Perception of continuous acoustic cues in speech revealed by the auditory N1 and P3 ERP components

Joe Toscano (jtoscano@illinois.edu)

Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign

Bob McMurray (bob-mcmurray@uiowa.edu)

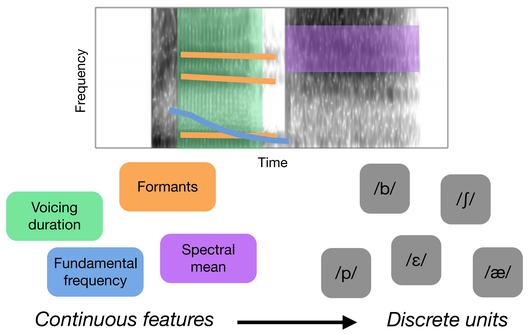
Dept. of Psychology and Dept. of Communication Sciences & Disorders, University of Iowa

Poster reprint

Psychological Science article with Experiment 1

INTRODUCTION

What information is contained in the speech signal?



Two stages of processing:

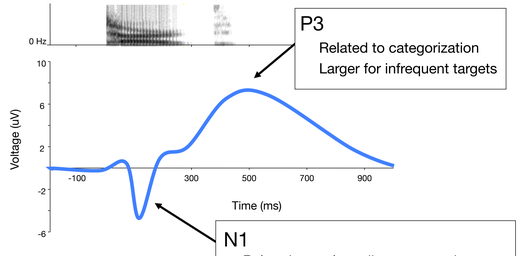
- 1) Initial encoding of speech sounds
- 2) Classification of sounds (categorization)

Is cue encoding based on auditory features or phonological categories?

- Listeners are sensitive to within-category differences, but responses reflect category structure (Liberman et al., 1957; Pisoni & Lazarus, 1974)
- Suggests that either (1) encoding is shaped by categories, or (2) behavioral tasks aren't measuring encoding

APPROACH

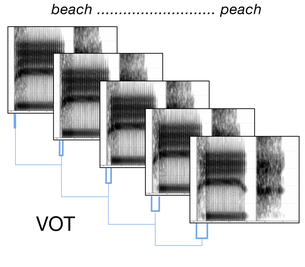
ERP approach to isolate encoding and categorization (Toscano, McMurray, Dennhardt, & Luck, 2010)



EEG recording

- N1: average of F3, Fz, and F4 channels
- P3: average of P3, Pz, P4 channels
- Average mastoid reference
- Impedance ≤ 5 kΩ

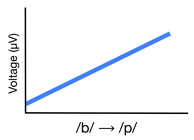
Stimuli varied along continuous acoustic dimensions



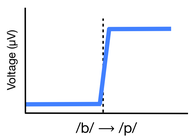
PREDICTIONS

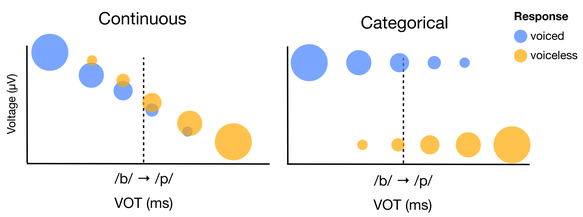
N1 predictions

- If encoding is based on auditory features, monotonic response
- If encoding is based on categories, nonlinear response centered on category boundaries



Analyzed data grouped by behavioral response to ensure effects are not an artifact of averaging across categorical differences





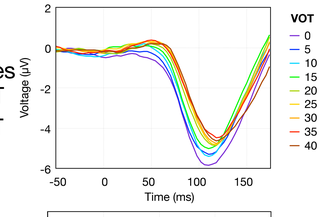
P3 predictions: If sensitivity to acoustic detail is maintained at late stages of processing, within-category variation in P3 amplitude

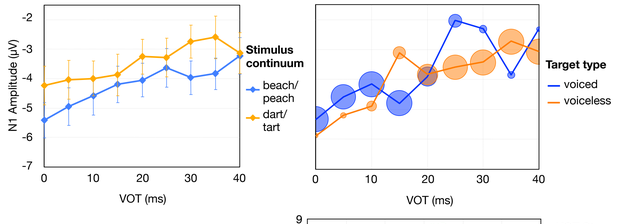
EXPERIMENT 1

Synthetic stimuli varying along two VOT continua (*beach-peach* and *dart-tart*); target detection task

N1 results

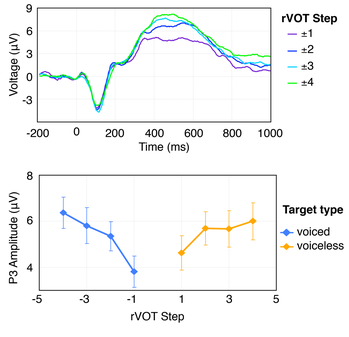
- Amplitude decreases with increasing VOT
- Linear effect of VOT on N1 amplitude





P3 results

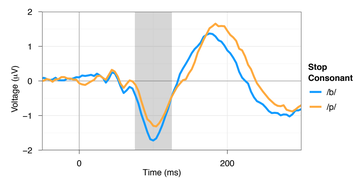
- P3 observed for infrequent targets
- Relative VOT (rVOT) computed from category boundaries
- Amplitude varied within each voicing category



EXPERIMENT 2

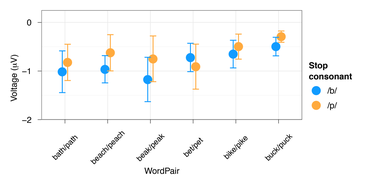
Can we measure differences in N1 amplitude for other types of speech sounds (e.g., natural speech instead of synthetic speech)?

