

Statistical learning, cross-constraints and the acquisition of speech categories: a computational approach

Joseph Toscano
joseph-toscano@uiowa.edu
Dept. of Psychology
University of Iowa

Bob McMurray
bob-mcmurray@uiowa.edu
Dept. of Psychology
University of Iowa

Infants learning the phonetic categories of their native language must recognize which distinctions are relevant to their language and which are not. While they initially discriminate both native and non-native phoneme contrasts, infants quickly learn to discriminate only those contrasts that are present in their language (Werker & Tees, 1984), and eventually form language-appropriate phonetic categories. One way they might do this is to take advantage of the statistics available in their linguistic environment. Previous work has shown that infants are indeed sensitive to and make use of the distributional statistics for speech sounds (Maye et al, 2002). Infants exposed a series of sound in which phonetic cues formed two clusters learned two categories. Infants exposed to a unimodal distribution learned only one.

We implemented this hypothesis in a computational model. Data representing the distribution of Voice Onset Times (VOTs) for one of several languages were fed into a statistical learning model. These data were based on the statistical distributions of VOT measured by Lisker and Abramson (1964). The model began with a set of Gaussian distributions located at random locations in VOT-space. On each generation, it was given a particular VOT. The model then adjusted the distributions, giving a greater weight to the distribution that best matched the input. Over successive generations, the model was able to fit the input distributions for a variety of languages differing in VOT boundaries and categories. Thus, this form of statistical learning, as implemented in a relatively simple learning device, can learn actual phonetic categories.

We next used the model to examine the role of cross-linguistic patterns on learning. Cross-linguistic similarities may place constraints on the properties of the phoneme categories that must be learned (see Newport & Aslin, 2000, for a similar argument). By varying the starting states of the distributions in the model and evaluating their effect on successful learning, we can determine the relative importance of the initial category locations on the model's performance. If the starting states correspond to categories that are common across languages, they may yield better performance by the model. However, if these starting states provide no advantage, the model's performance will be similar to the condition in which its initial categories are random. This would suggest that statistical learning is a sufficiently powerful mechanism for the acquisition of speech categories without the cross-linguistic constraints. Findings suggest that while statistical learning is sufficient for a most learning-situations, there may be a small benefit to cross-linguistic constraints.

References

- Lisker, A. S., & Abramson, L. (1964). A cross-linguistic study of voicing in initial stops: Acousical measurements. *Word, 20*, 384–422.
- Maye, J., Werker, J., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*, B101–B111.
- Newport, E. L., & Aslin, R. (2000). Innately constrained learning: Blending old and new approaches to language acquisition. In S. C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *Proceedings of the 24th Annual Boston University Conference on Language Development* (pp. 1–21). Somerville, MA: Cascadilla Press.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49–63.