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

W. Nogueira[1], H. Dolhopiatenko[1], I. Schierholz[1], B. Mirkovich[2], M. Bleichner[2], S. Debener[2] , A. Büchner[1]

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

18.07.2019
Lake Tahoe
USA

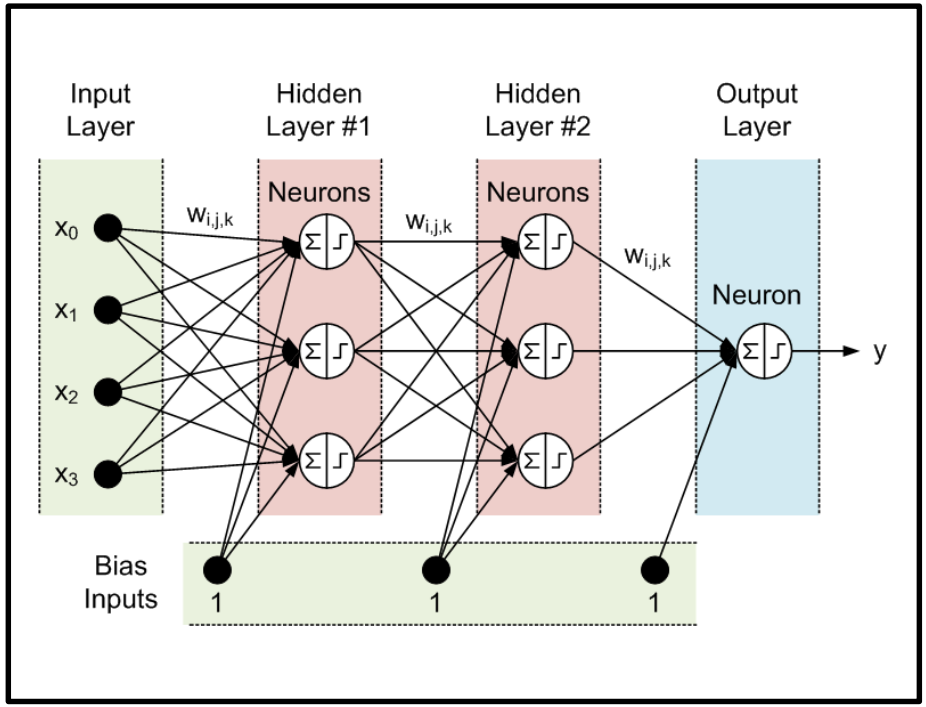
Motivation: Cocktail Party Effect

Speech
Streams

Cherry, 1953, JASA

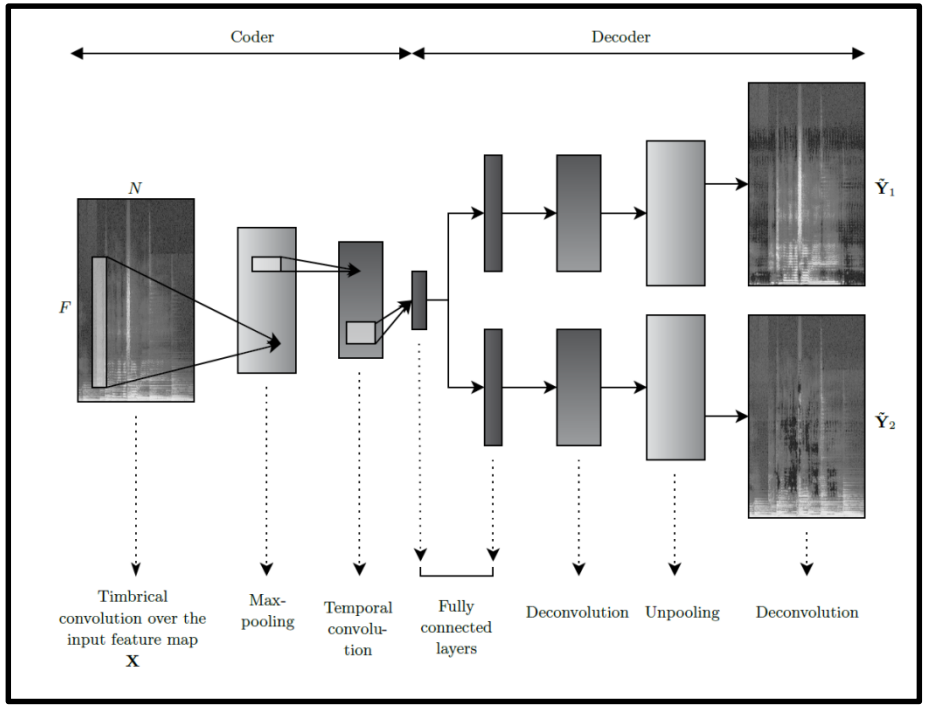
Methods:
Monaural Source Separation

- Monaural Source Separation [Huang et al. 2014;Goehring et al. 2017;Nogueira&Gajecki 2017]

