Automated Deduction in Historical Phonology

Dearest Dream of My Youth

陳朝陽
Chen Zhaoyang
zc23@illinois.edu

November 27, 2022

TL; DR

I wrote a small piece of program (~ 1,000 loc) that implements a couple dozen major phonological changes from Latin to Modern Spanish. The program is able to derive a good amount of Modern Spanish words from their Latin etymon.
This presentation is about

1. How it is done
2. Why do it in the first place

Reviving a field that has been silent for a while

One can see that very little of this project is new in the field by any means. In fact, an essentially identical project was done from the 1980s to early 2000s by S. Lee Hartman and is still accessible online as of now[1].
So why do it again?

[1] Steven Lee Hartman. *Phono (Version 4.0): Software for Modeling Regular Historical Sound Change*. 2003. URL: https://langnhist.weebly.com/files/theme/ver40.pdf (visited on 10/18/2022).

Seasoned Techniques

This project is a synthesis of three techniques: rule-based phonology, Romance historical linguistics, and some elementary programming. All of which are well-established techniques. But to synthesize all of them, as I have said, is something that has not been done for a good while.

Rule-based Phonology

Romance Historical Linguistics

Historical Phonology of the Spanish Language

The Standard ML Programming Language

This small program is written entirely in Standard ML, without using
SML is a relatively small programming language that has been primarily used in programming language implementation and automated theorem proving.
Its *Heimatland* was University of Edinburgh,

Why SML?

As of now, Python is the *de facto* LINGUA FRANCA of scientific computing, and, in our particular interest, computational linguistics.
One may ask: why choose a programming language that is little known[2] to the historical linguistics community, which may potentially impose a language barrier when communicating the results?

[2] Although some members of the formal semantics community within linguistics are familiar with Haskell, which is a close relative of SML.

Why SML? cont.

The answer is that ... it is mostly personal preference and *sectarian* reasons. Any modern general purporse programming language is capable of writing this small piece of program, I just happen to come from the ML camp of programming languages.

Implementing a Phonology

This project has a simple two-layer structure: the first layer that defines ways of constructing the *statics* of a phonology – namely the segmental inventory, syllable structure, and the phonological word of the Spanish language and her predecessors – and the second layer that defines ways of constructing the *dynamics* of a phonology – namely sound changes and how to compose them into chain shifts.

The Statics

Representing Features, Segments, Syllables, and the Phonological Word

The Dynamics

Rewriting of Syllabic Constituents, Syllables, and Phonological Words

Pearls of Sound Changes from Latin to Romance

In the remainder of this presentation, I am going to demonstrate

Latin → Proto-Romance: Romance Vowel Shifts

Arguably the most fundamental change from Late Latin to Proto-Romance is the transformation of its vowel system.
The transformation has the following components:

1. Loss of Vowel Quantity
2. The Great Vowel Merger
3. Merger in Atonic Vowels
4. Reduction in Final Vowels

Another important sound change, the loss of hiatus, unfortunately we are not going to cover in this presentation.

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	

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

The 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	

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

Romance Vowel Shifts: Rule Ordering

Loss of Quantity $\sqsubset$ Great Merger $\sqsubset$ Atonic Merger $\sqsubset$ Final Vowel Reduction[3]

[3] This notation means: $0 \sqsubset 1 \sqsubset 2 \sqsubset \ldots$

Latin → Proto-Romance: Fundamental Consonantal Shifts

Proto-Romance → Western Romance: Intervocalic Lenition

Proto-Romance → Western Romance: Degemination

Westerm Romance → Old Spanish: Debuccalization of [ɸ]

Old Spanish → Modern Spanish: The Spanish Sibilant Rearrangement

Mein liebster Jugendtraum

Old Chinese → Middle Chinese

Bibliography

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