

Neural Responses to Target Words Are Modulated by Semantic and Acoustic Similarity of Temporally Offset Distractors

Gidey W. Gezae¹, Sahil Luthra², Marysia Brown³, Eric Parker⁴, Hee So Kim⁴, Wusheng Liang⁴, Abigail Noyce⁴, Barbara Shinn-Cunningham⁴
¹Pennsylvania State University, ²Stony Brook University, ³University of Pittsburgh, ⁴Carnegie Mellon University

Segregation vs. Selection

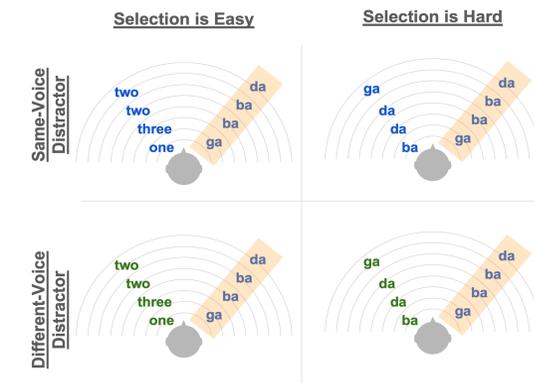


- Successful auditory attention requires both **segregating** simultaneous acoustic streams and **selecting** one for further analysis
- **Voice characteristics** such as pitch, rhythm and intonation differences facilitate auditory attention; in the literature, this is assumed to specifically benefit segregation^{1,2,3,4}
- However, studies showing a benefit of voice differences have typically used **linguistically similar streams**, which may also challenge selection

Does the benefit of different voice characteristics for auditory attention depend on selection difficulty?

Task & Experiment Design

- Target and distractors were temporally offset
- Target location fixed left/right ($\pm 30^\circ$ azimuth) for each participant
- Target timing (leading/lagging re: distractor) manipulated within subjects



EEG Methods

- Data were collected with a 32-channel BioSemi system
- Data were preprocessed in EEGLAB (bandpass filtered at 0.1-50 Hz, down-sampled to 256 Hz, artifacts rejected, eye/muscle components removed using ICA) and epoched (baselined to syllable onset)

Participants (N=22)

- Ages 18-28 (mean: 22)
- Hearing thresholds of ≤ 20 dB HL
- Normal/corrected-to-normal vision
- No known neurological disorders

Discussion

Voice differences did not facilitate performance when selection was easy, but preliminary ERPs indicate voice differences nonetheless alter neural coding

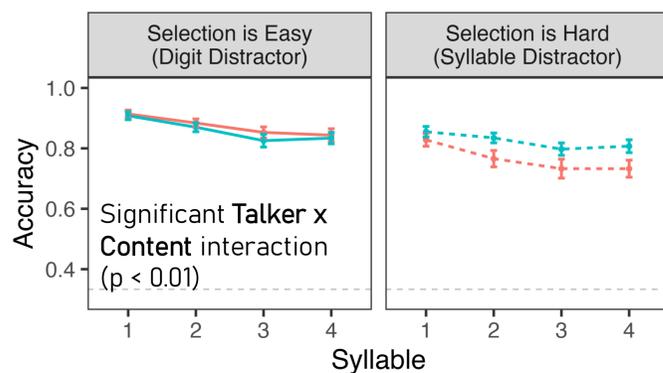
The benefit of voice (acoustic) differences between streams depends on selection demands (semantic similarity)

Target ERPs indicate how the brain filters out distracting information: when target and distractor are similar, less of target stream gets through

