Auditory N1 Amplitude Varies Across Multiple Acoustic and Phonological Dimensions in Speech

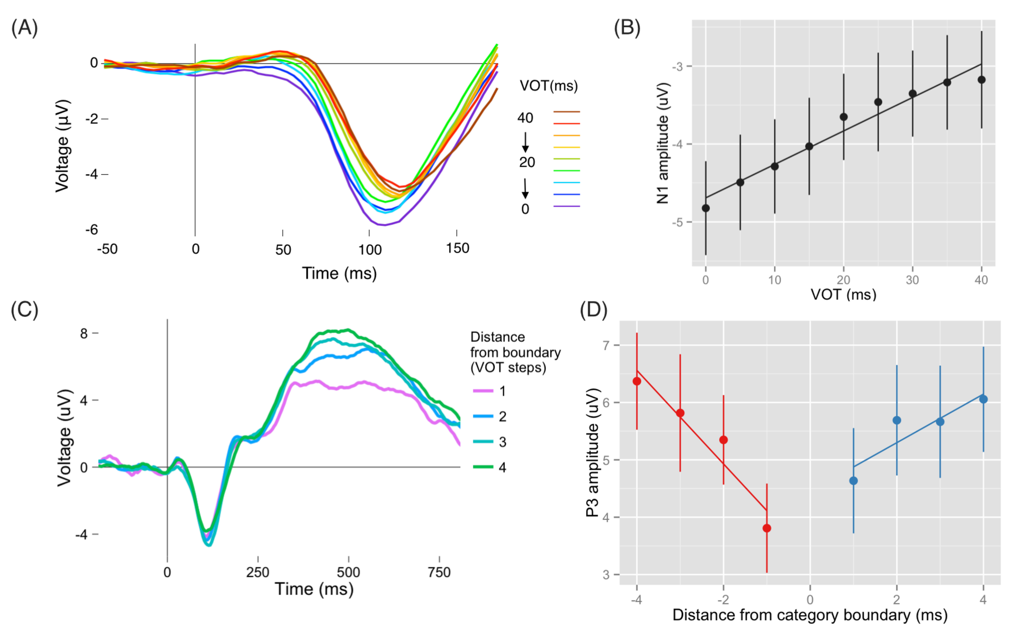
Olivia Pereira (opereira@villanova.edu)
Department of Psychology
Villanova University

Joe Toscano (joseph.toscano@villanova.edu)
Department of Psychology
Villanova University

WRAP Lab
http://wraplab.co/

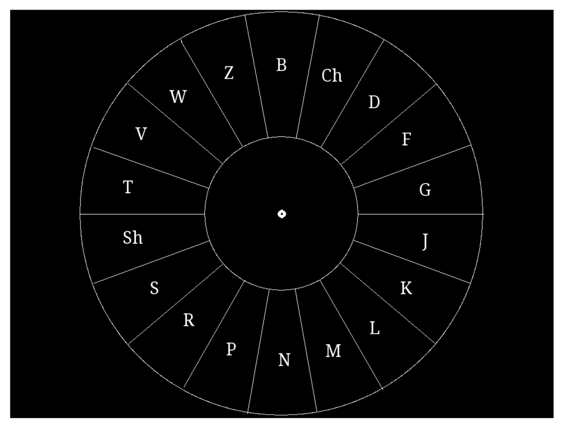
INTRODUCTION

- Early work suggested that fine-grained differences between phonetic cues were discarded early in perception, but recent evidence has refuted this claim [1-3]
- However, it is not clear how the perceptual system encodes specific phonetic cues at early stages
- Previous work has used the ERP technique to address this [4]:
  - Auditory N100 component varies linearly with differences along VOT continua
  - P300 varies as a function of phonological category, while retaining sensitivity to within-category differences
- We aimed determine how N1 varies for other cues across a wide range of consonant distinctions



