00752

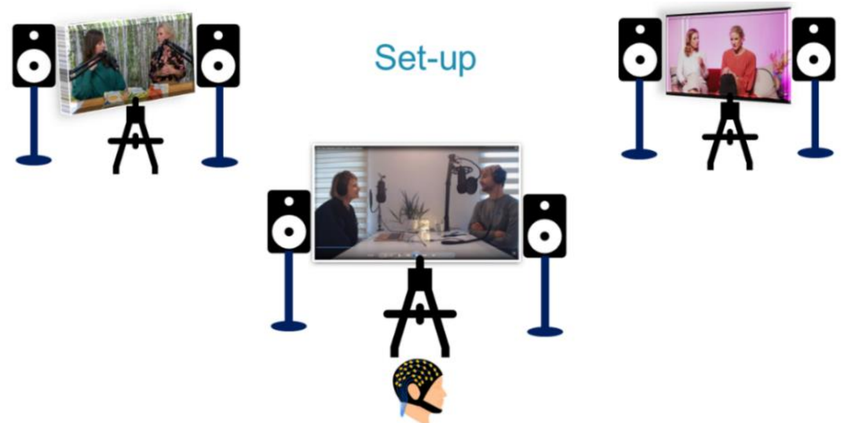
AUDITORY OBJECTIVE MEASURES

EEG-based Auditory Attention Decoding for neuro-steered hearing aids: decoding conversations in a multi-talker restaurant environment

Iris Van de Ryck, Nicolas Heintz, Iustina Rotaru, Simon Geirnaert, Alexander Bertrand, Tom Francart

