabilistic Approach to Diachronic Phonology*.

Sims-Williams, Patrick (2018). "Mechanising Historical Phonology". In: *Transactions of the Philological Society* 116.3, pp. 555–573.

Romance and Historical Linguistics

Spiegel, Irving (1952). *Old Judaeo-Spanish Evidence of Old Spanish Pronunciation*.

Hartman, Steven Lee (1974). "An outline of Spanish historical phonology". In: *Research on Language & Social Interaction* 7.1-2, pp. 123–191. DOI: 10.1080/08351817409370365.

Lloyd, Paul M. (1987). *From Latin to Spanish*. American Philosophical Society.

Kohen, Elli and Dahlia Kohen-Gordon (2000). *Ladino-English English-Ladino Concise Encyclopedic Dictionary (Judeo-Spanish)*. Hippocrene Books.

Penny, Ralph (2002). *A History of the Spanish Language*. Cambridge University Press.

Hualde, José Ignacio (2004). "Quasi-Phonemic Contrasts in Spanish". In: *WCCFL 23 Proceedings*. Ed. by Benjamin Schmeiser et al.

— (2005). *The Sounds of Spanish*. Cambridge University Press.

Quirk, Ronald J. (2005). "The "Appendix Probi" as a Compendium of Popular Latin: Description and Bibliography". In: *The Classical World* 98.4, pp. 397–409.

Chitoran, Ioana and José Ignacio Hualde (2007). "From hiatus to diphthong: the evolution of vowel sequences in Romance". In: *Phonology* 24, pp. 37–75. DOI: https://doi.org/10.1017/S095267570700111X.

Varol, Marie-Christine and Ralph Tarica (2008). *Manual of Judeo-Spanish: Language and Culture*. University Press of Maryland.
Alkire, Ti and Carol Rosen (2010). *Romance Languages: A Historical Introduction*. Cambridge University Press.
Hualde, José Ignacio (2010). *Ibero-Romance II: Astur-Leonese, Spanish, Navarro-Aragonese*. Oxford Research Encyclopedia of Linguistics. DOI: https://doi.org/10.1093/acrefore/9780199384655.013.718.
Clackson, James, ed. (2011). *A Companion to the Latin Language*. Blackwell Companions to the Ancient World. Wiley-Blackwell.
Hualde, José Ignacio and Mahir Şaul (2011). "Istanbul Judeo-Spanish". In: *Journal of the International Phonetic Association* 41.1, pp. 89–110.
Maiden, Martin, John Charles Smith, and Adam Ledgeway, eds. (2011). *The Cambridge History of The Romance Languages: Structures*. Vol. 1. Cambridge University Press.
Bermúdez-Otero, Ricardo (2013). *The Spanish lexicon stores stems with theme vowels, not roots with inflectional class features*.
Baddeley, Susan and Anja Voeste, eds. (2017). *Orthographies in Early Modern Europe*. De Gruyter Mouton.
Marr, Clayton (2017). *Towards a computational model of long-term diachronic change: simulating the development of Classical Latin to Modern French*.
Quirk, Ronald J. (2017). "Hypercorrection in the *Appdendix Probi*". In: *Philologus* 161.2, pp. 350–353. DOI: https://doi.org/10.1515/phil-2016-0119.
Dworkin, Steven N. (2018). *A Guide to Old Spanish*. Oxford University Press.
Coulson, Frank T. and Robert G. Babcock, eds. (2020). *The Oxford Handbook of Latin Palaeography*. Oxford University Press.
Recasens, Daniel (2020). *Palatalizations in the Romance Languages*. Oxford Research Encyclopedia of Linguistics. DOI: https://doi.org/10.1093/acrefore/9780199384655.013.435.
Manterola, Julen and José Ignacio Hualde (2021). "Old Basque had */χ/, not /h/". In: *Journal of Historical Linguistics* 11.3, pp. 421–456.
Barbato, Marcello (2022). *The Early History of Romance Palatalizations*. Oxford Research Encyclopedia of Linguistics. DOI: https://doi.org/10.1093/acrefore/9780199384655.013.750.
Mackenzie, Ian (2022). "The Genesis of Spanish /θ/: A Revised Model". In: *Languages* 7.191.
Jover, Francisco Gago, Andrés Enrique Arias, and Javier Pueyo Mena, eds. (2022a). *Biblia romanceada E8*. Spanish Biblical Texts. Digital Library of Old Spanish Texts. URL: http://www.hispanicseminary.org/t&c/bib/e8/framconc.htm (visited on 10/23/2022).
— eds. (2022b). *Evangelios y Epístolas paulinas*. Spanish Biblical Texts. Digital Library of Old Spanish Texts. URL: http://www.hispanicseminary.org/t&c/bib/lucena/framconc.htm (visited on 10/17/2022).

Computation

Backus, J. W. et al. (1960). "Report on the Algorithmic Language ALGOL 60". In: *Communications of the ACM* 3.5.
Geurts, Leto and Lambert Meertens (1978). "Remarks on Abstracto". In: *ALGOL Bull.* 42, pp. 56–63. DOI: https://dl.acm.org/doi/10.5555/1061701.1061714.
Bird, Richard S. and Lambert Meertens (June 1987). "Two Exercises Found in a Book on Algorithmics". In: *The IFIP TC2/WG 2.1 Working Conference on Program specification and transformation*, pp. 451–457. DOI: https://dl.acm.org/doi/10.5555/28195.28216.

Friedman, Daniel P. and Matthias Felleisen (1989). *The Little LISPer*. 3rd ed. The MIT Press.
Bird, Richard S. and Oege de Moor (1997). *Algebra of Programming*. Ed. by C. A. R. Hoare. Prentice Hall International Series in Computer Science. Prentice Hall Europe.
Milner, Robin et al. (1997). *The Definition of Standard ML*. Revised. The MIT Press. DOI: https://doi.org/10.7551/mitpress/2319.001.0001.
Pierce, Benjamin C. (2002). *Types and Programming Languages*. The MIT Press.
Marlow, Simon, ed. (2010). *Haskell 2010 Language Report*.
Harper, Robert (2016). *Practical Foundations for Programming Languages*. 2nd ed. Cambridge University Press.
Gibbons, Jeremy (2020). *The School of Squiggol*. URL: https://www.cs.ox.ac.uk/people/jeremy.gibbons/publications/squiggol-history.pdf (visited on 10/28/2022).
Gorman, Kyle and Richard Sproat (2021). *Finite-State Text Processing*. Ed. by Graeme Hirst. Synthesis Lectures on Human Language Technologies. Springer. DOI: 10.1007/978-3-031-02179-4.

Misc.

Aigner, Martin and Günter M. Ziegler (2010). *Proofs from THE BOOK*. 4th ed. Springer.
Coogan, Michael D. et al., eds. (2018). *The New Oxford Annotated Bible: New Revised Standard Edition*. 5th ed. Oxford University Press.
Underflow (2018). *To Clio*. URL: https://greekpagan.com/2018/09/07/to-clio/ (visited on 10/28/2022).