METHOD

- Subjects heard word-initial minimal pairs spanning 18 consonants (/b,ʧ,d,f,g,ʤ,k,l,m,n,p,r,s,ʃ,t,v,w,z/)
- Listeners indicated the sound that each word began with by clicking corresponding letter(s) on a display

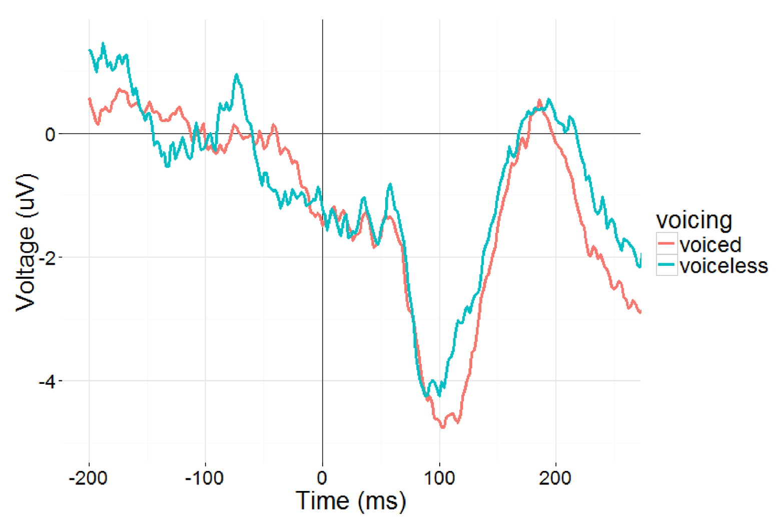


- EEG data recorded using a 64-ch Brain Vision actiCHamp system
- Electrodes placed at 10-20 sites (F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, T7, T8, O1, Oz, O2, CP1, CP2, FC1, FC2, A1, A2, FP1, FP2)
- EOG recorded via electrodes at FP1 and FP2 (vertical) and electrodes lateral to the external canthi of each eye (horizontal)
- Electrode impedences <5 kΩ
- Data referenced online to the left mastoid (A1), and referenced offlined to an average mastoid reference
- EEG was digitized at a sampling rate of 500 Hz and recorded in DC mode; offline bandpass filter from 0.1 to 125 Hz was applied in ERPLAB (Butterworth filter with 36 dB/octave roll-off)
- Mean ambient room temperature in recording chamber was 19.7° C and mean relative humidity was 29.6%

