Infant speech sound acquisition as error-driven discriminative learning of the speech signal

Jessie S. Nixon, Fabian Tomaschek
jessie.nixon@uni-tuebingen.de, fabian.tomaschek@uni-tuebingen.de
Quantitative Linguistics, University of Tübingen, Germany

How do infants learn the sounds of their language? During the first year of life, infants progress from being able to discriminate the sounds of all the world's languages to a system increasingly refined for their native language.

One proposal is that speech cues are learned through exposure to words and their semantic contrasts. This idea has been challenged by the argument that early infant vocabulary is limited and provides an insufficient number of contrasts. Subsequently, statistical acoustic clustering systems have become perhaps the most popular current models of first language acquisition. However, recent empirical evidence and computational modelling suggests that statistical clustering may be insufficient to explain speech sound acquisition.

In the present study, we propose that learning of speech sounds results from error-driven, discriminative learning of the unfolding speech signal. This requires neither semantic contrasts nor similarity-based acoustic clustering. We tested this by training a Naive Discriminative Learner (NDL) network to discriminate 10 ms acoustic sequences from the Karl-Eberhards-Corpus (outcomes/ output), based on acoustic cues from the surrounding signal (input). The trained network was then tested in a vowel continuum to compare the model with human responses in perception tasks.

The development of connection strenths led to

These results suggest that the increases and decreases in perceptual sensitivity to speech sounds observed in infants over the first year of life - before rich lexical knowledge develops - may be at least partially explained by error-driven learning of the speech signal.