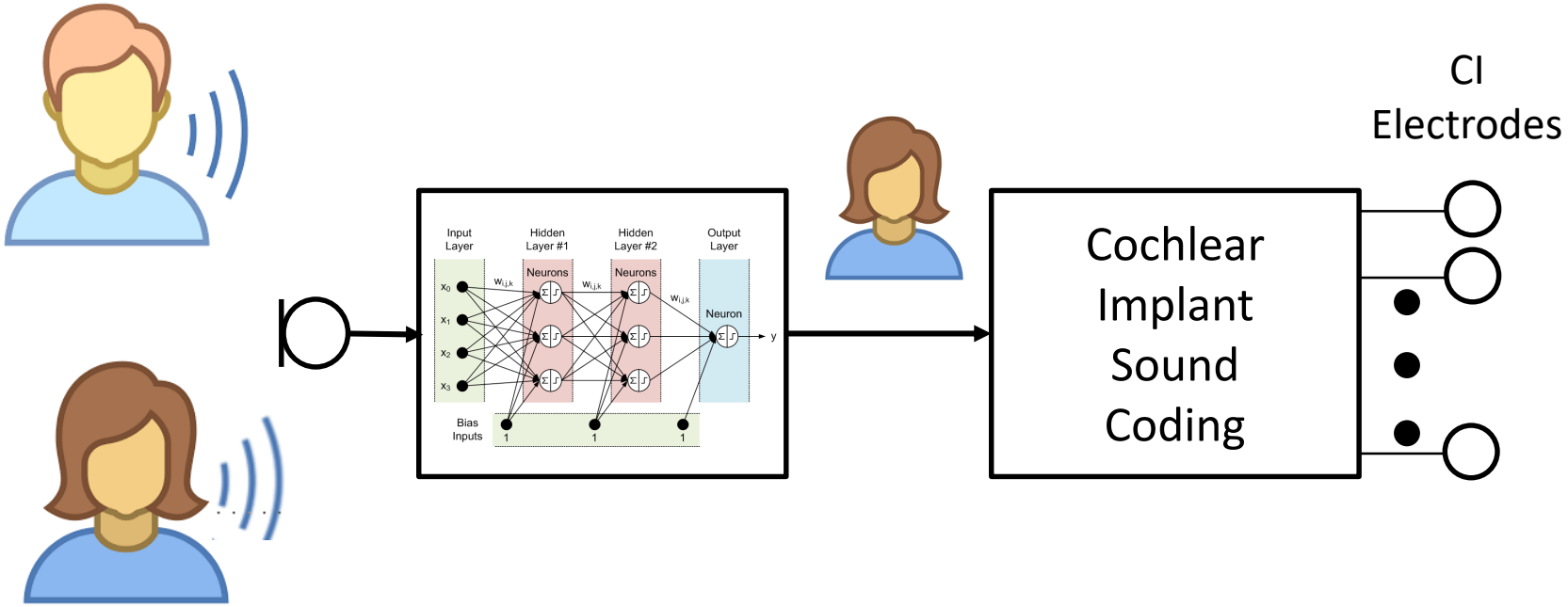


Methods:
Monaural Source Separation

- 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

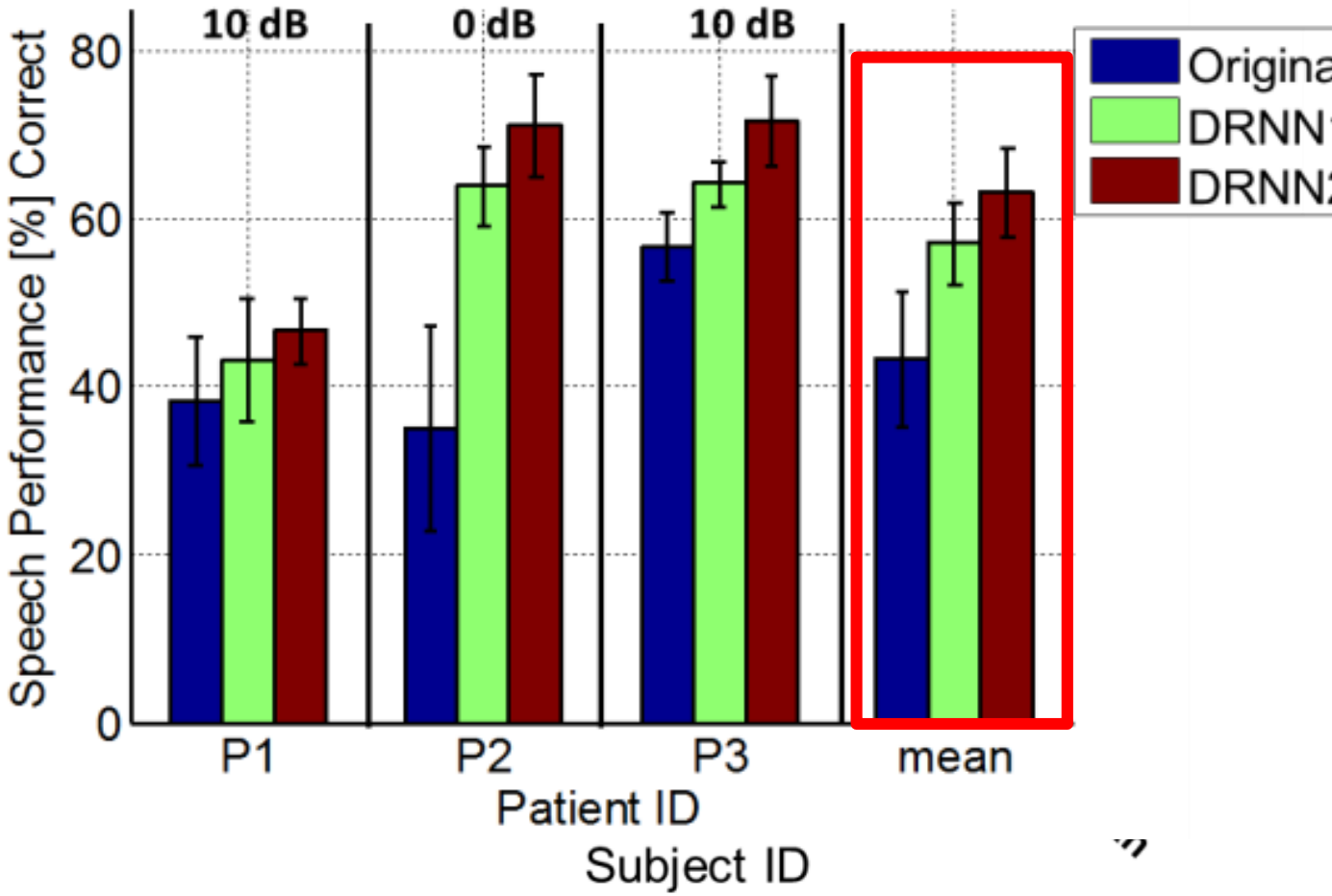
CI Users

Speech in Speech Task:

Target: Male Speaker (HSM)
Interference: Female Speaker

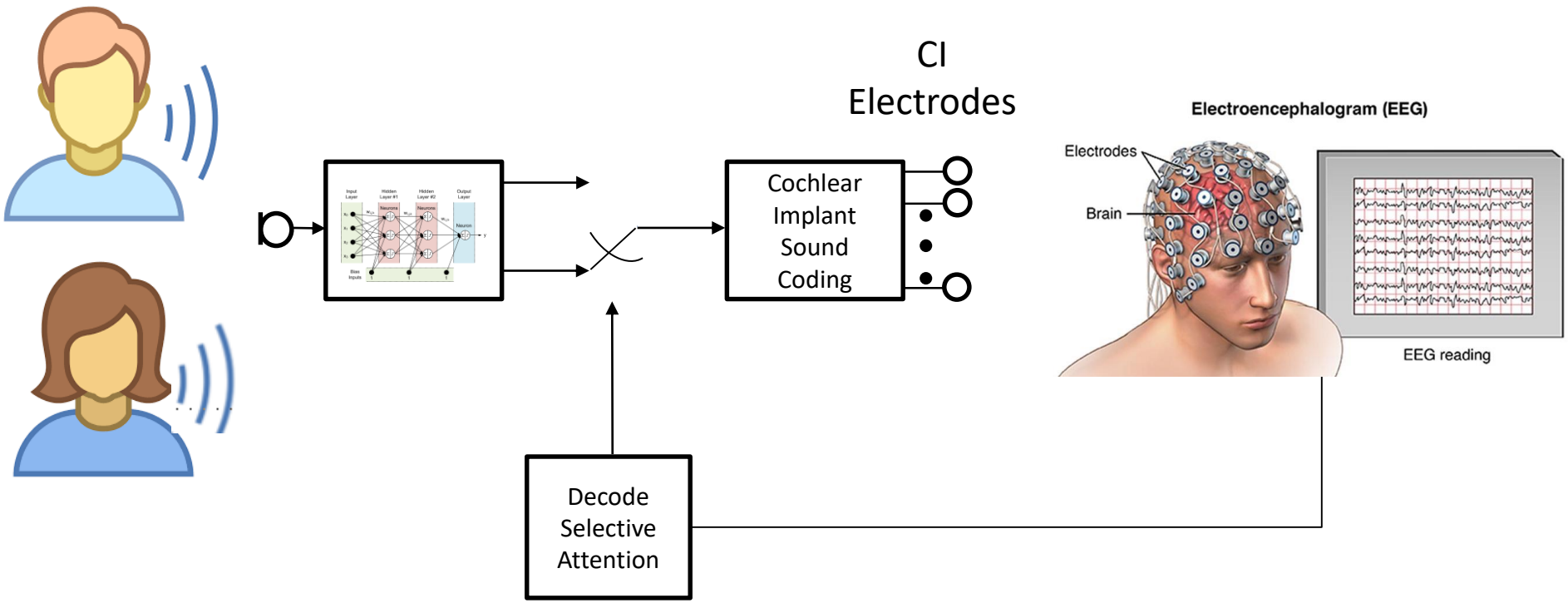
DRNN1: Complex DRNN
- 3 Layers
- 1024 nodes

DRNN2: Simple DRNN
- 1 Layer
- 256 nodes



Vision
Decode Selective Attention
Steer front end signal processing

- Which is the attended speaker: Male or Female Speaker?



Han, O'Sullivan, Luo, Herrero, Mehta, Mesgarani, 2019, Science

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.

