SEMI-SUPERVISED PHONETIC CATEGORY LEARNING

Till Poppels
Supervised by: Daniel Swingley, PhD

Email: tillpoppels@googlemail.com

Purpose

The learning problem:
•Extracting categorical speech sounds from continuous acoustics (Goudbeek, 2007)

Distributional Learning
•Makes use of statistical cues in the input (Maye & Gerken, 2000)
•Facilitated by lexical context (e.g. Feldman et al., 2009; 2011)

Purpose:
•Better understand the effect of lexical context on the efficacy of distributional learning
•Is it enhanced by the presence of referent pictures?

Training by Exposure

Participants: 36 Penn undergraduates in 2 conditions

Training Conditions: Lexical contexts either aligned or at odds with phonetic categories

Figure 1 Training vowels by first and second formant, coloured and shaped based on lexical context. During the training phase, subjects were exposed to vowels with either low or high second formants, embedded in one of 6 consonant contexts (e.g. /v_t/).

Picture-Word Mapping: one-to-one

Semi-supervised Learning: subjects received no feedback

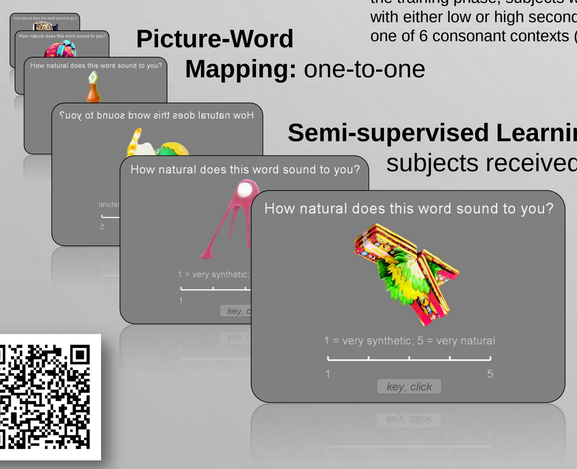

Figure 2 A cover task encouraged subjects to focus on their acoustic properties: on each trial they rated the 'naturalness' of the training words. Stimuli were paired with pictures to highlight their lexical nature.

Test: Categorization Task

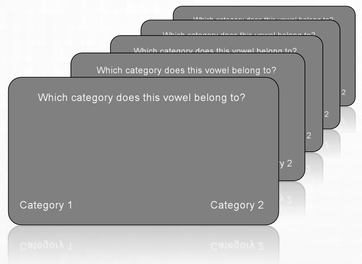

Figure 3 Test displays. Subjects were presented with one vowel per trial and sorted it into one of two categories.

- One vowel per trial
- On each trial: forced-choice categorization
- 2 blocks of 144 trials

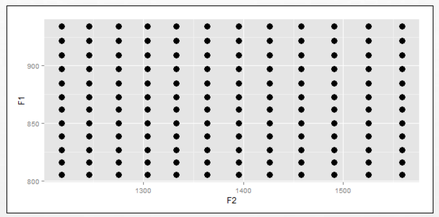

Figure 4 Test vowels by first and second formant.

Individual Performance

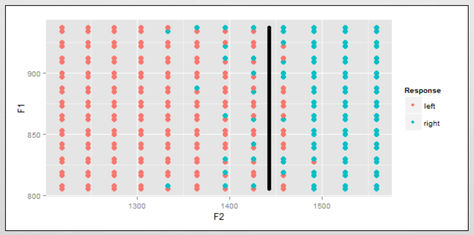

Figure 5 Example of post-test response pattern of one of the top-performers. Based on her individually optimal category boundary (black line), 89% of her responses are correct.

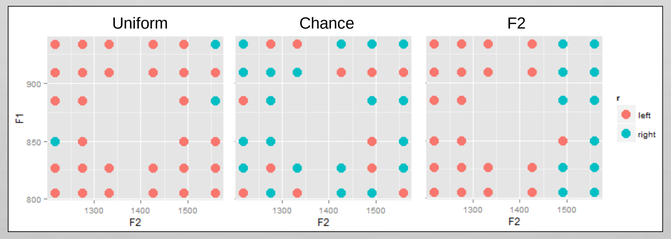

Figure 6 Subjects' pre-test perofrmance fell into one of 3 categories: near-uniform response pattern (left); chance-level categorization (center); F2-based category boundary (right).

Group Results

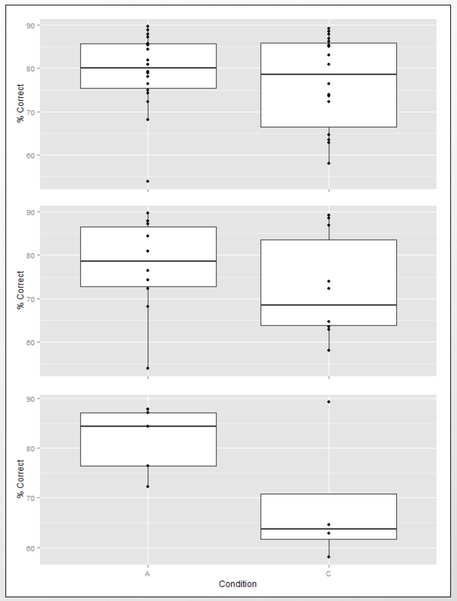

Figure 7 Boxplots comparing post-test performance across training conditions of all subjects (top; *n* = 36), subjects with chance-level or uniform pre-test performance (center; *n* = 20), and only subjects with chance-level pre-test performance (bottom; *n* = 9).

To evaluate the effect of training, we compared post-test performance across training conditions (aligned vs. control). Because too many subjects in both conditions performed at ceiling level, we reanalyzed the data based on individual pre-test performance. All group comparisons were non-significant, $p > .16$.

Future Research

- Replication with more difficult stimulus distribution – to rule out ceiling explanation
- Modify instructions to prevent systematic mis-construal of phonetic space boundaries

References

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). Learning phonetic categories by learning a lexicon. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2208–2213).

Feldman, N. H., Myers, E., White, K., Griffiths, T., & Morgan, J. L. (2011). Learners use word-level statistics in phonetic category acquisition. In *Proceedings of the 35th Annual Boston University Conference on Language Development* (pp. 197–209).

Goudbeek, M. (2007). *The acquisition of auditory categories*.

Maye, J., & Gerken, L. A. (2000). Learning phonemes without minimal pairs. In *Proceedings of the 24th Annual Boston University Conference on Language Development* (Vol. 2, pp. 522–533).