Decoding attended talker solely from listening-state EEG signals

Mohamed Elminshawi[*]*, Julia Kostina, Emanuël A. P Habets, Neeraj Kumar Sharma*

International Audio Laboratories Erlangen, Germany
(A joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg and Fraunhofer IIS, Germany)

[*]mohamed.elminshawi@audiolabs-erlangen.de

Abstract (max 300 words)

Evolutionarily, the human brain specializes in efficient reception, processing, and interpretation of sensory signals. Focusing on aspects of human listening, the brain extracts a multitude of information from an attended audio signal. The problem of decoding the attended audio source attributes while listening to an audio signal is commonly referred to as auditory attention decoding (AAD). Conventional auditory attention decoding (AAD) approaches require access to both EEG and individual audio source signals in order to decode the attended audio source attributes (O' Sullivan et al., 2015). In this work, we propose an alternate approach to AAD which is based on classifying attended audio source attributes using only the EEG signal. This enlarges the scope of AAD as such an approach does not require access to the individual audio source signals. In the study, we analyze two AAD tasks, namely, attended speaker decoding and attended direction decoding. We hypothesize that attended speaker and location signatures can be decoded directly from short-time EEG signals. We validate this hypothesis using two publicly available listening-state AAD datasets (Fuglsang et al., 2020, and Mundanad et al., 2021). We find that both speaker and spatial attention can be decoded with significantly above chance performance from EEG. This holds even when using short (0.1 - 0.5 sec) decision windows. We also find that in comparison to hearing impaired subjects, normal hearing subjects perform better on the speaker detection

task. Furthermore, the duration as well as temporal distribution of the test data is found to significantly impact the classification performance. Our findings contribute towards furthering the understanding of attended speaker decoding from EEG signals, and open new questions on what aspects of the speech influence this decoding.

[Abstract Word count: 279]

References

Journal article

O'Sullivan, James A., Alan J. Power, Nima Mesgarani, Siddharth Rajaram, John J. Foxe, Barbara G. Shinn-Cunningham, Malcolm Slaney, Shihab A. Shamma, and Edmund C. Lalor. "Attentional selection in a cocktail party environment can be decoded from single-trial EEG." Cerebral cortex 25, no. 7 (2015): 1697-1706.

S. A. Fuglsang, J. Märcher-Rørsted, T. Dau, and J. Hjortkjær, 'Effects of Sensorineural Hearing Loss on Cortical Synchronization to Competing Speech during Selective Attention', J. Neurosci., vol. 40, no. 12, pp. 2562–2572, Mar. 2020, doi: 10.1523/JNEUROSCI.1936-19.2020.

Narayanan, Abhijith Mundanad, Rob Zink, and Alexander Bertrand. "EEG miniaturization limits for stimulus decoding with EEG sensor networks." Journal of Neural Engineering 18, no. 5 (2021): 056042.