Simulating eternal optimization: Grimm's law

We show computer simulations of *Grimm's law,* the most famous circular sound change in history. The circularity of this change can be illustrated by how it changed the labial obstruents: p → ph → f → v → b → p. Similar changes occurred for the coronal and dorsal obstruents.

1. Ingredients of the general model

In order to get the simulations working we need a grammar model with at least three levels of representation: (1) the phonological surface form, expressed either in terms of unanalysed phonemes such as /p/ or in terms of feature combinations such as /lab, −voi, −cont/ (depending on the classification that our simulated learners will choose on the basis of the auditory learning data), (2) the auditory-phonetic form, expressed in continuous degrees of spectral (place) information, voicing information, noisiness information, and plosiveness (or continuancy) information (because this is what learners receive as their input), and (3) the articulatory-phonetic form, expressed in continuous degrees of gestural involvement of the muscles responsible for implementing place, voicing, noisiness and plosiveness (because the associated effort is one of the criteria that guides phonetic implementation as soon as the learners grow up and start to speak to the next generation of learners).

Boersma & Hamann (2008) showed that this three-level grammar model can account for chain shifts in sound change, without the need for any teleological mechanisms, provided that one additional assumption is made (following Smolensky 1996), namely that constraints and their rankings are used *bidirectionally*: the language user uses the same constraints in comprehension and production, with the same rankings. In the one-dimensional example by Boersma & Hamann, as well as in its two-dimensional extension by Van Leussen (2008), languages end up in a stable equilibrium; in the one-dimensional case (Boersma & Hamann used a back–front sibilancy continuum), the language always ends up in the same stable state, while in the two-dimensional case (Van Leussen used the vowel backness and vowel height continua), the language ends up in one of a finite number of possible stable states. In the present paper we employ an example with *four* continua (place, voice, noise, plosivity) and show that the language does *not* end up in a stable state but rather continues to evolve in a circular manner.

2. Ingredients of the specific simulation:

Every (simplified) simulated learner is assumed to be able to handle four auditory continua, namely place, voice, noise, and plosivity, each divided into 11 values (i.e. the values run from 0 through 10, in steps of 1). In the environment there are always nine phoneme categories, and these sound like [p, b, f, t, d, θ, k, g, x] in the environment with which the first generation of learners is confronted (listeners may perceive them in terms of the seven feature values lab/cor/dor/±voi/±cont instead); the exact auditory environment is modelled as 36 Gaussian distributions, i.e. the realizations of each of the nine categories can be expressed as a Gaussian distribution on each of the four auditory continua.

The simulated learner assigns an arbitrary label to each of the three phonemes and is assumed to have correct lexical representations that reflect these three categories faithfully. The learner is gifted with a total of 396 *cue constraints*, which connect each of the nine categories to each of the 11 auditory values of each of the four continua; in the initial state of the learner, all these cue constraints are ranked at the same height. The learner is also gifted with 14,641 *articulatory constraints*, one for each possible combination of place, voice, noise, and plosivity values; throughout the learner's acquisition period, these articulatory constraints are ranked according to known effort functions of the muscles involved.

The acquisition procedure follows lexicon-driven learning of perception in Stochastic OT. The language data consists of a series of pairs of auditory form and phonological category, randomly drawn from the distributions of the language environment. Of these, the auditory form is given to the learner by her ear, while the phonological category is given to the learner by her lexicon. That is, the learner will classify the incoming auditory form into one of the nine available categories; after this, the lexicon (on the basis of the available vocabulary and/or the semantic-pragmatic context) tells the learner what category she should have perceived; if this 'correct' category is different from the category that the learner actually did perceive, then she will take action by lowering the rankings of the cue constraints that favoured the learner's own perceived category and raising the rankings of the cue constraints that would have favoured the perception of the 'correct' category (Boersma 1997).

3. Results of the simulation

At the end of the simulated acquisition period, the learner has become an *optimal listener* of the language of her parents, i.e. she perceives any combination of auditory events as the category that was most likely intended by the speaker. However, the learner will subsequently utilize the *same* cue constraint ranking in her own productions, i.e. when she starts to speak. The result is *not* necessarily that she speaks in exactly the same way as her parents do: she has not become an *optimal speaker* of the language of her parents. In fact, if the parents' language has { /p/, /b/, /f/ } with equally voiceless pronunciations of /p/ and /f/, the children will automatically come to pronounce the /f/ category with slightly more voicing than their parents do, whereas the voicelessness of /p/ will not change. Observationally speaking, one could say that /f/ is allowed to move toward [v] because /f/ is not contrastively specified for voice. Within five generations, there will be free variation between [f] and [v]. Importantly, this free variation will be skewed towards [v] as a result of biases inherent in the articulatory constraints and in the *transmission noise* between speaker and listener: two thirds of learners come to pronounce /f/ as [v], one third as [f]. As a result, the following generation will reanalyse this phoneme as /v/, and tend to classify it as [+voi, +cont]. Phonologically, the { p, b, f } inventory has then changed into { p, b, v }. In later generations, this changes again into { ph, b, v }, then into { ph, p, v }, then into { ph, p, b }, then into { f, p, b }. The end result is that the original inventory is arrived at again, but with shifted colours (roughly, this is Latin { **p**ater, la**b**ies, **f**los } versus English { **f**ather, li**p**, **b**loom }).

The 'eternal optimization' witnessed in our simulations applies only if the decision criterion of listeners and speakers is constraint *ranking* (Optimality Theory); if the decision criterion is instead constraint *weighting* (Harmonic Grammar), the language will instead end up in one of a number of possible stable end states. Regarding the fact that Grimm's law happened in reality, this provides some support for constraint ranking over constraint weighting.

References

Boersma, Paul (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences* (University of Amsterdam) **21**: 43–58.

Boersma, Paul, and Silke Hamann (2008). The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* **25**: 217–270.

Leussen, Jan-Willem van (2008). Emergent optimal vowel systems. MA thesis, University of Amsterdam. *Rutgers Optimality Archive* **1006**.

Smolensky, Paul (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry* **27**: 720–731.