Measuring perceptual encoding and categorization of speech sounds using an ERP approach

Joe Toscano (jtoscano@illinois.edu)

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

Bob McMurray (bob-mcmurray@uiowa.edu)

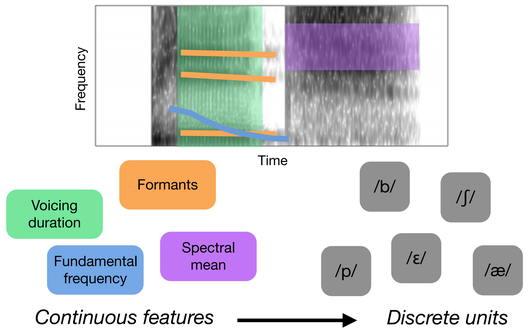
Dept. of Psychology and Dept. of Communication Sciences & Disorders, University of Iowa

Scan here for poster reprint

Psychological Science article with Experiment 1

Background

What information is contained in the speech signal?



Two stages of processing:

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

Is cue encoding based on auditory features or phonological categories?

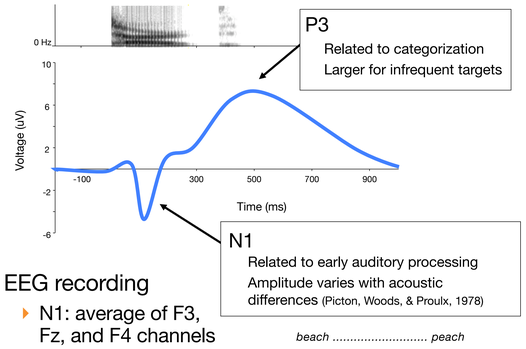
- Problem is typically understood in terms of categorical perception (Liberman, Harris, Hoffman, & Griffith, 1957)
- Listeners are sensitive to within-category differences, but responses reflect category structure (Pisoni & Lazarus, 1974)

This suggests that either

- 1) Encoding is shaped by categories
- 2) Behavioral tasks aren't measuring encoding

Approach

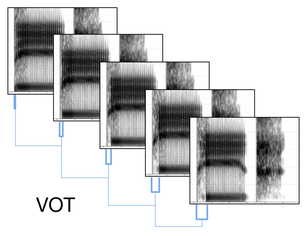
Measured ERPs to speech sounds to isolate encoding and categorization (Toscano, McMurray, Dennhardt, & Luck, 2010)



EEG recording

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

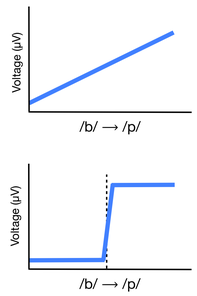
Stimuli varied along continuous acoustic dimensions



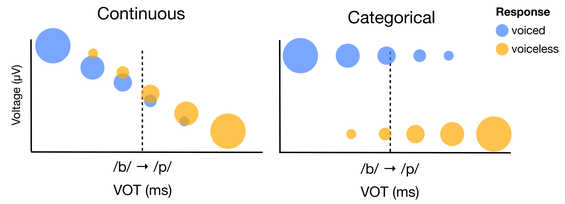
Predictions

N1 predictions

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



Also analyzed data grouped by behavioral response — ensures that effects are not are an artifact of averaging across categorical differences)



P3 predictions

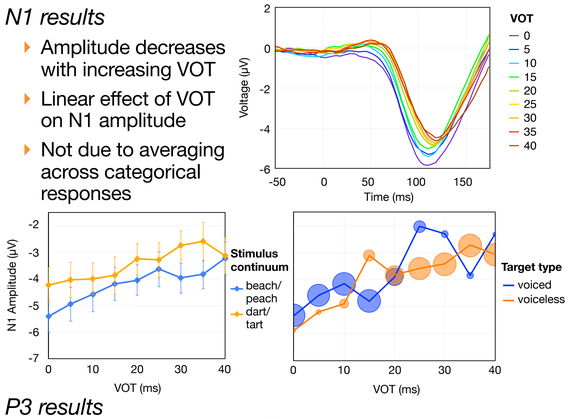
- If sensitivity to acoustic detail is maintained at late stages of processing, within-category variation in P3 amplitude (reflecting categorization)

Experiment 1

Presented synthetic speech that varied along two VOT continua (*beach-peach* and *dart-tart*)

