Why We Need a Sound Change Database

- Historical phonologists have used the comparative method for over 150 years to reliably document sound correspondences and reconstruct sound changes in many of the world's languages and language families.
- Despite the breadth of work done in the field, there currently exists no database (or even a print volume) of the world's recurrent regular sound changes to consult should one want to determine something as basic as the overall cross-linguistic frequency a particular sound change or whether a given sound change is more common than another.
- How cross-linguistically common is sound change?
- Which is more common: h > s or s > h?
- Historical linguists often texts provide examples of "common" and "rare" types of sound change, but do not elaborate on the cross-linguistic applicability of such claims, or how they arrived at these conclusions.

A standard historical linguistic method to develop a feeling for which kinds of changes are 'common' (or 'rare'), 'natural' or 'possible' and which are 'unnatural' (or 'impossible'). This feeling is typically developed through a combination of experience, working through the histories of languages, and inherited wisdom, gained through reading textbooks or discussion with others. It's a rather haphazard basis for science; it would be much easier if there were a systematic collection of types of changes, as found in the histories of the language of the world, organized on a phonological basis, from which we could extract generalizations on a formal basis about which changes happen commonly and what kinds of things are not found at all. This book aims to fill this need, and, as a historical phonologist, I'm grateful for it.

-Honeybone in his 2009 review of Martin Kippen's Komponentenlehre (p. 28).

- Most historical reconstructions document sound changes within the context of a given language, subgroup, or language family, often without considering possible phonetic bases for such changes.
- Typologies of individual sound patterns can be found in the phonological literature, and are often informed by phonetics, but they are rarely discussed in terms of other similar sound changes.
- A database could also aid in the investigation of the notable similarities between common types of historical sound change and common synchronic sound patterns, recurrent synchronic sound patterns observed in unrelated languages that are argued to share regular phonetically based sound change as a common source, and their potential parallel evolution constrained by the perceptual and articulatory properties of human speech claimed to be the force underlying their independent development (Evolutionary Phonology; Blevins 2004).

Based on the groundwork laid for the Handbook of Phonological Change, the current project's goals include:

- collecting as much empirical data as possible on the phonetic basis of sound change
- creating an electronic, publicly accessible database of regular sound change
- making the database flexible in structure, to accommodate questions asked with different starting assumptions (e.g., genetic relationships, feature systems, phonetic quality of segments, definitions of languages)

Existing Databases

Sound Change Databases

UNIDIA (Ben Haim and Flavier 2009)
- A database for deriving diagnostic universals
- "a data-based typology of sound change, universals, tendencies and sound change distributions"
- 10,349 sound changes, 352 languages

CGBOLD (Comparative Bantu Online Dictionary, UC Berkeley)
- An electronic database to support and enhance the theoretical, descriptive, and historical linguistic study of the languages of the important Bantu family
- Contains a list of reconstructed Proto-Bantu roots and reflexes in 500+ daughter languages

Sound Pattern Databases

PBase (Mithou 2008)
- A database of (syndromic) phonological patterns
- 738 patterns, 620 languages, 8 feature systems, 398 features

Proposed Database

Data Entry: Quality Over Quantity

Many databases include as much information as possible, and leave vetting the quality of the sound change results to the user. We plan to take a different approach by:

- Entering only regular, exceptionless sound changes
- Including the best-documented language families and subgroups
- Relying on language experts when determining the relevant regular sound changes
- Noting how many examples of each sound change are attested

Searchability

Users can investigate sound changes and inventories by general type, phonetic basis, target segment or structure, output segment or structure, language, language family, and other more specific attributes.

Search Results

A database could also aid in the investigation of the notable similarities between common types of historical sound change and common synchronic sound patterns, recurrent synchronic sound patterns observed in unrelated languages that are argued to share regular phonetically based sound change as a common source, and their potential parallel evolution constrained by the perceptual and articulatory properties of human speech claimed to be the force underlying their independent development (Evolutionary Phonology; Blevins 2004).

Based on the groundwork laid for the Handbook of Phonological Change, the current project's goals include:

- collecting as much empirical data as possible on the phonetic basis of sound change
- creating an electronic, publicly accessible database of regular sound change
- making the database flexible in structure, to accommodate questions asked with different starting assumptions (e.g., genetic relationships, feature systems, phonetic quality of segments, definitions of languages)

Individual Sound Changes

Language

Type

Language Family

Proto-Indo-European

Change Description

* m > / h

Output

n

Changes

Order

Environment

Language

Language Family

Subgrouping

Domain

Stress Typ

Other Options

Select

Change Type

Distributions

Other Options

Change Type

Feature System

Search

Don't find a widely accepted underlying phonetic channel? Try...

Altering want to know when one change is part of a series of ordered changes in an awarded seemingly complicated tree...

Details Still Under Consideration

How should language / language family-specific quirks be accounted for?

- The Proto-Indo-European larynxals
- Coding Proto-Mates-Polynesian *N for phonetic qualities
- The orthographic symbol is used for distinct (proto-)phonemes in different language families

What is the best way to represent ordered sound changes? Changes to inventories over time? If it's possible, would visual relative chronologies of changes be helpful?

Contact Information & Questionnaire

I welcome any and all WSC4 attendees may have regarding this project. If you have expertise or experiences you'd like to share, please feel free to email me at spradlin.lauren@gmail.com.

Additionally, please take a moment to answer a questionnaire about what you'd like to see in a sound change database as a user. It's available at https://www.surveymonkey.com/s/WDXCN6H.

Selected References


CGBOLD (Comparative Bantu Online Dictionary, UC Berkeley)
- A database of (syndromic) phonological patterns
- 10,349 sound changes, 352 languages

UNIDIA (Ben Haim and Flavier 2009)
- A database for deriving diagnostic universals
- "a data-based typology of sound change, universals, tendencies and sound change distributions"
- 10,349 sound changes, 352 languages