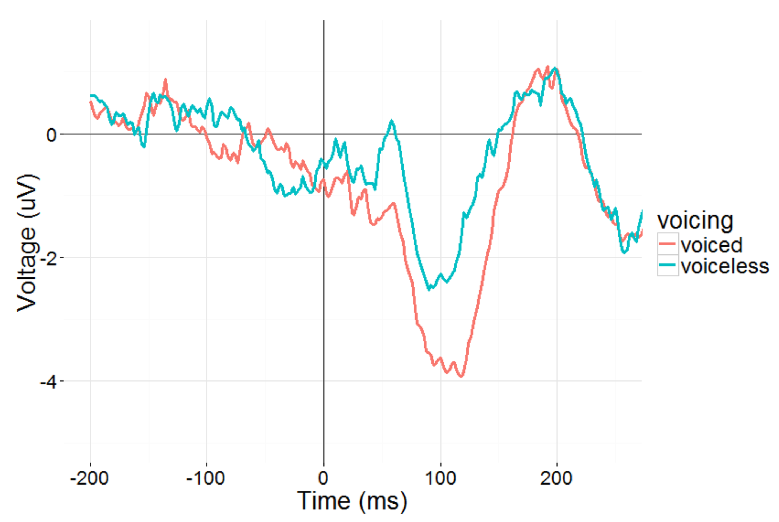
RESULTS

Voicing Contrasts

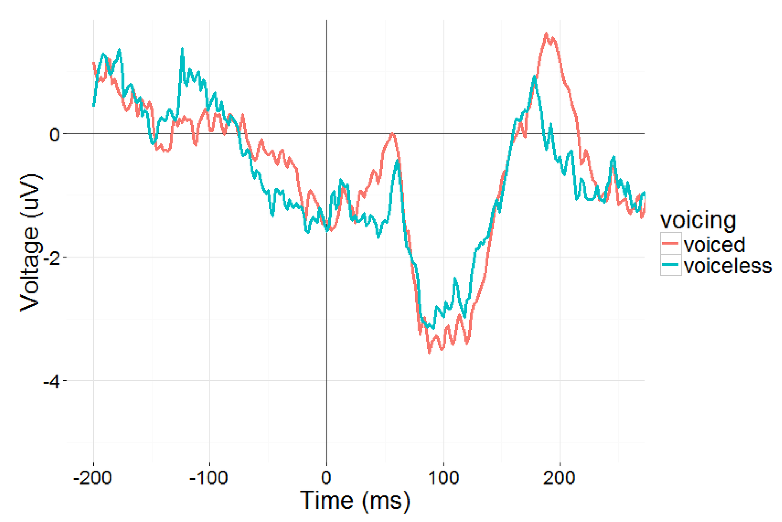
Bilabial Stop Consonants (/b/ vs /p/)



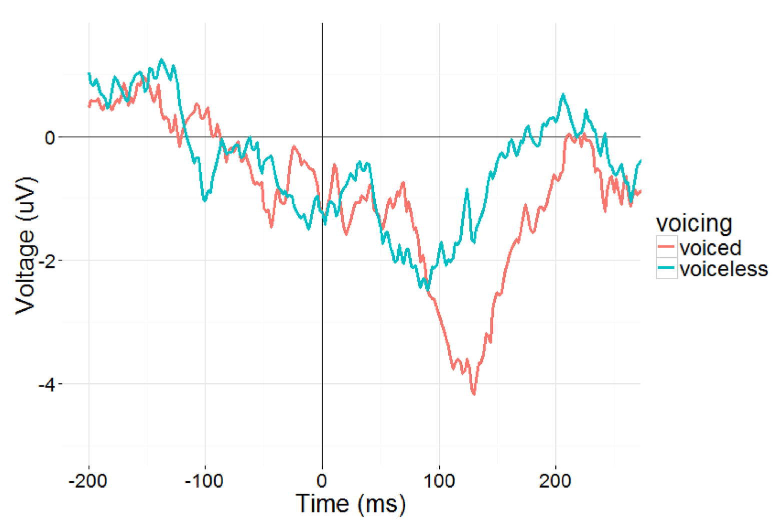
Alveolar Stop Consonants (/d/ vs /t/)



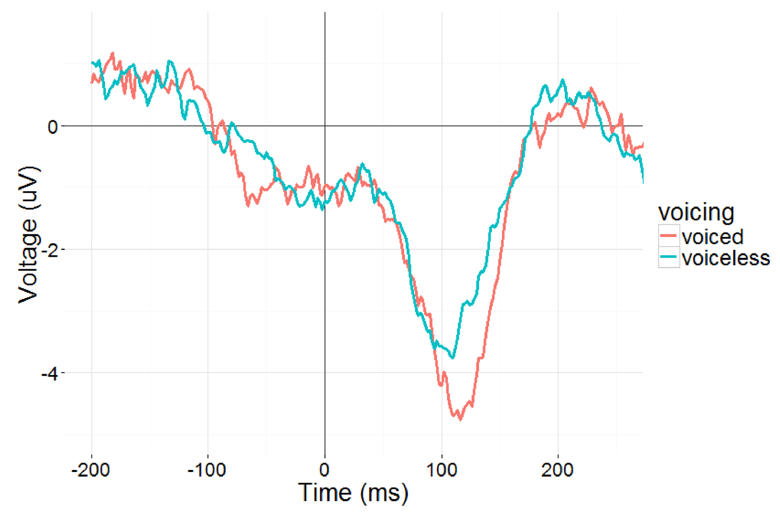
Velar Stop Consonants (/g/ vs /k/)



