The Cocktail Party Effect: Is our ability to listen to one conversation impacted by an acoustically and linguistically similar conversation nearby?

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

[1]Carnegie Mellon University, [2]Stony Brook University, [3]University of Pittsburgh, [4]Pennsylvania State University

Introduction

To pay attention in noisy, cocktail party-like environments, a listener must segregate the auditory streams they are hearing and select one to attend to.

Low-level voice differences like pitch, rhythm, and intonation have been shown to aid with segregation[1,2,3,4], as well as high-level semantic ones such as word content.

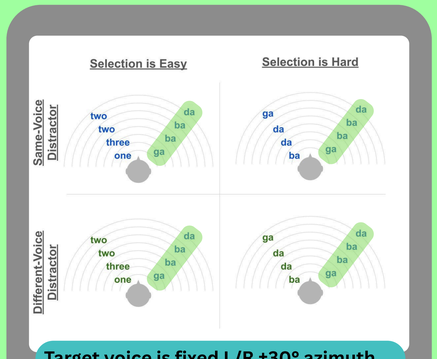
Previous EEG studies have shown that low-level filtering takes place relatively early on, but it's unknown whether semantic differences similarly support early filtering.

Could the benefit of vocal differences vary based on selection difficulty?

Experimental Design

What were the four target syllables?

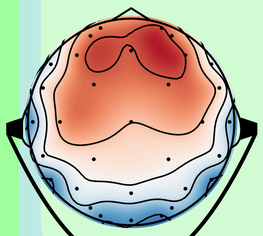
- Target and distractor streams were temporally offset
- Counterbalanced whether target led or lagged behind distractor



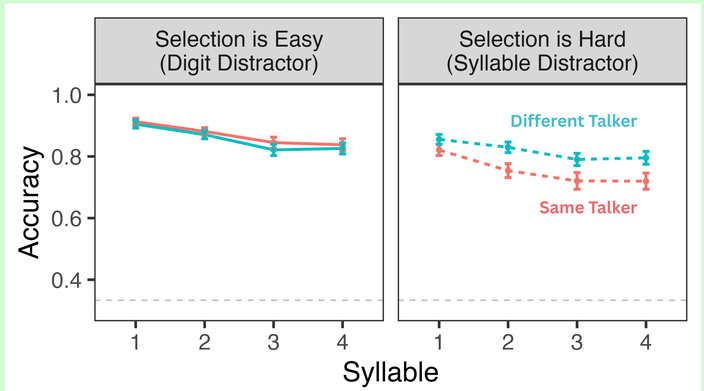
Target voice is fixed L/R ±30° azimuth

Participants (n = 30)

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



Behavioral Analysis

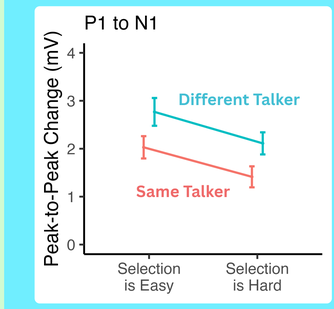


Mean Accuracy: 83%

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

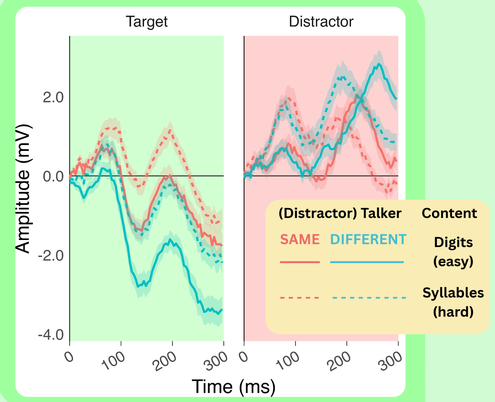


EEG Analysis

- Differences in early ERP[5] responses were shown between distractor conditions
- When distractors were linguistically similar to targets, earlier neural responses were observed
- In addition, P1-N1 peak-to-peak amplitudes elicited by the target were also smaller in this condition

EEG Methods

- Data were collected using a 32 channel BioSemi system
- Data were then 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)



Discussion

Vocal differences did not benefit performance when target selection was easy. However, ERPs indicate that these vocal differences may still alter neural processing.

The benefit of vocal differences depends on selection difficulty, which is influenced by linguistic similarity between streams.

Target ERPs indicate how well the brain is filtering out distracting information. When targets and distractors are linguistically similar, less of the target stream makes it to processing.

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

1. 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.
2. 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.
3. 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.
4. 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.
5. 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.