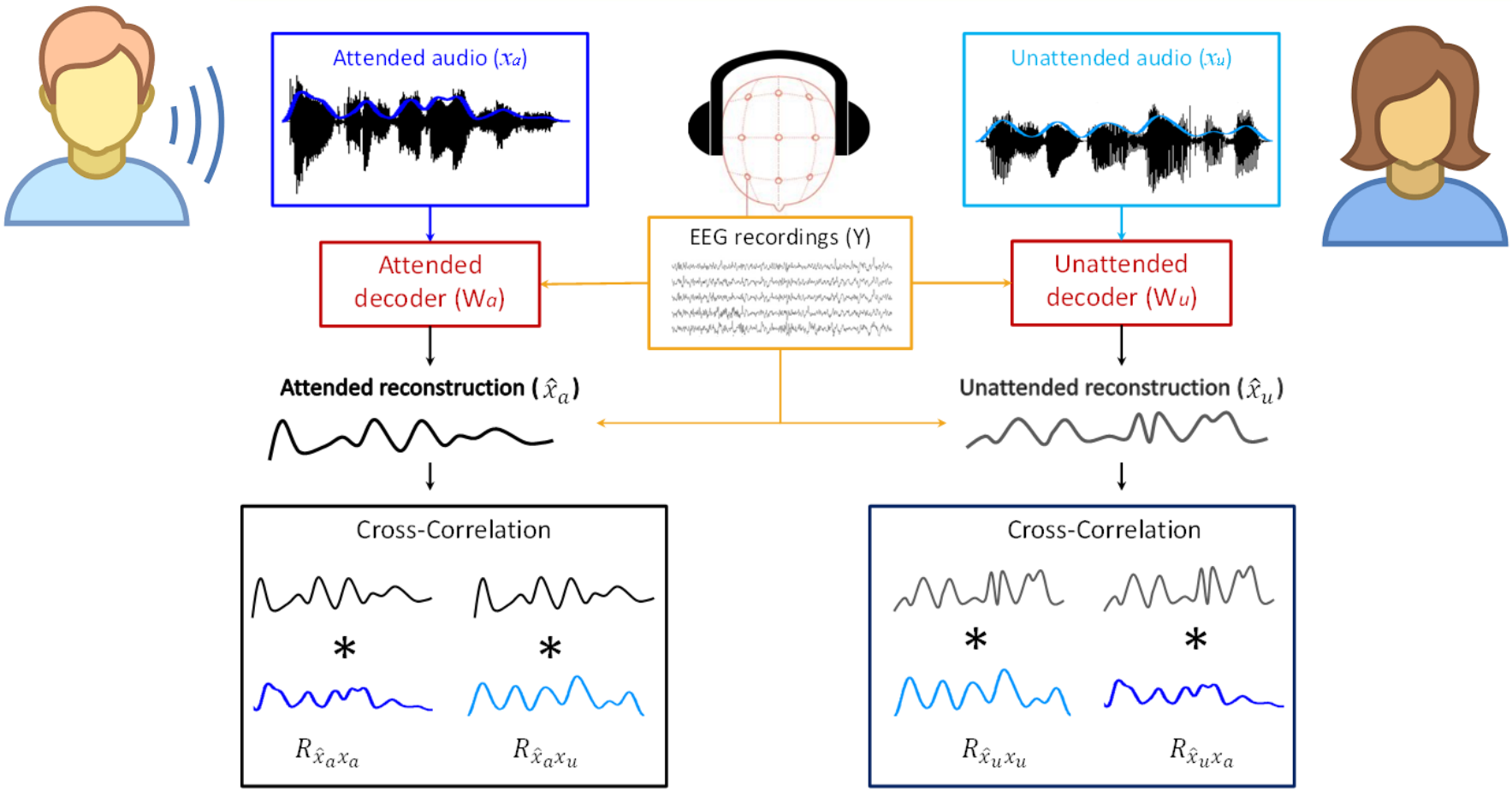
- Goal: Investigate whether it is possible to identify the attended speech stream in a multi-speaker scenario in cochlear implant (CI) users

Ongoing EEG recordings and Cochlear Implants

- 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

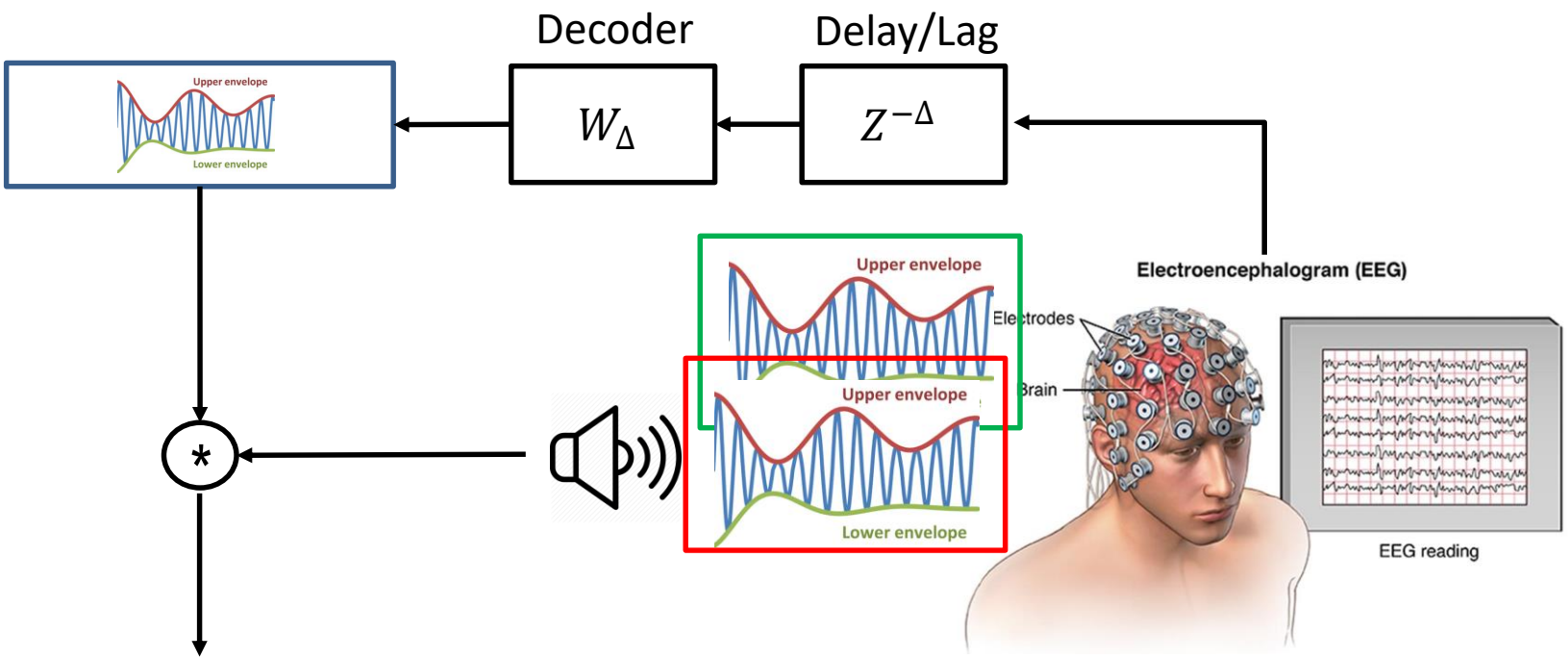
Methods: Decoding Selective Attention Temporal Response Function (TRF)



Adapted from O'Soulivan, 2014

Nogueira et al 2019, IEEE Transactions Biomedical Engineering

Methods: Detection Selective Attention Temporal Response Function (TRF)