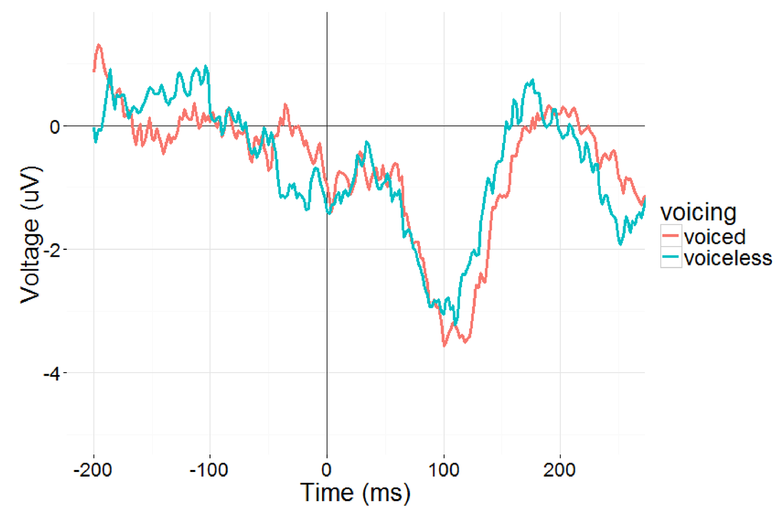
Labiodental Fricative Consonants (/v/ vs /f/)



Alveolar Fricative Consonants (/z/ vs /s/)

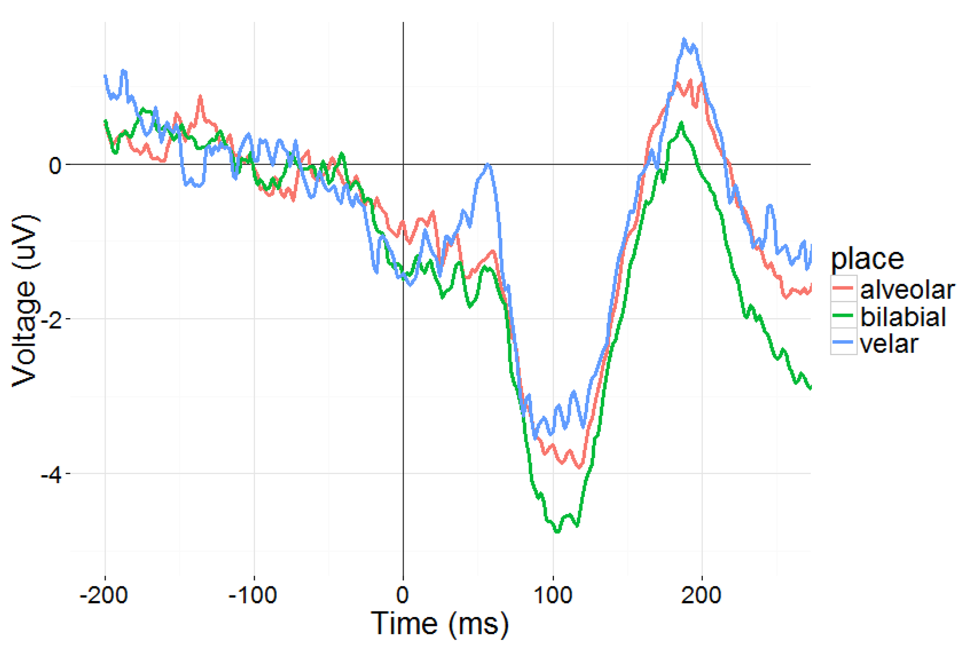


Alveopalatal Affricate Consonants (/ʤ/ vs /ʧ/)