Results

Behavioral Analysis

✓ Mean accuracy: 83%

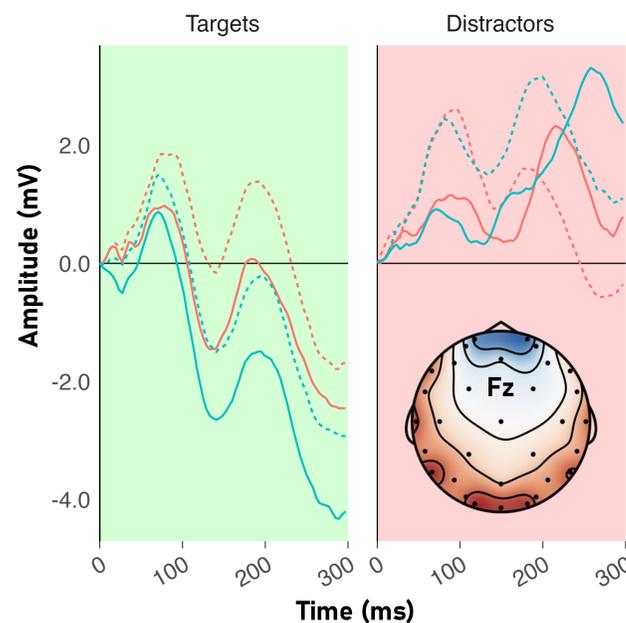


Distractor Content
 Digits (Easy) Syllables (Hard)

Distractor Talker
 Same (Orange) Different (Blue)

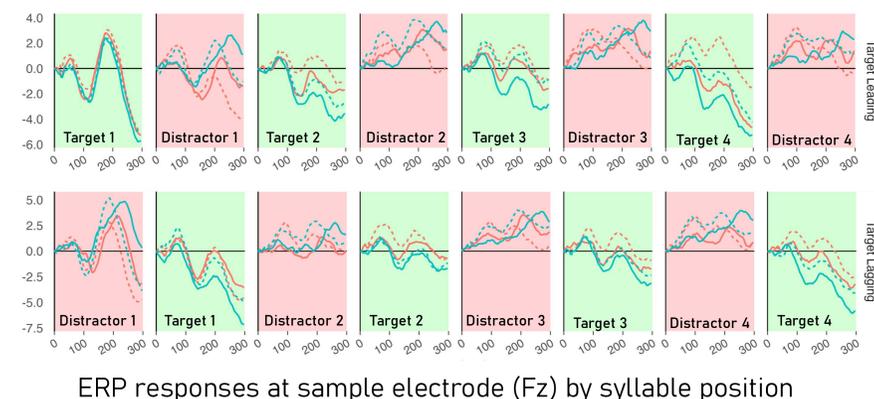
If selection is easy, talker differences have no benefit; if selection is hard, talker differences benefit performance

Event-Related Potentials Analysis



ERP responses at sample electrode (Fz), averaged over positions (excluding first window)

- ✓ ERP responses to target stimuli (always syllables spoken in male voice) differed based on whether **temporally offset distractors** were similar or different
- ✓ We observed **large peak-to-peak differences** in P1 and N1 amplitudes to target stimuli when distractors differed from the target, and smaller differences when they were similar.
- ✓ Distractors that are linguistically similar to targets drive stronger, earlier neural responses (but this may be driven partly by acoustic properties of the stimuli)



ERP responses at sample electrode (Fz) by syllable position

Conclusion

When selection is **difficult**, hearing the **same voice** in both streams leads to an **accuracy cost**

Future Work

- ✓ Data collection is ongoing (target N = 30)
- ✓ Future work might consider which aspect of voice differences (e.g., pitch, accent) influence listener ability to selectively attend

Acknowledgements

This work was supported by a grant from the Office of Naval Research (ONR N00014-23-1-2065, PI: Shinn-Cunningham). GWG was supported by a SURIEA fellowship from ASA. SL was supported by NIH NRSA F32DC020625. Iris Ta assisted with data collection.

References

¹Assmann, P. F. (1999). Fundamental frequency and the intelligibility of competing voices. In Proceedings of the 14th International Congress of Phonetic Sciences (pp.179-182). Oakland, CA: University of California Press.

²Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. The Journal of the Acoustical Society of America, 109(3), 1101-1109.

³De Cheveigné, A., Kawahara, H., Tsuzaki, M., & Aikawa, K. (1997). Concurrent vowel identification. I. Effects of relative amplitude and F difference. The Journal of the Acoustical Society of America, 101(5), 2839-2847.

⁴Shackleton, T. M., & Meddis, R. (1992). The role of interaural time difference and fundamental frequency difference in the identification of concurrent vowel pairs. The Journal of the Acoustical Society of America, 91(6), 3579-3581.

⁵Joos, K., Gilles, A., Van de Heyning, P., De Ridder, D., & Vanneste, S. (2014). From sensation to percept: The neural signature of auditory event-related potentials. Neuroscience & Biobehavioral Reviews, 42, 148-156