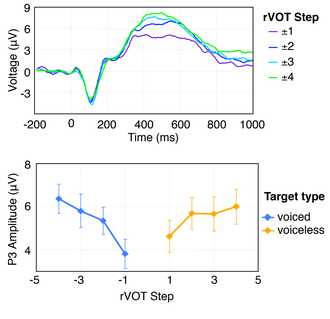
N1 results

- Amplitude decreases with increasing VOT
- Linear effect of VOT on N1 amplitude
- Not due to averaging across categorical responses



P3 results

- Listeners performed a target detection task to produce a P3 response
- P3 observed on infrequent target trials
- Relative VOT (rVOT) computed from category boundaries
- Amplitude increases with distance from boundary and varies within each category

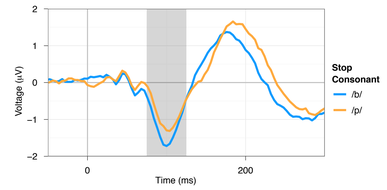


Experiment 2

N1 results indicate that listeners are sensitive to continuous acoustic differences at early stages

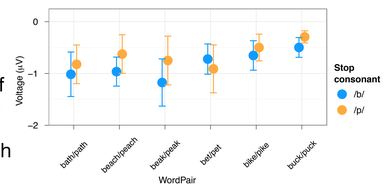
Can we also measure differences in N1 amplitude for other types of speech sounds (e.g., natural speech instead of synthetic speech)?

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



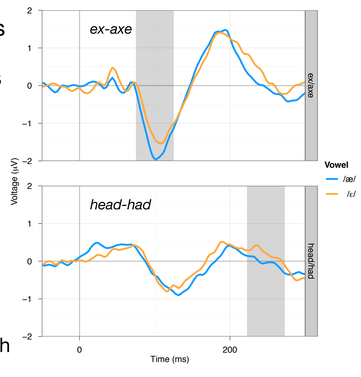
Results

- N1 amplitude smaller for /p/ than for /b/
- Same pattern of results that was observed for synthetic speech



Experiment 3

Can we observe effects for contrasts cued primarily by spectral differences (e.g., vowels)?



Results

- Larger N1 for /æ/ in axe/ex stimuli
- had/head difficult to measure due to overlap from preceding /h/
- Later experiment addressed this with Adjar (Woldorff, 1993)

Direction of effect suggests that the N1 does not simply reflect overall acoustic frequency differences

Larger for

- Voiced sounds (lower F0; Experiments 1 and 2)
- Vowels with higher formant frequencies (Experiment 3)

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.

Experiment 4

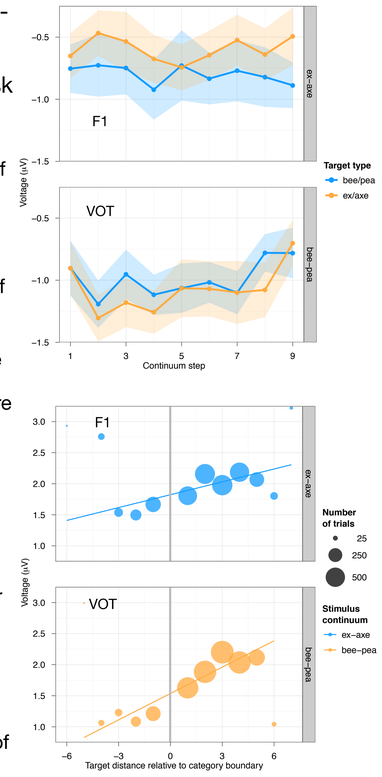
Presented naturally-produced F1 and VOT continua in target detection task

N1 results

- No overall effect of F1; significant effect for target-response trials
- Direction of effect matches results of Experiment 3
- Suggests that encoding of some cues may be difficult to measure due to orientation of generators for those cues

P3 results

- Significant effect across continuum (and boundary) for both continua
- P3 amplitude decreased with distance from boundary
- Similar to results of Experiment 1



Conclusions

A 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
- both natural and synthetic speech

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

How do listeners map cues onto categories?

- Listeners initially encode continuous acoustic cues independently of phoneme categories
- Acoustic sensitivity is maintained after cue encoding
- Overall, this supports models of speech perception that harness fine-grained acoustic information

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.