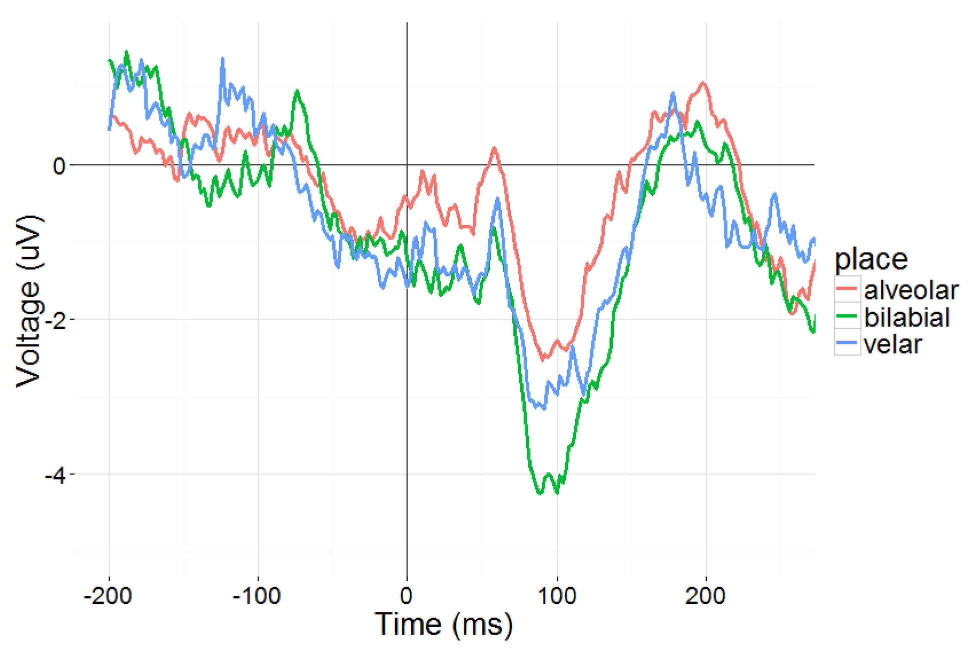


Place of Articulation Contrasts

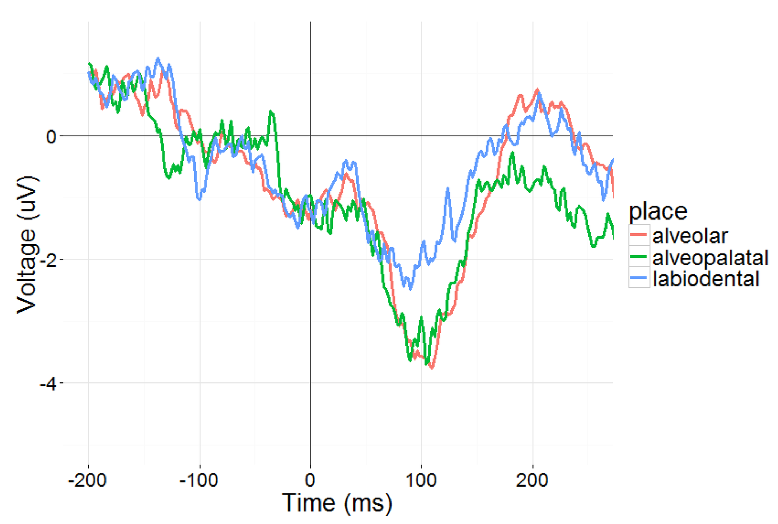
Voiced Stops (/d/ vs /b/ vs /g/)



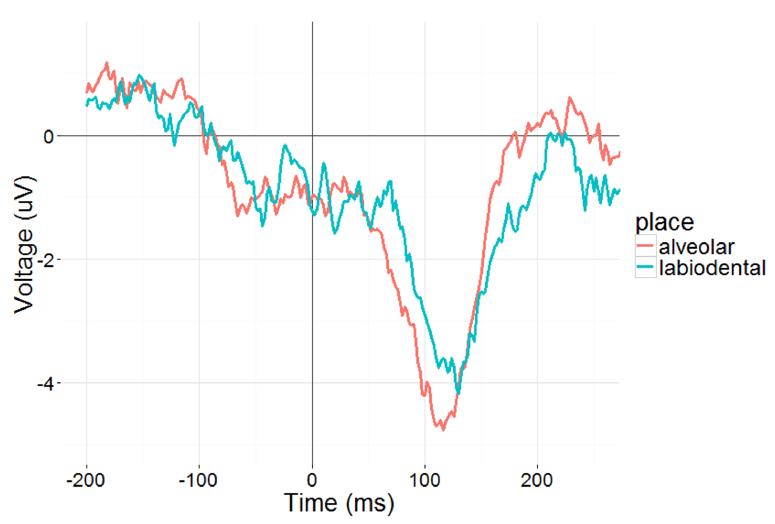
Voiceless Stops (/t/ vs /p/ vs /k/)