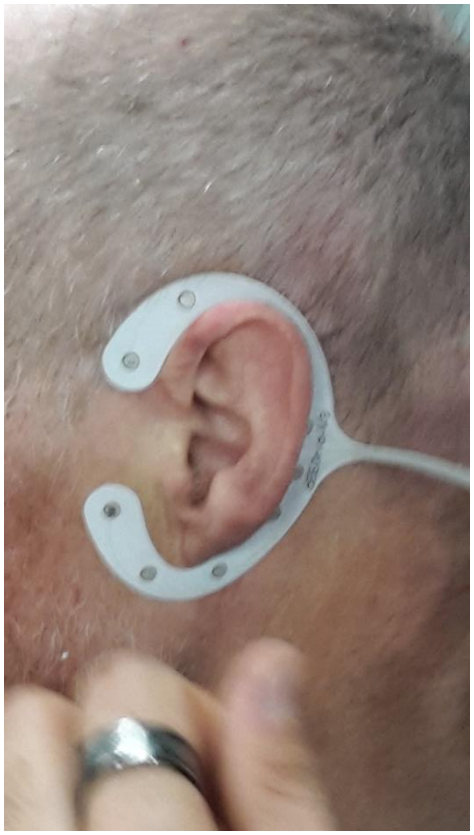
From:http://drmridha.com/services/eeg

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

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

If $C_{x_a\hat{x}_a} > C_{x_u\hat{x}_a}$ ➡ Attended signal decoded

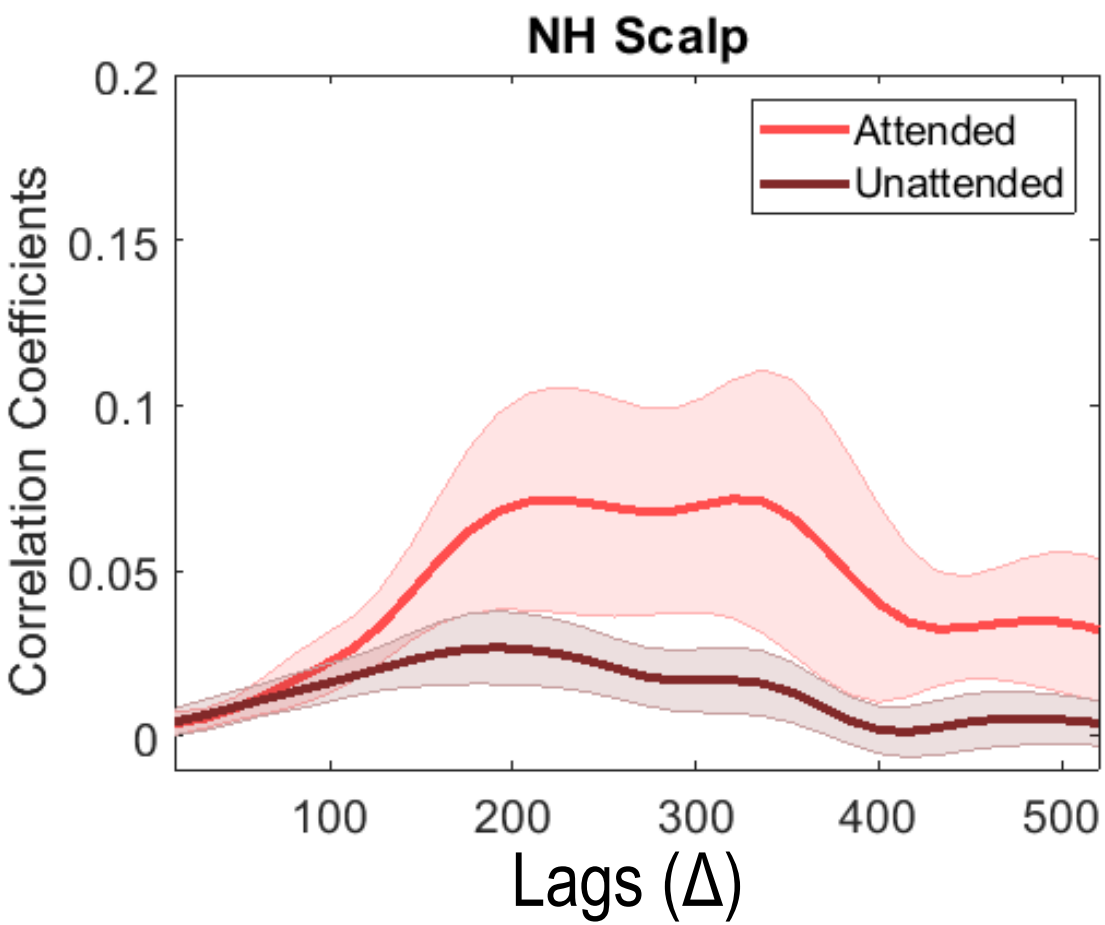
High-Density Scalp Electrodes + cEEGrid electrode array

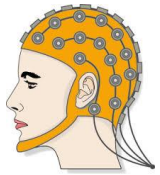


Collaboration Prof. S. Debener, B. Mirkovich and M. Bleichner, University of Oldenburg

Nogueira et al 2019, Frontiers Neuroscience.

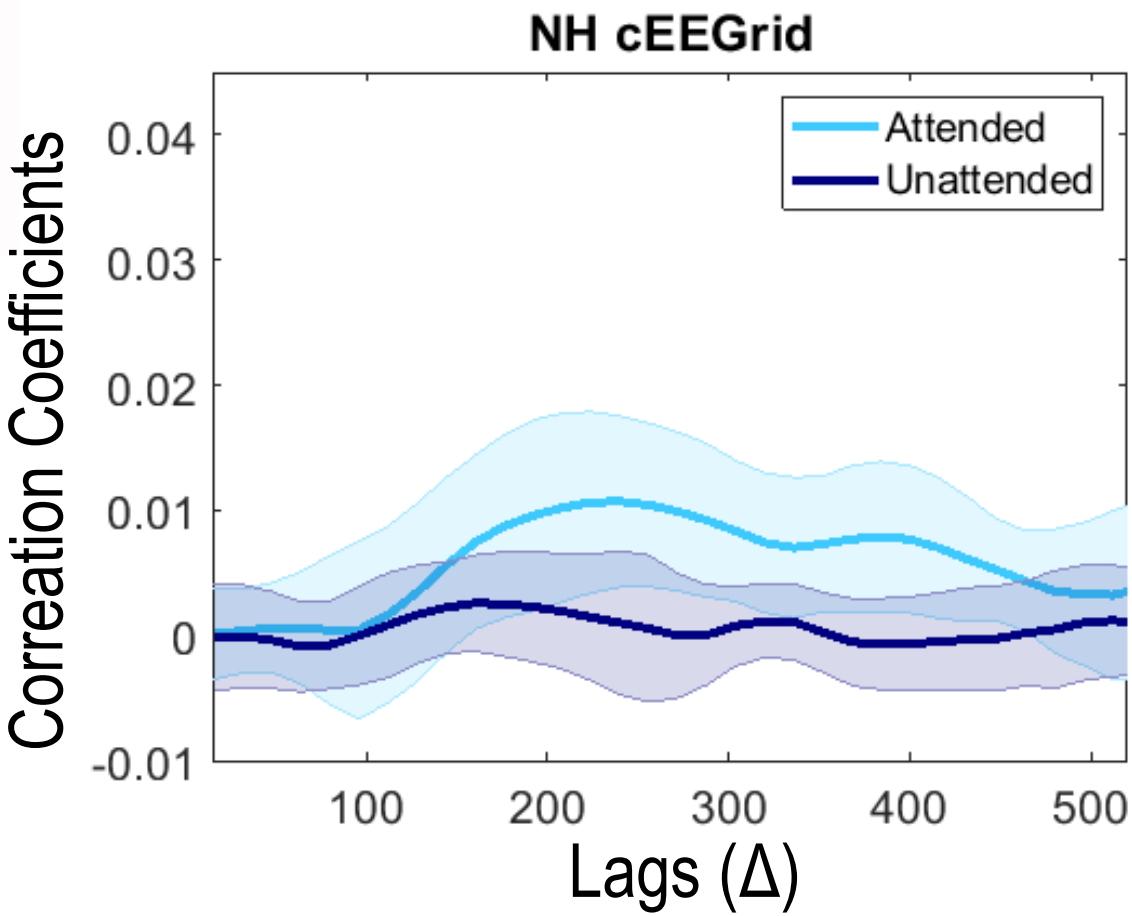
Selective Attention Correlation Coef.
High Density
N=10 NH listeners
Recording 48 minutes

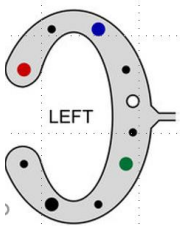




Nogueira et al 2019, Frontiers Neuroscience.