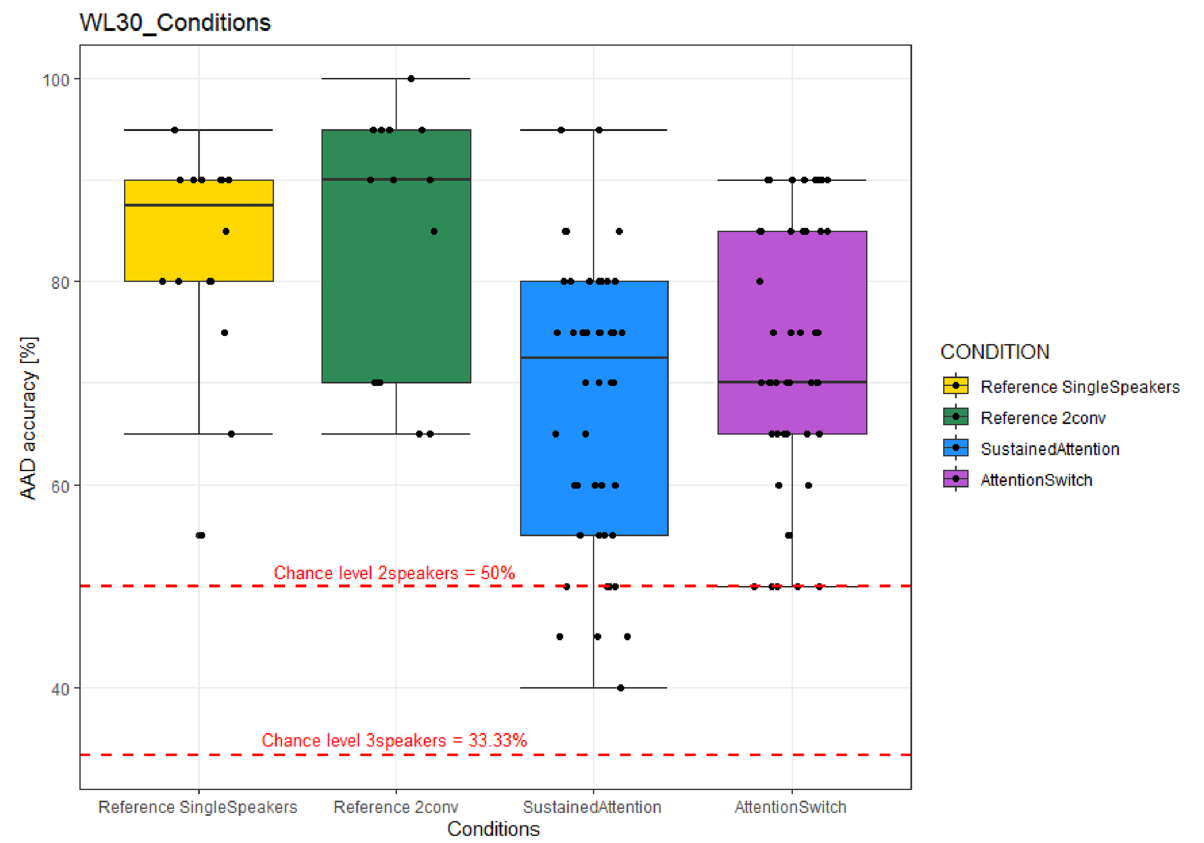
Abstract

Individuals with hearing impairment, even with the use of hearing aids, face difficulties understanding speech in multi-talker environments. Hearing aids and cochlear implants partly solve this problem through signal-processing algorithms that enhance the target speaker and suppress other noise sources. However, a fundamental problem appears in a multi-talker scenario: how does an algorithm decide which speaker the listener is actually attending to? This can be solved by decoding the auditory attention from measured brain signals (using EEG), and accordingly enhancing the attended speaker. This method is known **as auditory attention decoding (AAD)**. Previous studies have concentrated on scenarios with two speakers in which participants were required to focus on the attended speaker while ignoring the unattended speaker. However, in real-life situations, more than two speakers frequently talk at the same time (**cocktail party scenario**). This study serves as a means to comprehend the potential of AAD in real-life scenarios with numerous speakers. Therefore, we shift towards a new paradigm in which we decode clusters of turn-taking speakers instead of individual speakers.



Results

AAD accuracies are above chance levels, above 50% for the 2 reference conditions and above 33,33% for the sustained and attention switch conditions. AAD accuracies can reach 100% in the reference conditions with a median of 87% (ref. single speakers) and 90% (ref. 2 conversations). For the conditions with sustained attention (3 conversations) medians are around 70 to 75% and for the attention switch conditions, they are around 72%. We did not find a significant difference between decoding 2 speakers compared to 2 conversations. There is also no significant difference between sustained attention and attention switch conditions. In addition, there is no significant difference between the locations of attention (trials). This could mean that conversations can be decoded just like single speakers, even in this complex scenario.



Objectives

In this study we wanted to investigate whether we can still decode AAD from brain signals in such an acoustically challenging environment.

- Does conversations-based AAD work as good as 'classic' 2-speaker AAD?
- Does an attention switch influence the AAD performance?
- Does the locus of attention affect AAD performance?

	Condition	Content	Focus
1	Reference Single Speakers	Classic 2 speaker AAD	Left or Right
2	Reference 2 conversations	2 conversations instead of 2 speakers	Left or Right
3	Sustained Attention	3 conversations – attention only to 1 of the 3 conversations (separate trials)	Left or Front of Right
4	Attention Switch	3 conversations – switch attention from one conversation to another conversation	Left or Front of Right

Conclusion

Twenty young normal-hearing subjects were tested in a soundproof booth. We measured EEG during 4 different conditions. Conditions included 2 or 3 speakers or conversations. Results so far indicate promising AAD results for conversations. The AAD algorithm is able to perform in an acoustically complex scenario. This is one step closer towards neuro-steered hearing aids.

Methods and Materials

We simulated a restaurant scenario with 6 speakers having 3 simultaneous conversations (each with 2 turn-taking speakers), using 3 screens and 6 loudspeakers in a soundproof test booth and conducted an EEG experiment to investigate the AAD performance. Conditions were 10 minutes long, in which subjects focused their attention on 1 conversation and ignored the other speakers. The conversations were podcasts of 2 speakers that we selected. We compared conditions with sustained attention with conditions with 1 attention switch after 5 minutes. Additionally, there were 2 reference conditions, one with 2 single speakers and another condition with 2 conversations instead of 3. The attended conversation can be in the frontal position, left or right (trials). Brain activity was measured with 64-channel EEG.

References

- Choudhari, V., Han, C., Bickel, S., Mehta, A. D., Schevon, C., McKhann, G. M., & Mesgarani, N. (2024). Brain-controlled augmented hearing for spatially moving conversations in multi-talker environments. https://doi.org/https://doi.org/10.1101/2024.02.05.579018
- Das, N., Bertrand, A., & Francart, T. (2018). EEG-based auditory attention detection: Boundary conditions for background noise and speaker positions. *Journal of Neural Engineering*, *15*(6). https://doi.org/10.1088/1741-2552/aae0a6
- O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., Shamma, S. A., & Lalor, E. C. (2015). Attentional Selection in a Cocktail Party Environment Can Be Decoded from Single-Trial EEG. *Cerebral Cortex*, *25*(7), 1697–1706. https://doi.org/10.1093/cercor/bht355
- Van Eyndhoven, S., Francart, T., & Bertrand, A. (2017). EEG-Informed Attended Speaker Extraction from Recorded Speech Mixtures with Application in Neuro-Steered Hearing Prostheses. *IEEE Transactions on Biomedical Engineering*, *64*(5), 1045–1056. https://doi.org/10.1109/TBME.2016.2587382