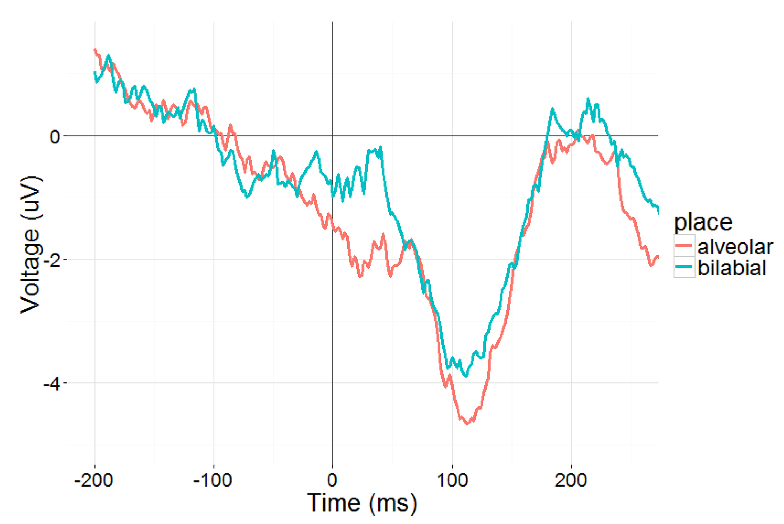
Voiceless Fricatives (/f/ vs /ʃ/ vs /s/)



Voiced Fricatives (/z/ vs /v/)



Voiced Nasals (/n/ vs /m/)



DISCUSSION

- Widespread differences in N1 amplitude as a function of variation in both voicing and place of articulation for stops, fricatives, and nasals
- **Voicing**
  - Consistently larger N1 for voiced consonants than voiceless consonants
  - Replicates and extends findings from Toscano et al. (2010) [4]
  - N1 may be indexing degree of low-frequency voicing energy that occurs consistently across consonants with different manners of articulation
- **Place of articulation**
  - Pattern of responses varies depending on manner of articulation
  - N1 amplitude is greater for alveolar fricatives and nasals than for bilabial and labiodental sounds
  - No clear distinction between /s/ and /ʃ/
  - For stops, two different patterns emerge:
    - Voiced stops organized based on place of articulation (/b/ → /d/ → /g/)
    - Voiceless based on acoustic differences in burst frequency (/p/ → /k/ → /t/)
    - Difference has implications for the nature of representations used to encode speech sounds (articulatory v.s. acoustic features)
- **Conclusions**
  - Overall, results suggest that the N1 can serve as a useful tool for studying cue encoding at early stage of perception across a wide range of speech distinctions
  - Future work will examine N1 response to stimuli varying along voice onset time and burst frequency continua for stop consonants.

ACKNOWLEDGEMENTS & REFERENCES

- We would like to thank Nicole Johnson, Chris Burley, and Jackie Coelho for assistance with data collection, and Emma Folk for stimulus recording and subject recruitment.

- **References**

  - Liberman AM, Harris KS, Hoffman HS, Griffith (1957). *Journal of Experimental Psychology, 54*, 358-368.
  - Pisoni DB, Tash J (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics, 15*, 285-290.
  - Miller JL (1994). On the internal structure of phonetic categories: A progress report. *Cognition, 50*, 271-285.
  - 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.