

# Listening at a cocktail party: Does our ability to follow one conversation depend on the acoustic and linguistic similarity of other conversations nearby?

Marysia Brown<sup>1</sup>, Sahil Luthra<sup>2</sup>, Eric Parker<sup>3</sup>, Gidey W. Gezae<sup>4</sup>, Hee So Kim<sup>3</sup>, Wusheng Liang<sup>3</sup>, Abigail Noyce<sup>3</sup>, Barbara Shinn-Cunningham<sup>3</sup>

<sup>1</sup>University of Pittsburgh, <sup>2</sup>Stony Brook University, <sup>3</sup>Carnegie Mellon University, <sup>4</sup>Pennsylvania State University

Voice differences only help improve attention when target selection is challenging.

## INTRODUCTION

To pay attention to a target source in a noisy environment, we need both to **segregate** the streams we are hearing and **select** one for further analysis.

**Voice differences** across streams can help with segregation<sup>1,2,3,4</sup>

Studies showing a benefit in voice differences have typically used **linguistically similar streams**<sup>1,2,3,4</sup>, which can make selection harder.

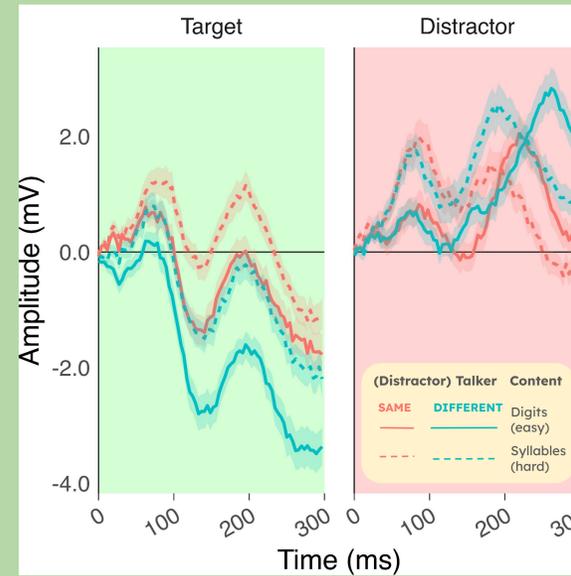
Do voice differences provide as much benefit for auditory attention when selection difficulty changes?

## EEG RESULTS

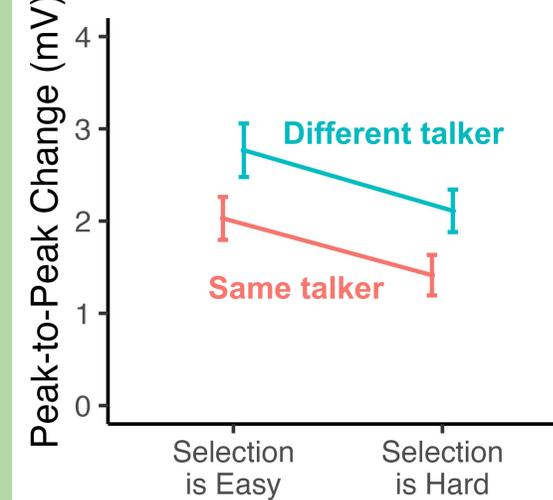
### Data Collection and Processing

Collected with a 32 channel BioSemi system

Data 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)



### P1-N1 (Response to Target)



- Early ERP responses<sup>5</sup> to target stimuli differed based on whether distractors were similar or different
- Distractors that are linguistically similar to targets drive stronger, earlier neural responses (but this may be driven partly by acoustic properties of the stimuli)
- When targets and distractors were similar, P1-N1 peak-to-peak amplitudes elicited by the target were smaller

## EXPERIMENTAL DESIGN

**TASK:** What are the four target syllables?



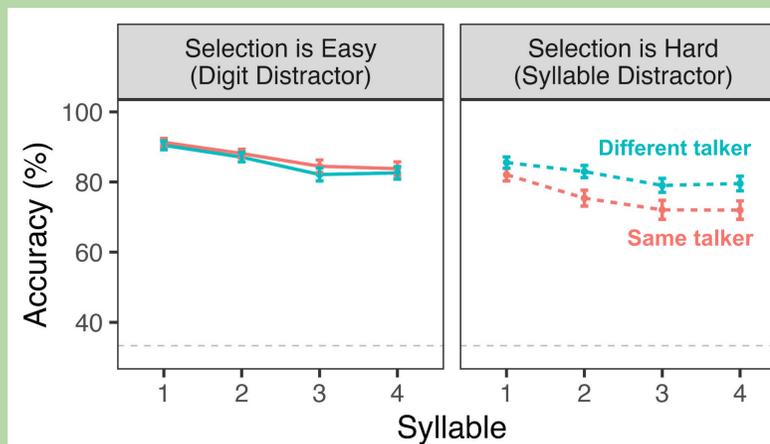
- Target and distractors were temporally offset
- Counterbalanced whether target led or lagged behind distractor

### PARTICIPANTS

(n=30)

- Ages 18-28
- Native English speakers
- No known neurological disorders
- No hearing impairment (thresholds  $\leq 20$ dB HL)

## BEHAVIORAL ANALYSIS



Mean Accuracy: **83%**

Significant **Talker x Content** interaction ( $p < 0.01$ )

If selection is **easy**, talker differences have **no benefit**;

If selection is **hard**, talker differences **benefit** performance

## DISCUSSION

Differences in voice did not improve performance when selecting the target stream was easy. However, ERP results suggest that voice differences still change how the brain processes the information.

The advantage of voice differences between streams depends on how difficult the selection process is, which is influenced by how linguistically similar the streams are.

**Target ERPs show how the brain filters out distracting information. When the target and distractor are more similar in meaning, less of the target stream is successfully processed.**

### Acknowledgements:

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

### References:

- Assmann, 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.
- Brungart, 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.
- de 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.

<sup>4</sup> Shackleton, 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.

<sup>5</sup> Joos, 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.