 DOI: 10.3217/978-3-85125-467-9-37

Speech envelope tracking using around-the-ear EEG

B. Mirkovic[1,2*], M. G. Bleichner[1,2], S. Debener[1,2], M. De Vos[3]

[1]Department of Psychology, Neuropsychology Lab, University of Oldenburg, Oldenburg, Germany;
[2]Cluster of Excellence Hearing4all, Oldenburg, Germany;
[3] Institute of Biomedical Engineering, Department of Engineering, University of Oxford, UK;

* E-mail: bojana.mirkovic@uni-oldenburg.de

Introduction: Envelope tracking can be used to predict the direction of attention to natural speech based on continuous electroencephalography (EEG) [1] and is therefore a useful approach to control hearing devices with neural signals. In our previous work, we showed that this method also works with a small number of electrodes with a minimal decrease in decoding accuracy [2]. Here we compare the quality of speech decoding using around-the-ear EEG electrodes (cEEGrid) [3] connected to a mobile EEG amplifier and high density EEG system and propose an unobtrusive Brain-Computer Interface (BCI) based on continuous EEG signals.

Material and Methods: 18 participants with self-reported normal hearing took part in this study. Participants were presented with two 50-minute long concurrent continuous speech streams and instructed to attend to only one of them. After each ten minutes of experiment, participants answered 10 multiple-choice questions pertaining to the attended story. The EEG was recorded simultaneously with cEEGrid and high-density 84-channel EEG cap (Fig.1A). Decoding accuracy was calculated following the procedure described in [1-2], with 60s training and validation trial intervals. The temporal profile of the neural response to presented speech, temporal response function (TRF), was obtained using a univariate mapping between the stimulus and corresponding EEG [4] in the -200ms to 800ms time lag range.

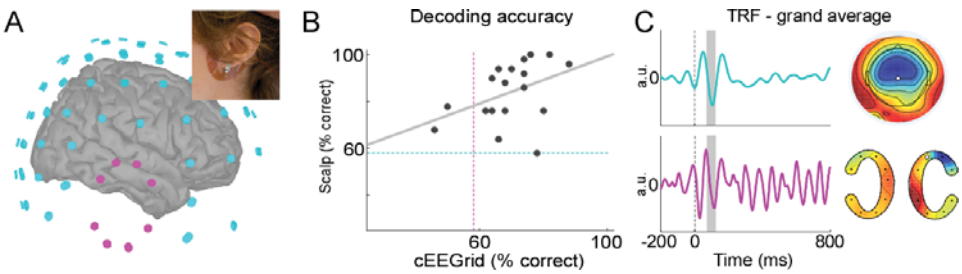

Figure 1. *A) cEEGrid (in the right corner) is a semi-disposable C shaped electrode grid printed on flexible biocompatible material [3]. Recordings were acquired from high density 84 channel EEG (blue) and 18 channel cEEGrid (pink) in parallel. B) Decoding accuracies for scalp and cEEGrid weakly correlate (r=0.43,p=0.07). Dashed lines represent chance level. C) On the left, temporal response functions (TRFs) for scalp Cz and cEEGrid R3 channel are shown. Corresponding channels are indicated in white on the scalp and cEEGgrid topography, shown on the right. The topographies are based on TRF profiles for all scalp and cEEGrid channels at ~100ms time lag (grey shaded area) and exhibit the pattern of N1-P2 AEP component.*

Results: On average, both scalp and cEEGrid decoding accuracy were significantly above the statistical chance level of 59% (cEEGrids 69.4%, scalp EEG 84.5%), as shown on Fig 1B. Only two individuals performed below chance for the cEEGrid. All individuals had above chance questionnaire performance (68% to 100%, mean of 85.7%), validating the attention instruction. No significant correlation emerged between questionnaire performance and decoding accuracy. Both scalp and cEEGrid TRF revealed a temporal and spatial pattern with components also visible in the AEP (Fig. 1C).

Discussion: Electrode placement around the ear is sufficient to indicate the attended speaker in this concurrent speaker paradigm. The usefulness of the cEEGrid data is further confirmed by investigating the physiologically interpretable TRF pattern, which, while showing the negative peak at 100ms for both setups, contains clearer responses for scalp recordings. The decoding accuracy of the cEEGrid recordings is lower compared to the scalp recordings, which might be due to the lower number of electrodes and the lower spatial coverage.

Significance: We show that tracking the auditory attention to continuous speech can be done with an unobtrusive, user-friendly electrode setup located around the ear. Combined with a lightweight mobile EEG amplifier this result is a further step towards controlling hearing devices with a BCI.

References:

[1] 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. Attentional selection in a cocktail party environment can be decoded from single-trial EEG. Cereb. Cortex. 2014; doi:10.1093/cercor/bht355

[2] Mirkovic B, Debener S, Jaeger M, De Vos M. Decoding the attended speech stream with multi-channel EEG: implications for online, daily-life applications. Journal of neural engineering. 2015 Jun 2;12(4):046007.

[3] Debener S, Emkes R, De Vos M, Bleichner M. Unobtrusive ambulatory EEG using a smartphone and flexible printed electrodes around the ear. Sci Rep. 2015 Nov 17;5:16743.

[4] Crosse M J, Butler J S, and Lalor E C. Congruent visual speech enhances cortical entrainment to continuous auditory speech in noise-free conditions. The Journal of Neuroscience 35.42 (2015): 14195-14204.