Annual report 2017 of ACLC research group: *Bidirectional Phonology and Phonetics*

Coordinator: Paul Boersma (back-up: Silke Hamann)

*Web page***:**
- [The members' personal web pages](#)
- [BiPhon](#)

Current external funding:
- Paul Boersma and Mirjam de Jonge: NWO-PhD's in the humanities (€160,000) (2012–2018).

Participants in 2017:
- Paul Boersma (ACLC), senior researcher, coordinator
- Silke Hamann (ACLC), senior researcher
- David Weenink (ACLC), senior researcher
- Jan-Willem van Leussen (ACLC), PhD candidate in Boersma's Vici-project
    - subproject: *The emergence of French phonology*, October 2009 – March 2014
- Mirjam de Jonge (ACLC), PhD candidate (0.8 fte)
    - project: NWO NWO-PhD's in the humanities *Primitives of phonological representations*, September 2012 – July 2018
- Klaas Seinhorst (ACLC), PhD candidate (0.5 fte)
    - project: *The learnability of phoneme inventories*, September 2012 – February 2019
- Jeroen Breteler (ACLC), PhD candidate (1.0 fte)
    - project: *Deconstructing pitch accent: a new perspective on word-prosodic typology*, November 2013 – November 2017; he will defend his thesis on 30 May 2018
- Itsik Pariente, external PhD candidate
    - Project: *The phonology of pharyngeal consonants in modern Hebrew*
- Dirk-Jan Vet, electronic engineer

Description of the research group:
We explain the typology of sound systems by modeling phonology as well as phonetics bi-directionally (i.e. we model the speaker as well as the listener), and by modeling the acquisition and cross-generational evolution of all this. We either model this in a symbolic framework based on strict constraint ranking (Optimality Theory), or in a distributed

framework based on artificial neural networks. If we employ a symbolic framework, we employ at least five representations (one 'semantic', two phonological, two phonetic) and four constraint families that connect these representations to each other. We model the processes of comprehension and production and their acquisition and evolution explicitly with computer simulations, and we test aspects of this model by performing laboratory experiments with adults and infants.

Research highlights in 2017:

- Silke Hamann and Ilaria Colombo showed that orthographic effects in loanword adaptation can be explained by an Optimality-Theoretic comprehension model in which native orthographic constraints (grapheme–phoneme relations) compete with native phonotactic restrictions as well as with native cue constraints (sound–phoneme relations).
- Paul Boersma showed that a simple Deep Belief Network, non-exhaustively trained on basilar excitation patterns from a typical five-vowel language, without supervision, manages to acquire appropriate categorical behaviour: if after training a vowel sound from this language is played to the network, the bidirectional network will turn the auditory input into one of five possible auditory forms, namely the category centre closest to the input.
- Jeroen Breteler computed a typology of Bantu tone shift and tone spread, given his constraint set. It turned out that the typology predicted several language types that do not actually exist. However, when trying to simulate the learnability of the predicted language types, it turned out that the virtual children happened to be capable of learning the attested languages, but not the unattested ones. This is evidence that typologies should not only be predicted on the basis of representability, but also on the basis of learnability.
- Jan-Willem van Leussen ran learning simulations on a large corpus of French, with a model that comprises six levels of representation (three "above" the underlying form), through which it finds optimal paths on the fly. The virtual learners show all kinds of liaison effects, including ones that are typical of human children.

Societal relevance:

—