# Annual report 2016 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) (2012–2016).

*Participants in 2016:*

- 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
- Sophie ter Schure (ACLC/CSCA), PhD candidate in the Brain & Cognition project
    - *Models and tests of early category formation: interactions between cognitive, emotional, and neural mechanisms* on the subject of *Category learning across linguistic and object representation domains* (0.8 fte), September 2010 – September 2014; defended her thesis in April 2016
- Mirjam de Jonge (ACLC), PhD candidate (0.8 fte)
    - project: NWO NWO-PhD's in the humanities *Primitives of phonological representations*, September 2012 – September 2016
- 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
- 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 2016:*

- Sophie ter Schure defended her PhD thesis in April 2016.
- Paul Boersma showed in several talks that a simple "deep" artificial neural network (a stack of two Restricted Boltzmann Machines) exhibits categorical behaviour. The network is first non-exhaustively trained on a three-peaked distribution of an auditory continuum, without supervision. If after this training a sound from anywhere on the continuum is played to the network, the bidirectional network will turn the auditory input into one of three possible auditory forms, namely the category centre closest to the input.
- Learning experiments by Seinhorst showed that the learnability of a phonological pattern is negatively correlated with its complexity, and that learners tend to reduce the complexity of their input. These findings predict that language change moves towards simpler systems, but investigation of a small number of attested sound changes reveals that phoneme inventories sometimes actually become more complex over time: a bias towards simplicity that has been established in the lab can probably be overruled by articulatory and perceptual pressures in spoken language.
- Jeroen Breteler spent the Autumn 2016 period at the University of Delaware, studying with Jeff Heinz, who had discovered that phonological patterns turn out to be surprisingly simple if you describe them as sets or set functions using logical statements. Breteler contributed new research to these efforts, showing that graph copying allows for the expression of a limited set of disjunctions, which further reduces the required expressivity of the formalism.

*Societal relevance:*

—