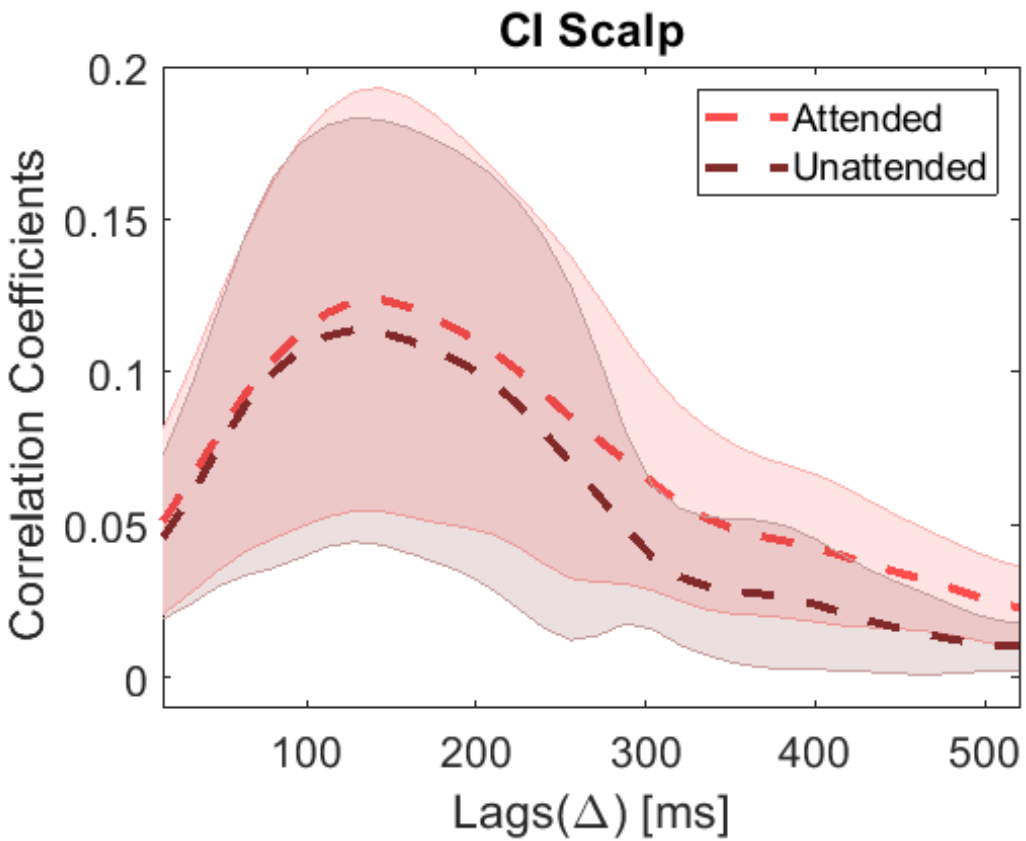
Selective Attention Correlation Coef.
cEEGgrid
N=10 NH listeners
Recording 48 minutes

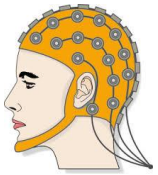




Nogueira et al 2019, Frontiers Neuroscience.

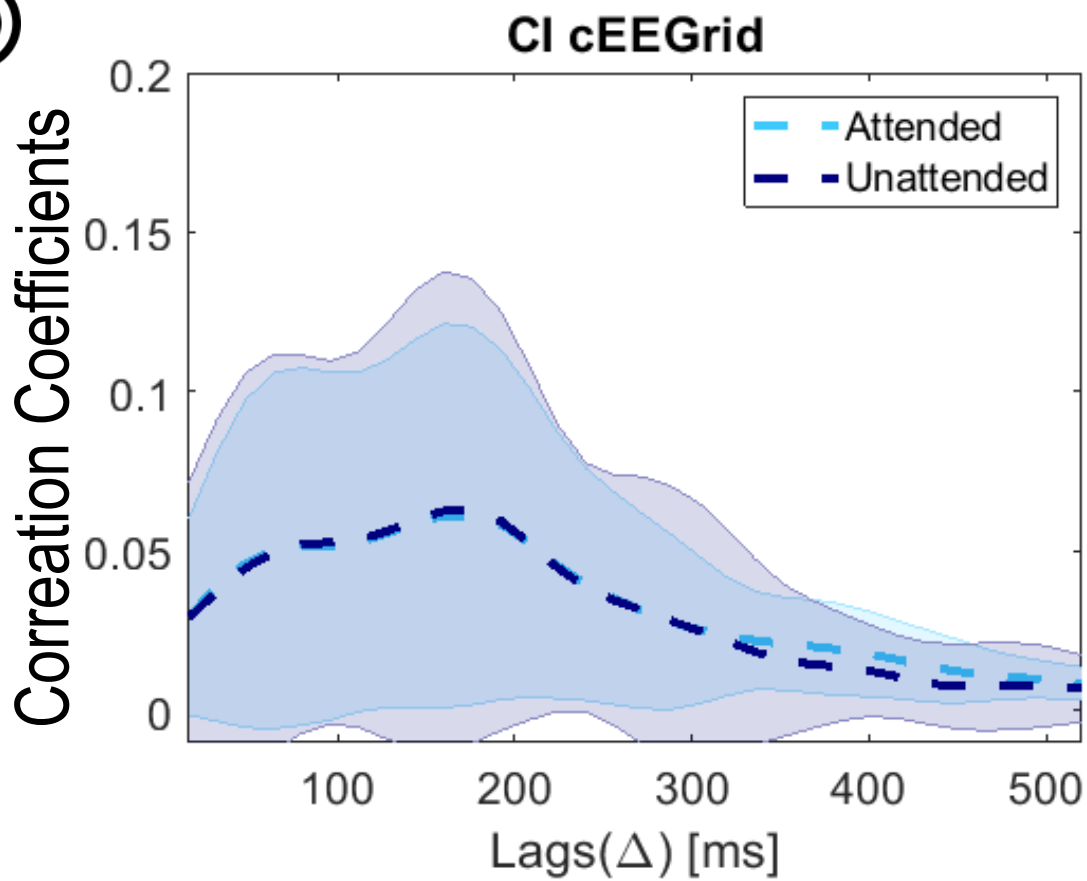
Selective Attention Correlation Coef.
High density EEG
N=10 NH listeners
Recording 48 minutes (CI)

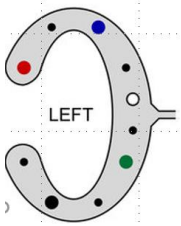




Nogueira et al 2019, Frontiers Neuroscience.

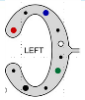
Selective Attention Accuracy
cEEGgrid
N=10 NH listeners
Recording 48 minutes (CI)





