Decoding Selective Attention in Normal Hearing Listeners and Bilateral Cochlear Implant Users with Concealed Ear EEG

W. Nogueira¹, H. Dolhopiatenko¹, I. Schierholz¹, B. Mirkovich², M. Bleichner², S. Debener², A. Büchner¹

¹ Hearing4all, Dept. of Otolaryngology, Hannover Medical University, Hannover, Germany
² Hearing4all, University of Oldenburg, Oldenburg, Germany

18.07.2019
Lake Tahoe
USA
Motivation: Cocktail Party Effect

Speech Streams

Cherry, 1953, JASA
Monaural Source Separation [Huang et al. 2014; Goehring et al. 2017; Nogueira & Gajecki 2017]
Methods:
Monaural Source Separation

- **Deep Recurrent Neural Network** [Huang et al. 2014; Goehring et al. 2017; Nogueira & Gajecki 2017]
- **Deep Convolutional Autoencoder** [Gajecki & Nogueira, 2018, JASA]
Methods
Front-End Source Separation Algorithm
Methods

DRNN: Deep Recurrent Neural Network

CI Users

Speech in Speech Task:

Target: Male Speaker (HSM)
Interference: Female Speaker

DRNN1: Complex DRNN
  - 3 Layers
  - 1024 nodes

DRNN2: Simple DRNN
  - 1 Layer
  - 256 nodes
Which is the attended speaker: Male or Female Speaker?
Motivation

- It has been shown that EEG recordings can be used to identify the attended speech stream in a multi-speaker scenario (cocktail party) in normal hearing (NH) listeners (e.g. O‘Soulivan, 2014; Mirkovich, 2015)

- The attended and unattended speech streams are differentially represented in the cortical activity (Mesgarani and Cheng, 2012)
  - Cortical response to unattended speech being suppressed and the response to attended speech being amplified.

- Goal: Investigate whether it is possible to identify the attended speech stream in a multi-speaker scenario in cochlear implant (CI) users
▪ Large artifact introduced by the CI
  ▪ Artifact is related to the envelope of the audio signal

▪ In CIs the cortical activity may be more smeared
Methods: Decoding Selective Attention Temporal Response Function (TRF)

Adapted from O’Soulivan, 2014

Nogueira et al 2019, IEEE Transactions Biomedical Engineering
Methods: Detection Selective Attention Temporal Response Function (TRF)

Decoder

\[ W_\Delta \]

Delay/Lag

\[ Z^{-\Delta} \]

From: http://drmridha.com/services/eeg

\[ C_{x_a \hat{x}_a} = \frac{E\{ (x_a[k] - \mu_{x_a[k]}) (\hat{x}_a[k] - \mu_{\hat{x}_a[k]}) \}}{\sigma_{x_a[k]} \sigma_{\hat{x}_a[k]}} \]

\[ C_{x_u \hat{x}_a} = \frac{E\{ (x_u[k] - \mu_{x_u[k]}) (\hat{x}_a[k] - \mu_{\hat{x}_a[k]}) \}}{\sigma_{x_u[k]} \sigma_{\hat{x}_a[k]}} \]

If \( C_{x_a \hat{x}_a} > C_{x_u \hat{x}_a} \) \rightarrow \text{Attended signal decoded}
High-Density Scalp Electrodes + cEEGrid electrode array

Collaboration Prof. S. Debener, B. Mirkovich and M. Bleichner, University of Oldenburg
Selective Attention Correlation Coef.
High Density
N=10 NH listeners
Recording 48 minutes
Selective Attention Correlation Coef.
cEEGgrid
N=10 NH listeners
Recording 48 minutes

Correlation Coefficients

NH cEEGrid

Lags (Δ)

Attended
Unattended
Selective Attention Correlation Coef.
High density EEG
N=10 NH listeners
Recording 48 minutes (CI)
Selective Attention Accuracy

cEEGgrid

N=10 NH listeners

Recording 48 minutes (CI)

Correlation Coefficients

CI cEEGrid

Correlation Coefficients

Lags(Δ) [ms]
Selective Attention Accuracy
cEEGgrid vs Scalp Electrodes
Accuracy [%]: Based on 48 Trials/subject

Artifact Analysis

Nogueira et al. 2019, Frontiers Neuroscience.
Summary/Discussion

- Signal pre-processing (based on deep learning) may help in cocktail party type scenarios .... But which is the attended speaker?
- Decoding selective attention may be used to steer signal processing algorithms towards the attended speaker
- Good accuracy detecting attended speech stream in normal hearing subjects with high density EEG
- Slightly lower accuracy in CI subjects with high density EEG
  - No artifact removal
- Worse accuracy with concealed around the ear EEG
- Next steps: Selective attention accuracy as a measure of speech performance?
Thank you
Cognitive-Driven Binaural Speech Enhancement System for Hearing Aid Applications, (Aroudi, Doclo)
- Individualized Electrical Stimulation Patterns with Auditory Prostheses and Closed-Loop Systems (A. Bahmer)
- From Surgery to Sound Perception - Signal Processing in Cochlear Implants (Koning, Litvak, Hamacher)
- Future Trends in Hearing Implants, (P. Nopp)
- Developing Behind-The-Ear EEG Sensing for Hearing Devices (M. Bleichner, S. Debener)
- Electrochemical Protocols Upgrade Conventional Noble Metal Electrodes to Long-Term Stable Sensors at the Tissue/Electrode Interface (A. Weltin)
- Next Generation Cochlear Implants Require Microsecond Binaural Synchronization, (N. Rosskothen-Kuhl)
- New Electrical and Ultrasound Stimulation Technologies for Treating Hearing Disorders and Tinnitus (H. Lim)
Speech Performance vs Correlation Coefficients

