

## CATEGORY LEARNING IN AUDITION, TOUCH, AND VISION

SABINE VAN DER HAM, BILL THOMPSON, BART DE BOER

*Artificial Intelligence Lab, Vrije Universiteit Brussel  
Brussels, Belgium  
sabine@ai.vub.ac.be*

The evolution of linguistic cognition is a notoriously difficult problem: as all learning mechanisms are intertwined with human development and behavior, it is hard to tease apart which aspects of language are the result of cultural processes and which are evolved cognitive traits (De Boer, 2015). We explore whether humans have specializations related to language, and specifically, whether statistical learning of categories is specialized for language-like signals. While it is well-known that statistical learning is not restricted to language learning, the way the mechanism operates in different domains may not be the same, and may be affected by perceptual and cognitive constraints (Conway & Christiansen, 2005).

We present an experiment in which participants learn, categorize, and reproduce signals in the acoustic, visual, and tactile modalities. Participants are trained on bimodal distributions of tones, images and buzzes with some variation in duration, resulting in a ‘long’ and ‘short’ category (Maye, Werker, & Gerken, 2002). After training, participants rate individual signals on a 6 point scale from ‘definitely short’ to ‘definitely long’. The production task consists of creating 3 signals for both categories by pressing the mouse button. The signal is presented as long as the button is pressed. If there is indeed a specialization for learning language-like signals, then we expect that participants reveal: a) more certainty in the categorization task in the auditory and visual conditions, and b) better estimation of the peaks in the distributions, resulting in more reliable reproductions of the categories in the auditory and visual conditions. Alternatively, in normal hearing adults there may be a linguistic training effect, for instance from distinctions between long and short vowels in their native language, in which case better performance in the auditory condition compared to the other conditions is expected, following earlier perception studies (Jones, Poliakoff, & Wells, 2009).

A within subjects design provided us with categorization and production data from 29 participants in all modalities. Results from a logistic regression reveal three interesting trends. First, statistical learning of categories and (reliably) reproducing them is possible in all domains, including the tactile modality. Second, and most importantly, the categorization and production processes are remark-

ably similar in the tactile and auditory domains (odds ratios 0.99 and 1.10; 95% confidence intervals 0.84-1.15 and 0.94-1.28, respectively), but not in the visual domain (OR = 1.30; CIs = 1.12-1.53), suggesting that there is no cognitive specialization for learning language-like signals, nor that there is an effect from previous language experience. Finally, comparing the categorization and production data reveals an interesting tension: on the one hand, we demonstrate that across all three modalities participants were able to induce representations of distinct categories that respect the statistics of the training distributions. On the other hand, we consistently find more variation in duration among participants' productions than was present in the training distributions. This pattern of variation differs between modalities, with more variation in the visual modality. We fitted a Bayesian model of inference for Gaussian distributions to the data in order to investigate whether this variation reflects a meaningful component of the learning process, and to provide a quantitative characterization of the pattern of differences in how participants formed categories in this experiment. The model correctly predicts the overall pattern of categorization behaviour that we find in the empirical data.

Our study further explores the extent to which statistical learning is domain general, as it is becoming increasingly clear that we cannot approach it as a unitary mechanism (Frost, Armstrong, Siegelman, & Christiansen, 2015). Teasing apart differences between modalities will help us find out how the perceptual biases of each modality potentially affect these learning mechanisms, which is necessary for understanding whether certain forms of learning are specialized for language.

### Acknowledgements

This research was supported by European Research Council Starting Grant number 283435, awarded to B. de Boer.

### References

- Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(1), 24.
- De Boer, B. (2015). Biology, culture, evolution and the cognitive nature of sound systems. *Journal of Phonetics*, 53, 79–87.
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in cognitive sciences*, 19(3), 117–125.
- Jones, L. A., Poliakoff, E., & Wells, J. (2009). Good vibrations: Human interval timing in the vibrotactile modality. *The Quarterly Journal of Experimental Psychology*, 62(11), 2171–2186.
- Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82(3), B101–B111.