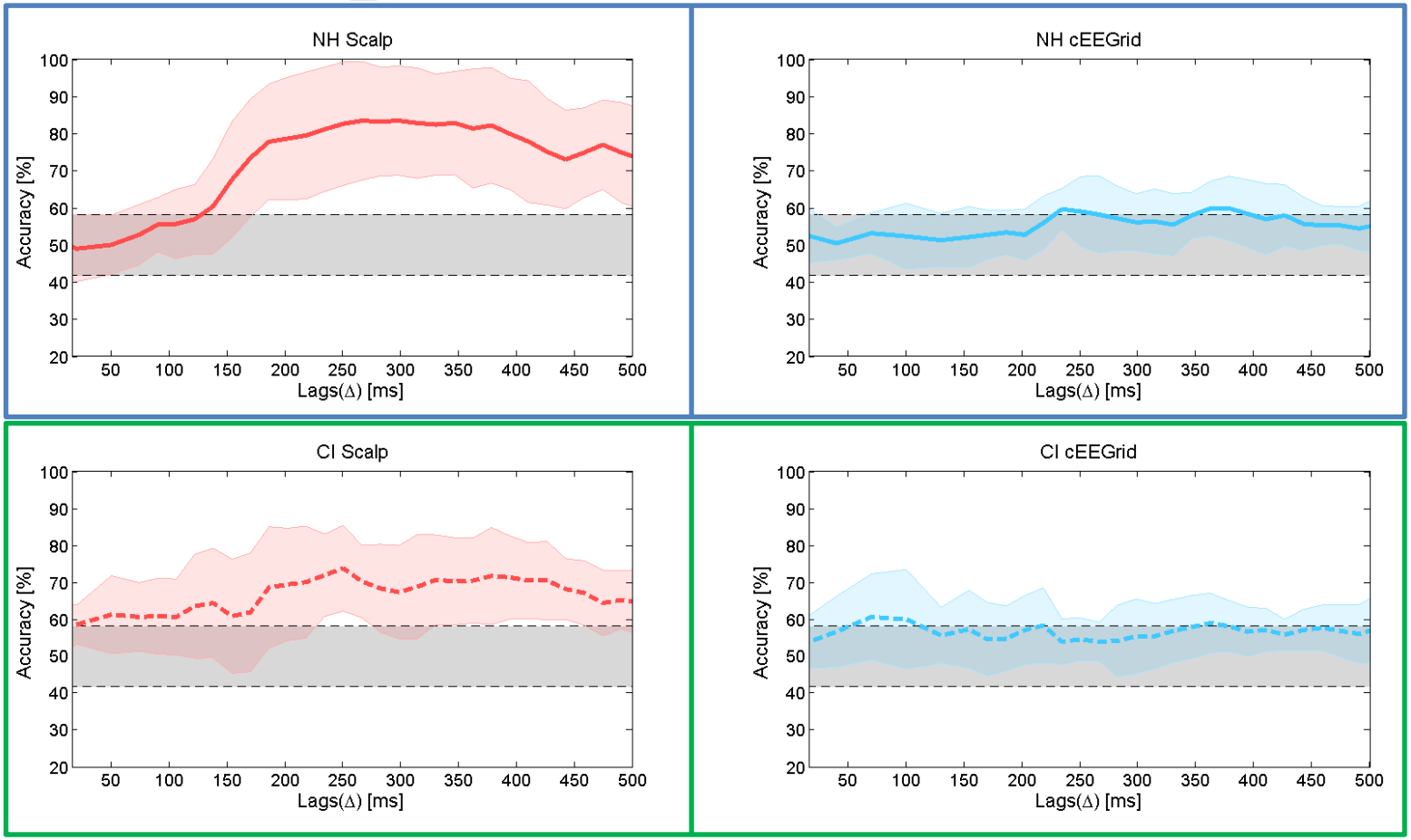
Nogueira et al 2019, Frontiers Neuroscience.

Selective Attention Accuracy cEEGgrid vs Scalp Electrodes Accuracy [%]: Based on 48 Trials/subject



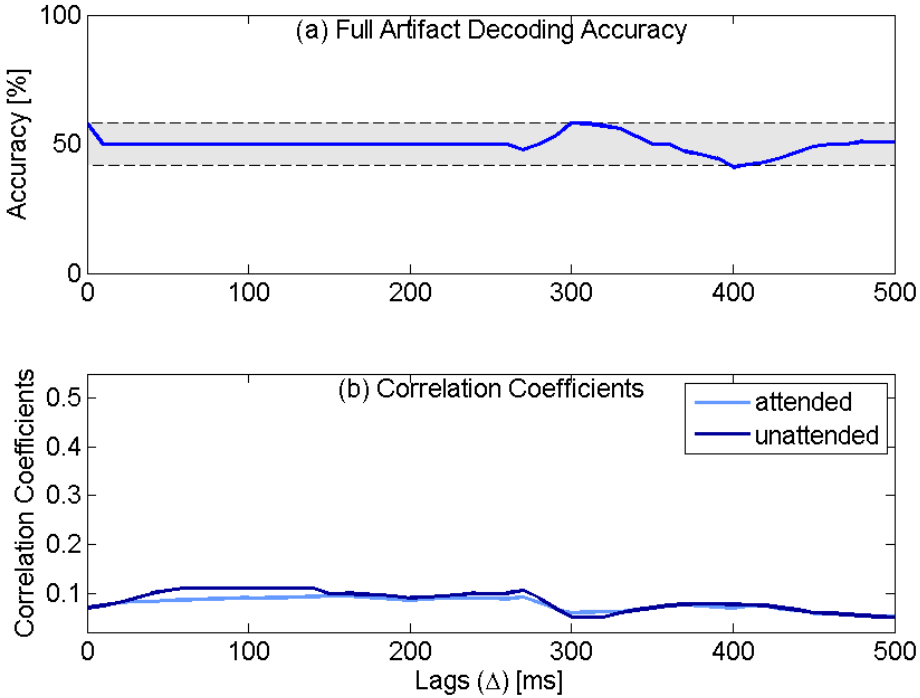
NH

CI



Nogueira et al 2019, Frontiers Neuroscience.

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?

Hearing 4all

EXZELLENZCLUSTER IM AUDITORY VALLEY ZUKUNFT HÖREN

APG Auditory Prosthetic Group

MHH Medizinische Hochschule Hannover

Irina Schierholz
EEG

Florian Langner
Loudness Model

Benjamin Krüger
EAS masking

Marina Imsiecke
EAS masking – Forward Masking

Tom Gajecki
Binaural Sound Processing

Collaborators:
Bojana Mirkovic (Oldenburg)
Martin Bleichner (Oldenburg)
Stephan Debener (Oldenbug)

Thomas Lenarz

Andreas Büchner

Thank you

DFG Deutsche Forschungsgemeinschaft

IEEE EMBC (July, 26 in Berlin) Hearing4all – Hearing Technologies and Diagnostics

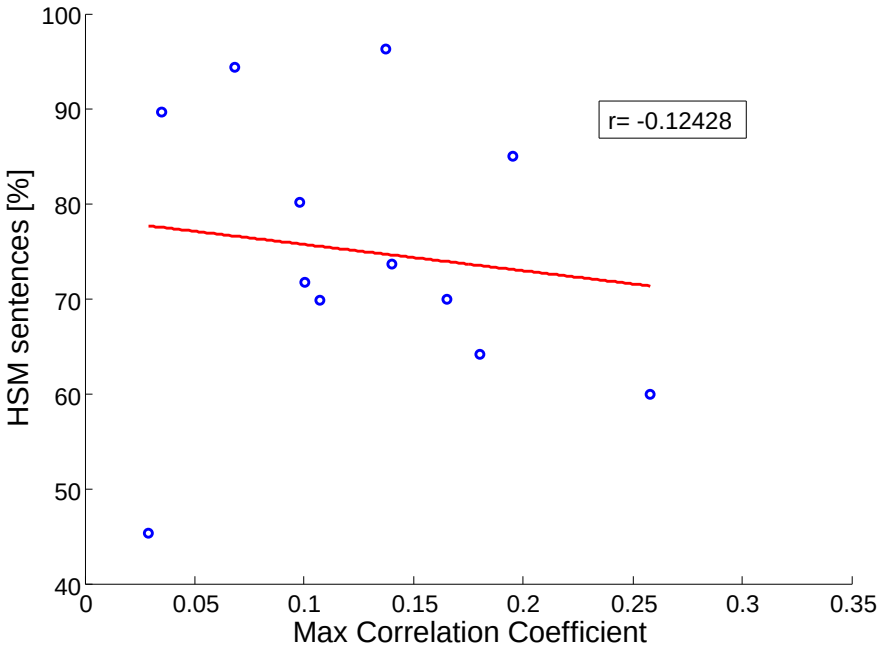
- Cognitive-Driven Binaural Speech Enhancement System for Hearing Aid Applications, (**Aroudi, Doclo**)
- Individualized Electrical Stimulation Patterns with Auditory Prostheses and Closed-Loop Systems (**A. Bahmer**)
- From Surgery to Sound Perception - Signal Processing in Cochlear Implants (**Koning, 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**)

