Rule Ordering and Relative Chronology

David R. Mortensen

April 18, 2024

Sound changes are historical events:

(1) a. They happen as a specific time interval
    b. They leave the word different than they found it

As a result, they occur in an order. Linguists call this order their RELATIVE CHRONOLOGY.

Let's go back to the Great Vowel Shift (GVS) (Figure 1)

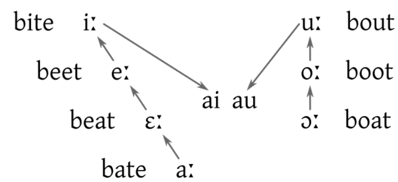

Figure 1: The Great Vowel Shift (image credit: Goran tek-en)

First, diphthongization must have occurred. Otherwise, when /eː/ and /oː/ were raised, they would have merged with /iː/ and /uː/ (the distinction between them would have been lost). Subsequently, /ɛː/ and /ɔː/ were raised to /eː/ and /oː/.

A similar sequence of changes (called a chain shift) occurred in Huishu.

(2) Huishu chain shift

  a. **Coloring:** Vowels assimilate to plosive codas
    i. PTk *jip > jep 'sleep'
    ii. PTk *jak > jok 'ashamed'
    iii. PTk *hak > hok 'big'
    iv. PTk *mik > mok 'eye'
    v. PTk *cik > tsok 'wash'
    vi. PTk *hrit > rejt 'heavy'

    but
    i. PTk *ʃap > sap 'have the ability'
    ii. PTk *nap > nap 'snot'

  b. **Debuccalization:** Plosive codas become glottal stops
    i. PTk *hwok > huk > huʔ 'pig'
    ii. PTk *maːk > maʔ 'brother in law'
    iii. PTk *kaːp > kaʔ 'shoot'
    iv. PTk *hak > hok > hoʔ 'big'
    v. PTk *nap > nap > naʔ 'snot'
    vi. PTk *jip > jep > jeʔ 'sleep'
    vii. PTk *mik > mok > moʔ 'eye'
    viii. PTk *cik > tsok > tsoʔ 'wash'
    ix. PTk *hrit > rejt > rejʔ 'heavy'

  c. **Epenthesis:** -k codas inserted after high vowels
    i. PTk *ri > lik 'medicine'

ii. PTk *rɨ > ruk 'chop'
iii. PTk *ru > ruk 'bone'

d. **Vowel Raising 1:** mid vowels are raised to high vowels
i. PTk *lo > lu 'buy'
ii. PTk *ko > ku 'nine'
iii. PTk *hwe > hwi 'late'

e. **Vowel Rasising 2:** low vowels are raised to mid vowels
i. PTk *la > le 'star'
ii. PTk *da > re 'sharpen'

Even though there is no written record of the history of Huishu (unlike English), it is possible to reconstruct—with unassailable logic—the history of changes. This can be done through counterfactual reasoning.

(3) a. If **Debuccalization** happened before **Coloring**, there could be no explanation for the difference in the rhyme of, e.g., jeʔ 'sleep' (< *jip) and moʔ 'eye' (< *mik).

b. If **Epenthesis** happened before **Debuccalization** then the epenthetic /k/s would have debuccalized.

c. If **Vowel Raising 1** happened before **Epenthesis**, formerly mid vowels (/e/ and /o/) would have developed coda /k/.

d. If **Vowel Raising 2** happened before **Vowel Raising 1**, PTk *a would be reflected as /i/ because it would become /e/ (Vowel Raising 2) before Vowel Raising 1 applied, which would raise it to /i/.

Ordering Relations between Sound Laws

For every pair of sound laws in the history of a language, one of five relationships hold:

- Feeding: change A creates environments where change B will apply
- Bleeding: change A eliminates environments where change B would otherwise apply
- Counterfeeding: change B *would* create environments in which change A would apply were it ordered before A
- Counterbleeding: change B *would* destroy environments in which change A would apply were it ordered before A
- No interaction

For example, the English apocope rule that deleted unstressed mid central vowels in certain context (as in /walkəd/ > /walkd/) fed the voicing assimilation rule (/walkd/ > /walkt/) because it created voiced-voiceless plosive clusters where the assimilation rule could apply.