Max Correlation Coefficient

HSM sentences [%]

r = -0.12428

Difference of attended and unattended correlation coefficients

HSM sentences [%]

r = 0.037848
Position of electrodes

FIGURE 3 | Comparison of electrode locations. (A) Selective attention decoding accuracy using all electrodes (eye electrodes 29 and 30 and scalp EEG electrodes spared at the position of the cEEGrid are excluded from the analysis); (B) Selective attention decoding accuracy using 19 electrodes on the top of scalp; (C) Decoding accuracy using 14 electrodes in the area next to the ear.
Artefact Model

Electric Field Model

\[ V_e(x, y, z) = \frac{I}{4\pi\sigma\sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}}, \]

\( V_e(x, y, z) \): Voltage at position \( x, y, z \) caused by stimulation of CI electrode \( i \) with current \( I \)
Decoding Selective Attention with an Artifact Model

Correlation coefficients

Full artifact decoding accuracy
Speech reconstruction from the single-trial EEG data was obtained by training decoders using a regularized least square estimation method.

- Parameter Lag ($\Delta$) \( \rightarrow \hat{x}_{a,u}[k] = \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} w_{n,l} y_n[k + \Delta + l] \)

- \( J_{LS}(\mathbf{W}_a) = \mathbb{E}\{|\mathbf{x}_a[k] - \mathbf{W}_a^T \mathbf{Y}[k]|^2\} \rightarrow \text{minimize} \)

- \( K: \) time; \( n: \) electrode; \( y_n: \) recorded EEG; \( w_{n,l}: \) decoder; \( \hat{x}_{a,u}: \) reconstructed signal

EEG recording matrix with delayed versions (Lags)

\[
Y = \begin{bmatrix}
    r_1(0) & r_1(1) & \cdots & r_1(\tau_{\max}) & \cdots & r_1(T) \\
    \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\
    0 & 0 & \cdots & r_1(0) & \cdots & r_1(T - \tau_{\max}) \\
    \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\
    r_n(0) & r_n(1) & \cdots & r_n(\tau_{\max}) & \cdots & r_n(T) \\
    \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\
    0 & 0 & \cdots & r_n(\tau_{\max}) & \cdots & r_n(T - \tau_{\max})
\end{bmatrix}
\]
Lags, typically expressed in [ms] account for delay between EEG and input sound.

\[
Y = \begin{bmatrix}
  r_1(0) & r_1(1) & \cdots & r_1(\tau_{\text{max}}) & \cdots & r_1(T) \\
  \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & r_1(0) & \cdots & r_1(T - \tau_{\text{max}}) \\
  \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
  r_n(0) & r_n(1) & \cdots & r_n(\tau_{\text{max}}) & \cdots & r_n(T) \\
  \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
  0 & 0 & \cdots & r_n(\tau_{\text{max}}) & \cdots & r_n(T - \tau_{\text{max}})
\end{bmatrix}
\]
Study 1: Investigate whether the degradation of spectral resolution as produced by a CI impacts the classification of selective attention
Methods:
Study 1: Procedure

- N=12 normal hearing subjects
- Recording 24 minutes using Vocoder (23 min training, 1 min testing + CV)
- Recording 24 minutes using Original (23 min training, 1 min testing + CV)

- No significant difference between Original speech or Vocoded Speech
- Envelope is preserved

Grey area: Chance Level
Study 2: Investigate whether CI artifact impacts the classification of selective attention
N=12 CI users and 12 NH listeners

- Recording 48 minutes (CI), 24 minutes (NH)

- Decoding accuracy in CIs is above chance level
  → Less influence from the artifact than expected
Correlation Coefficient Analysis

(a) Attended Decoder

NH

Vocoder

CI


18.07.2019
Study 3: Investigate whether selective attention can be decoded in CI users using a mobile EEG system
Methods:
Study 3: Procedure

- N=12 CI users and 12 NH listeners
- Recording 48 minutes (CI), 24 minutes (NH)

Decoding accuracy in CIs is above chance level
→ Less influence from the artifact than expected
Study 3: cEEGrid electrode array
Debener et. al. 2015
Cochlear Implants
Blind Source Separation

Front-end signal processing to address the “cocktail party” scenario

Speech in Speech Task:

Target: Male Speaker (HSM)
Interference: Female Speaker
Selective Attention Accuracy
High Density Scalp Electrodes


![NH Scalp](chart)
![NH cEEGrid](chart)

![CI Scalp](chart)
![CI cEEGrid](chart)
Selective Attention Accuracy
cEEGgrid vs Scalp Electrodes

Nogueira et al. 2019, Frontiers Neuroscience.