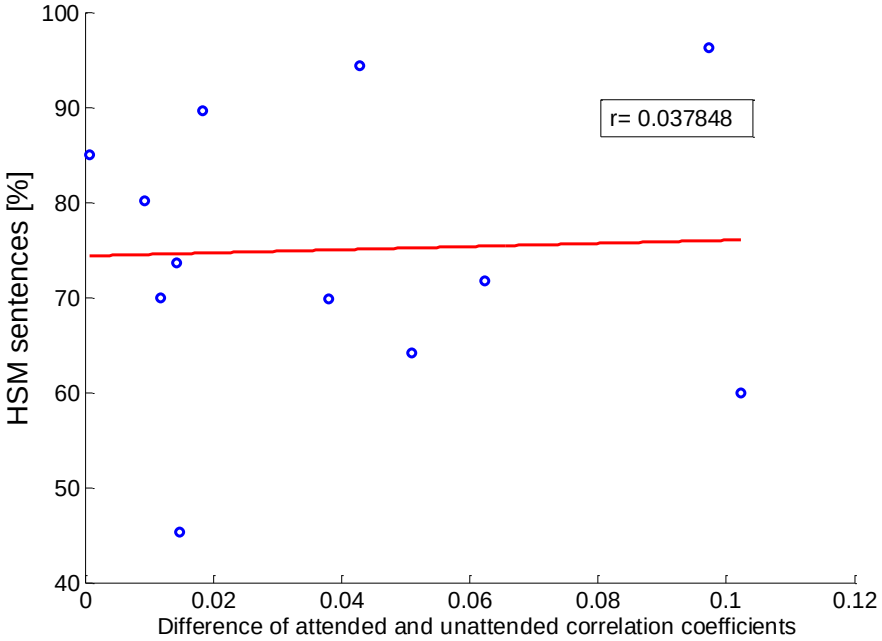
ENGINEERING in MEDICINE and BIOLOGY
23-27 July
41st EMBC CONFERENCE 2019 BERLIN

The APG organizes a mini-symposium during the IEEE EMBC conference in Berlin on July 26th.
If you are around Berlin, don't doubt to take part of this event. You can get daily tickets for the IEEE EMBC.

The mini-symposium is sponsored by the Hearing4all Cluster of Excellence.

Speech Performance vs Correlation Coefficients





Position of electrodes

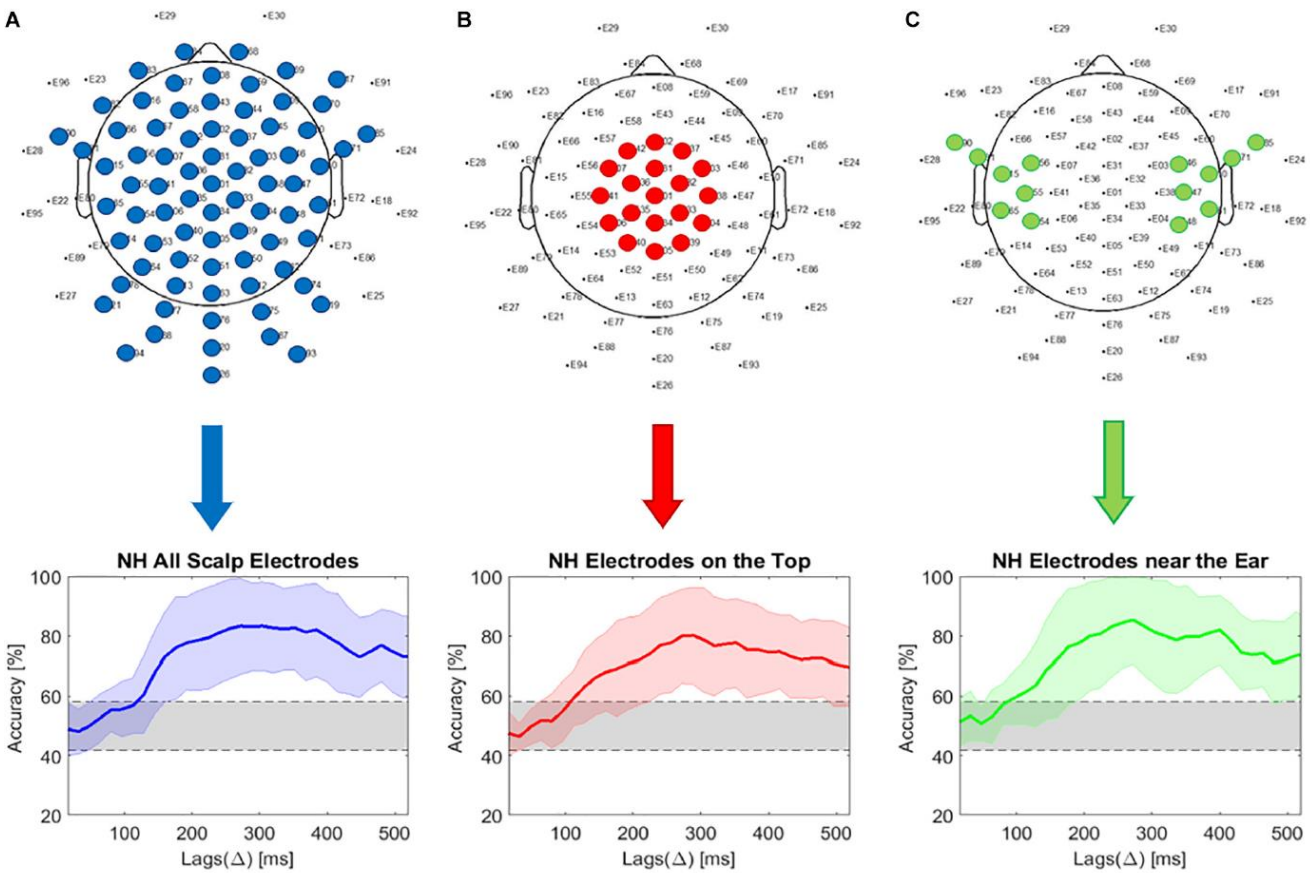
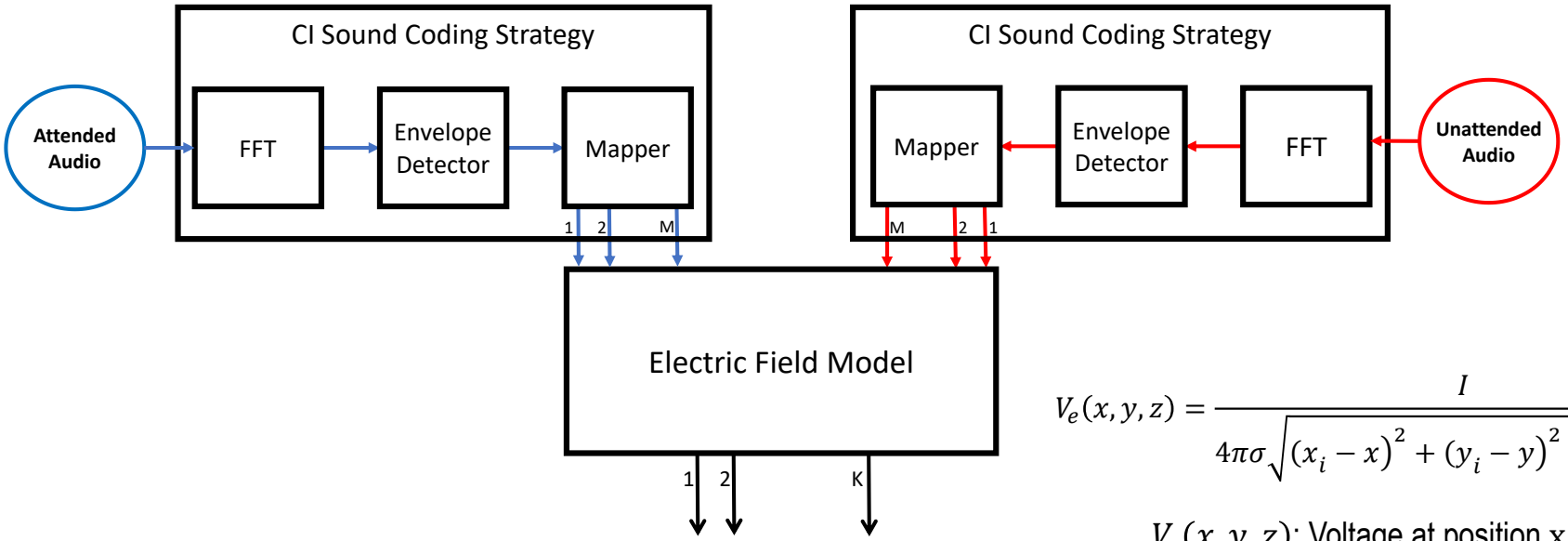