Examined ERPs to naturally-produced stop consonants (/b/-/p/ minimal pairs)



Results

- N1 amplitude smaller for /p/ than for /b/
- Same pattern of results as Experiment 1

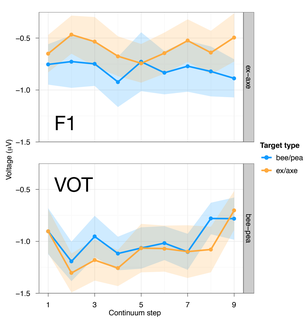


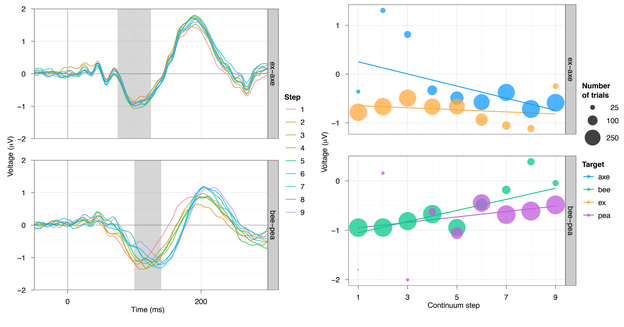
EXPERIMENT 3 — N1

Examined different acoustic and phonological dimension — vowel differences indicated by F1 frequency

N1 Results

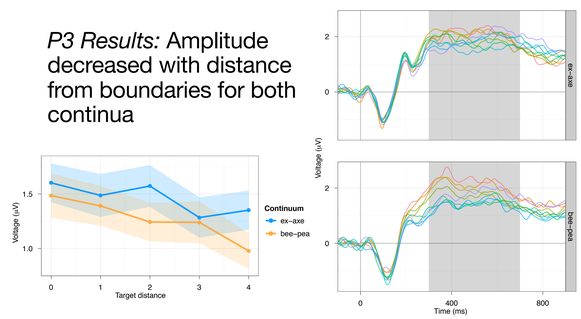
- No overall effect of F1; significant effect for target-response trials
- Encoding may be difficult to measure for some cues due to orientation of generators

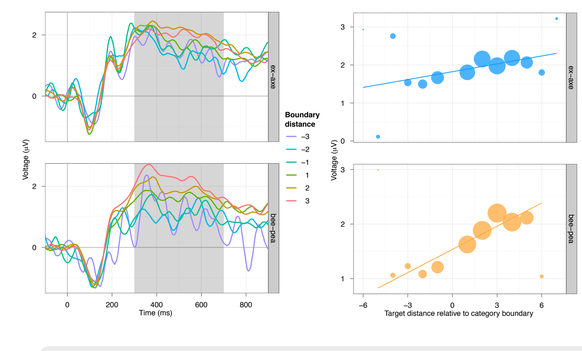




EXPERIMENT 3 — P3

P3 Results: Amplitude decreased with distance from boundaries for both continua





CONCLUSIONS

New tool for measuring online speech processing

- Auditory N1 reflects cue encoding
- P3 reflects categorization

Can be applied to various acoustic cues and classes of phonemes, and both natural and synthetic speech

However, some acoustic cue differences may not be observable in ERP response

What can this tell us about speech perception?

- Continuous acoustic cues encoded independently of phoneme categories
- Listeners maintain acoustic sensitivity after cue encoding
- Overall, supports models that harness fine-grained acoustic information in the signal

ACKNOWLEDGEMENTS

We would like to thank Steve Luck and Joel Dennhardt, who worked with us on Experiment 1; Toby Mordkoff for use of lab resources and helpful theoretical discussions of these issues; and Dan McEchron and Lisa Vangsness for assistance with participant recruitment, data collection, and analysis. Funding: NIH DC008089.

REFERENCES

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358-368.

Picton, T. W., Woods, D. L., & Proulx, G. B. (1978). Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalography and Clinical Neurophysiology, 45*, 198-210.

Pisoni, D. B., Lazarus, J. L. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *Journal of the Acoustical Society of America, 55*, 328-333.

Toscano, J. C., McMurray, B., Dennhardt, J., & Luck, S. J. (2010). Continuous perception and graded categorization: Electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychological Science, 21*, 1532-1540.

Woldorff, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology, 30*, 98-119.