Let's look at the Huishu sound laws:

(4) a. Coloring, Debuccalization: **Counterbleeding**
  b. Debuccalization, Epenthesis: **Counterfeeding**
  c. Epenthesis, Vowel Raising 1: **Counterfeeding**
  d. Vowel Raising 1, Vowel Raising 2: **Counterfeeding**

So far, we have examples of Feeding, Counterfeeding, and Counterbleeding. What about Bleeding? In Yawelmani (an indigenous language of North America), glottal stops between vowels were first deleted, then vowels before glottal stops were lengthened. This means that the deletion rule destroyed environments where the lengthening rule would have otherwise applied and thus bleeds that rule.

Rule Ordering in Synchronic Phonology

Following Saussure, language can be looked at from two perspectives:

(5) a. Diachronic
  b. Synchronic

DIACHRONIC linguistics looks at language as it develops over time. The sound changes we have talked about so far in this lecture are diachronic. We are talking about events in history.

SYNCHRONIC linguistics looks at language as a system that exists at a particular time. Most of our discussion about morphology, phonetics, and orthography has been synchronic. It is about describing the behavior of an idealized speaker of a language at a single point in time (again, an idealization).

Synchrony is always a product of diachrony and diachrony is always filtered through synchrony. Consider the case of English plurals/third-person singular non-pasts. These are formed by suffixing /z/ to words (orthographically ⟨s⟩):

(6) a. pigs /pɪgz/
  b. dogs /dɑgz/
  c. lions /lajənz/
  d. bids /bɪdz/

First, a sound change inserted a vowel between sibilants in English. That means that sequences of /s/, /z/, /ʃ/, /ʒ/, /t͡ʃ/ and /d͡ʒ/ followed by /z/ (etc.) were broken up by an epenthetic /ə/ (or /ɨ/):

(7) a. horses /hɔɹsəz/
b. fish /fɪʃəz/
c. rose /ɹowzəz/
d. rage /ɹejd͡ʒəz/

This resulted in allomorphy. A sound change in English then changed voiced fricatives and plosives to their voiceless equivalents when there were preceded by voiceless sounds. This resulted in a new kind of allomorphy:

(8) a. cats /kæts/
b. colts /kowlts/
c. bats /bæts/
d. sits /sɪts/

Both patterns of allomorphy resulted from historical changes. However, the patterns that they introduced exist synchronically, in the behavior of speakers. We know this because—if we ask English-speaking children to form the plurals of nouns or third-person singular non-pasts of verbs—they recapitulate these changes:

(9) a. There are two /wʌg**z**/
b. There are two /kæk**s**/
c. There are two /bæz**əz**/

English speakers must be aware, synchronically, of these patterns because they can apply them PRODUCTIVELY. One way of modeling this is as a sequence of string-rewrite rules that look suspiciously like our sound laws:

(10) a. $\emptyset \rightarrow \text{ə} \;/\; \begin{bmatrix}+\text{strident}\end{bmatrix} \;\_\_\; \begin{bmatrix}+\text{strident}\end{bmatrix}$
b. $\begin{bmatrix}-\text{sonorant}\end{bmatrix} \rightarrow \begin{bmatrix}-\text{voice}\end{bmatrix} \;/\; \begin{bmatrix}-\text{voice}\end{bmatrix} \;\_\_$

The only real difference, in terms of notation, is that we right the arrows as $\rightarrow$ rather than > to show that they are synchronic rather than diacronic.

Exercise: What is the relationship between these two rules?

Modeling Rule Ordering Computationally

Representing Rule Ordering with FSTs

We mentioned in the last lecture that sound changes (and, by extension, synchronic phonological rules) can be modeled as finite state transducers. These are simple machines that that a string as an input and return a string as an output. In fact, FSTs are probably somewhat more powerful than is required for sound laws or phonological rules.

