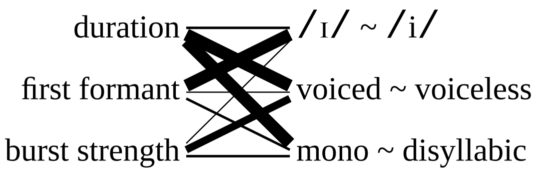
Optimality-Theoretic modelling of microvariation in phonological perception and production

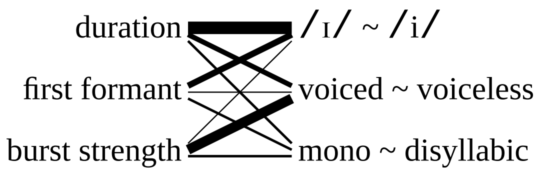
Paul Boersma, University of Amsterdam
Paola Escudero, Utrecht University and McGill University

We will show that the use of continuous constraint families in Optimality Theory can account for fine-grained distinctions between two closely related dialects. In line with the idea that language-dependent 'phonetic' phenomena are to be accounted for within phonological theory (Steriade 1995, Jun 1995, Flemming 1995, Hayes 1996, Kirchner 1998, Boersma 1998), we apply Optimality-Theoretic modelling with continuous constraints to microvariation in the perception and production of phonological contrasts.

As an example, we contrast Scottish Standard English with Southern British English. Both dialects share the following three phonological contrasts: that between the vowels /ɪ/ and /i/, that between voiced and voiceless plosives, and that between monosyllabic and disyllabic words. For word-final and word-medial vowel-plosive sequences, speakers implement these contrasts by varying at least the following three continuous auditory dimensions: the duration of the vowel, the height of the first formant, and the strength of the burst of the plosive. In a production experiment with a Scottish and a Southern speaker, we found that both speakers used most cues for all three contrasts, as the following figures show:



Scottish speaker



Southern speaker

The thickness of the lines shows the extent to which each speaker used each auditory cue to signal each of the three contrasts. The cue weightings in the mapping are different in the two dialects: both speakers signal the /ɪ/-/i/ distinction by means of the first formant, but the Southern speaker uses duration strongly as well; and only the Scottish speaker uses vowel duration as a strong cue for the voicing of the following consonant and for the number of syllables in the word.

The task of the perception process of the listener is to map raw auditory data on discrete phonological categories, which in our case constitutes a true multiple-to-multiple mapping (see the figures). We assume that the optimal listener will do this in such a way that she will be most likely to perceive the speaker's intended categories correctly. Given, for instance, the Scottish speaker's variation in production, a Scottish *maximum-likelihood listener* would identify the three intended categories correctly in 79.7% of the cases.

A true maximum-likelihood listener would have to keep track of the statistics of all possible auditory events (triplets of continuous duration, height, and burst strength values). We argue instead that such an ideal perception process can be modelled quite closely by means of an Optimality-Theoretic *perception grammar* (Boersma 1998) that only consists of simple constraints against the mapping of the three auditory cues to each of the six phonological categories, e.g.: "a duration of 83 ms is not /ɪ/", "a first formant of 310 Hz is not voiceless", "a burst strengh of 0.3 is not disyllabic", "a duration of 110 ms is not voiced", "a first formant of 460 Hz is not monosyllabic", "a burst strength of 1.5 is not /i/". We approximated the continuity of the three auditory cues by dividing each continuum up into 11 regions, so that we had 11 (subdivisions) × 3 (auditory continua) × 6 (categories) = only 198 constraints in total.

We used the Gradual Learning Algorithm (Boersma & Hayes 2001) to determine the optimal ranking of the 198 constraints for a Scottish and a Southern listener. Both simulated listeners started with all constraints ranked at the same height, so that they randomly mapped every possible auditory event to the three binary contrasts, thus achieving correct perception in only 12.5% of the cases. We then fed each listener with 40,000 auditory events based on the productions of a real speaker of her dialect (Scottish or Southern), together with information of what the correct categories were in every case. The simulation resulted in near-optimal behaviour: after learning, the virtual Scottish listener scored 75.3% correct, quite close to the theoretical optimum of 79.7%.[1]

Analogous results are found for the speaker's production grammar: her use of the various cues will be such that the listener is most likely to perceive the intended categories correctly. When applied to the production grammar, the Gradual Learning Algorithm takes care of ranking the faithfulness constraints in an appropriate listener-oriented way, under the restrictions imposed by the other (e.g. articulatory) constraints.

The success of the simulations first shows us that intricate phonological categorization (i.e. the integration of multiple auditory cues to multiple phonological categories) can indeed be modelled by Optimality-Theoretic perception and production grammars, confirming the idea that language-dependent 'phonetic' phenomena are to be accounted for within phonological theory. Most importantly, however, the resulting Scottish and Southern perception and production grammars showed clearly different constraint rankings that reflect the clear differences between the attested weightings of the auditory cues in these two closely related dialects. Only a relatively large fine-grained set of constraints allows Optimality Theory to model the attested differences between closely related phonological systems, in perception as well as in production.

References

Boersma, Paul (1998). *Functional phonology: Formalizing the interactions between articulatory and perceptual drives*. PhD dissertation, University of Amsterdam. The Hague: Holland Academic Graphics.

Boersma, Paul ,& Bruce Hayes (2001). Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* **32**: 45-86.

Escudero, Paola, & Paul Boersma (2001). Modelling the perceptual development of phonological contrasts with Optimality Theory and the Gradual Learning Algorithm. *Proceedings of the 25th Penn Linguistics Colloquium*.

Flemming, Edward (1995). *Auditory Representations in Phonology*. Doctoral thesis, UCLA.

Hayes, Bruce (1999). Phonetically-driven phonology: the role of Optimality Theory and Inductive Grounding. In Michael Darnell, Edith Moravcsik, Michael Noonan, Frederick Newmeyer & Kathleen Wheatley (eds.) *Functionalism and Formalism in Linguistics*, Vol. I: *General Papers.* Amsterdam: John Benjamins. 243–285.

Jun, Jongho (1995). Place assimilation as the result of conflicting perceptual and articulatory constraints. *WCCFL* **14**: 221–237.

Kirchner, Robert (1998). *Lenition in phonetically-based Optimality Theory*. PhD dissertation, UCLA.

Steriade, Donca (1995). *Positional neutralization*. Ms., Department of Linguistics, UCLA.

---

[1] In the simpler case of the mapping of multiple cues to a *single* contrast, simulated learners show true optimal maximum-likelihood behaviour (Escudero & Boersma 2001). In the current, more intricate case, the actual optimum can only approximately be attained with the current simple constraint set. To achieve the remaining 4.3% of correctness, the perception grammar would have to contain the uninterestingly high number of 11 × 11 × 11 × 6 = 7986 constraints.