One interesting thing about FSTs is that they can be composed: one can combine two FSTs to produce a new FST that is equivalent to feeding an input to the first transducer, collecting the output, feeding it to the second transducer, and collecting the output from that. This is basically the same as function composition. This means that it is easy to model relative chronology and rule ordering with finite state transducers. Equivalent formalisms have been applied to historical linguistics[1] and FSTs have also been used to model sound change.[2]

[1] Clayton Marr and David Mortensen. Large-scale computerized forward reconstruction yields new perspectives in french diachronic phonology. *Diachronica*, 2022

[2] Tomotheus A. Bodt and Johann-Mattis List. Reflex prediction: A case study of Western Kho-Bwa. *Diachronica*, 2022. ISSN 0176-4225. DOI: 10.1075/dia.20009.bod

Learning Rule Ordering as a Code Generation Problem

As mentioned in the last lecture, ongoing work shows that LLMs can generate sound changes given a set of input-output pairs (by treating the rule discovery as a code generation problem). It is also true that such models can infer ordering among rules (relative chronology). This can be done with comparatively little data (only a few hundred items).

This opens up an interesting pipeline (see Figure 2).

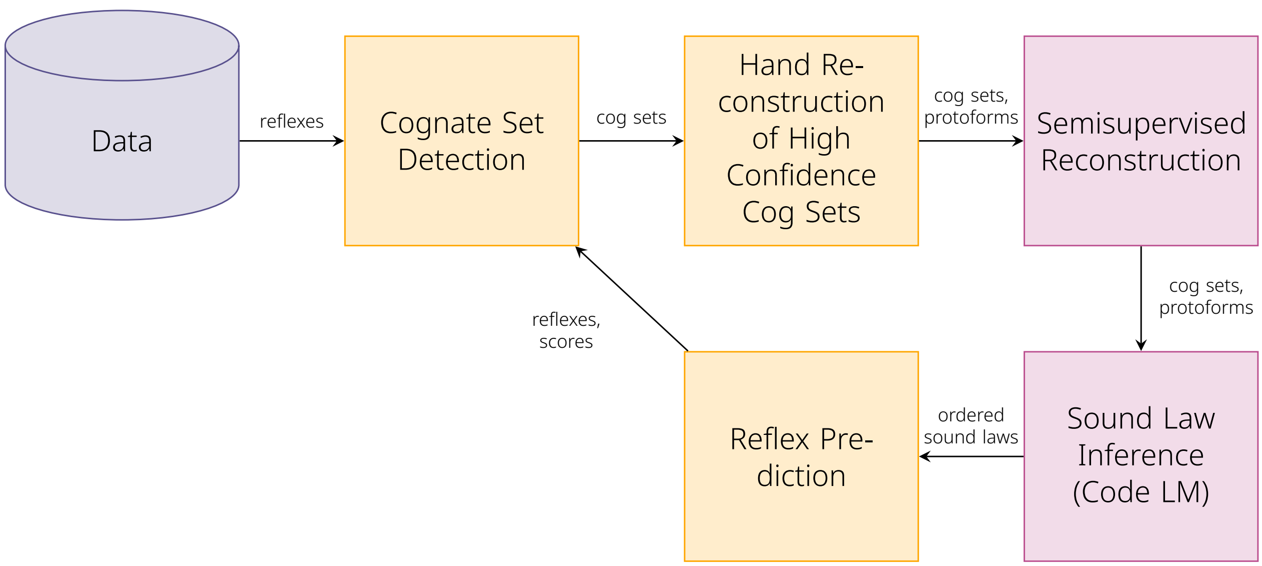

Figure 2: Proposed pipeline for human-in-the-loop reconstruction. Processes in purple have already been implemented.

References

Tomotheus A. Bodt and Johann-Mattis List. Reflex prediction: A case study of Western Kho-Bwa. *Diachronica*, 2022. ISSN 0176-4225. DOI: 10.1075/dia.20009.bod.

Clayton Marr and David Mortensen. Large-scale computerized forward reconstruction yields new perspectives in french diachronic phonology. *Diachronica*, 2022.