FIGURE 3 | Comparison of electrode locations. **(A)** Selective attention decoding accuracy using all electrodes (eye electrodes 29 and 30 and scalp EEG electrodes spared at the position of the cEEGrid are excluded from the analysis); **(B)** Selective attention decoding accuracy using 19 electrodes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

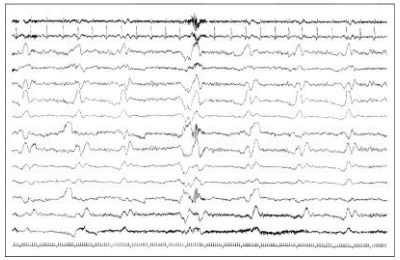
Nogueira et al 2019, Frontiers Neuroscience.

Artefact Model

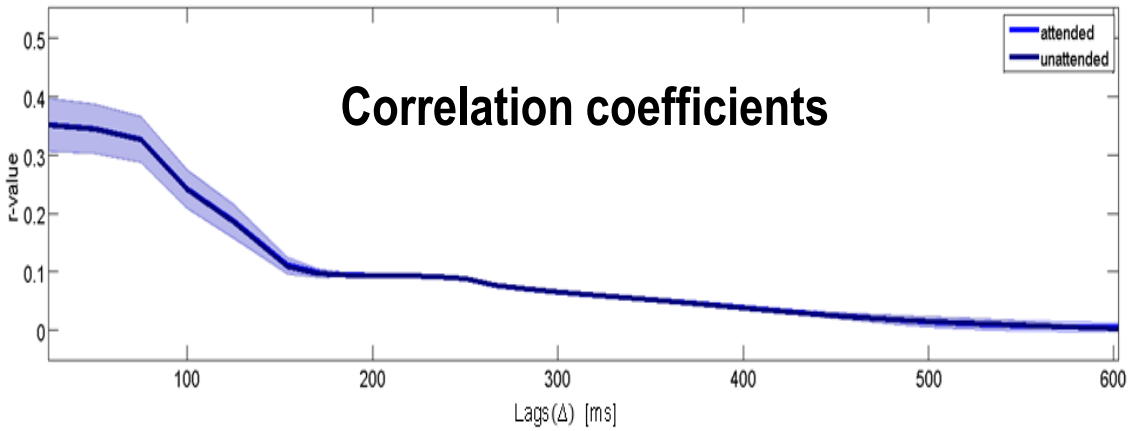


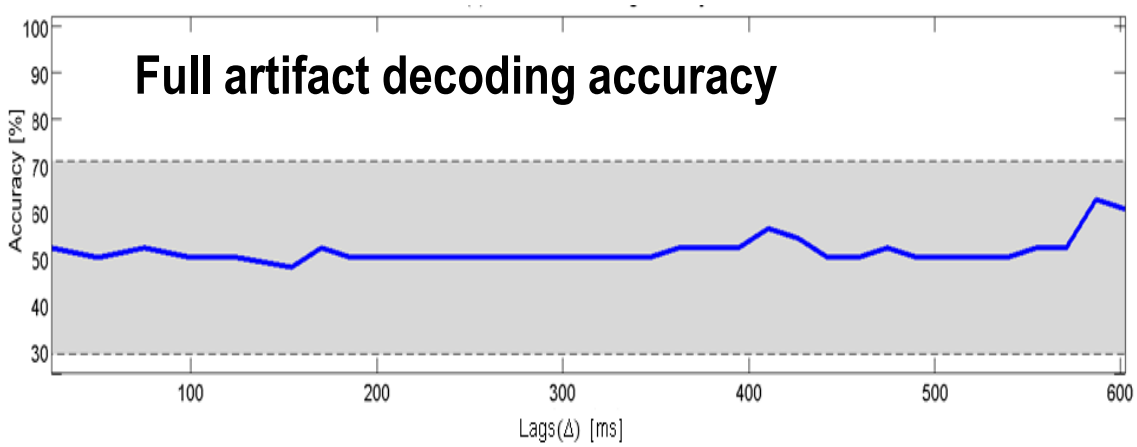
$$V_e(x,y,z) = \frac{I}{4\pi\sigma\sqrt{(x_i - x)^2 + (y_i - y)^2 + (z_i - z)^2}},$$

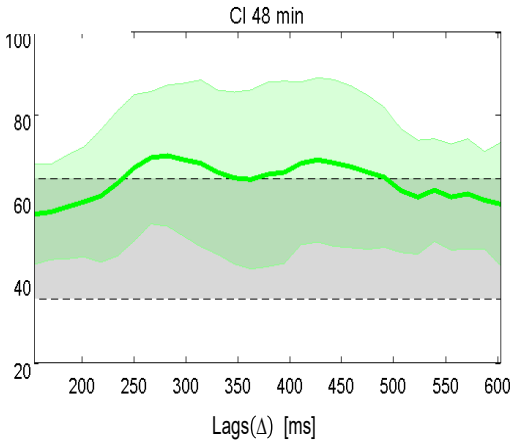
$V_e(x,y,z)$: Voltage at position x,y,z caused by stimulation of CI electrode i with curent I



Decoding Selective Attention with an Artifect Model







Least Square Estimation Method

- Speech reconstruction from the single-trial EEG data was obtained by training decoders using a regularized least square estimation method.
  - Parameter Lag (Δ) → $\hat{x}_{a,u}[k] = \sum_{n=0}^{N-1} \sum_{l=0}^{L-1} w_{n,l} y_n[k + \Delta + l]$
  - $J_{LS}(\boldsymbol{W}_a) = E\{|x_a[k] - \boldsymbol{W}_a{}^T \boldsymbol{Y}[k]|^2\}$ → minimize
  - K: time; n: electrode; y_n:recorded EEG; $w_{n,l}$:decoder; $\hat{x}_{a,u}$:reconstructed signal

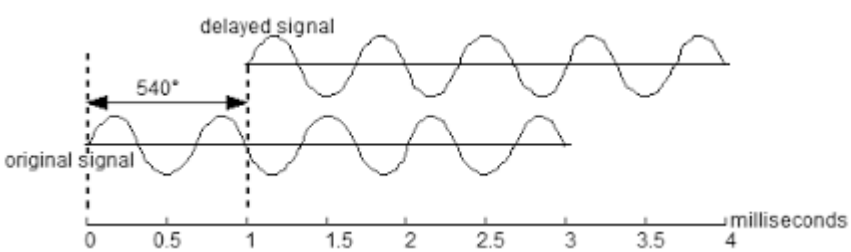
- EEG recording matrix with delayed versions (Lags)

$$Y = \begin{bmatrix} r_1(0) & r_1(1) & \cdots & r_1(\tau_{\max}) & \cdots & r_1(T) \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & \cdots & r_1(0) & \cdots & r_1(T - \tau_{\max}) \\ \vdots & \vdots & & & & \vdots \\ r_n(0) & r_n(1) & \cdots & r_n(\tau_{\max}) & \cdots & r_n(T) \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & & r_n(\tau_{\max}) & & r_n(T - \tau_{\max}) \end{bmatrix}$$

Detection Selective Attention
Temporal Response Function (TRF)

Lags, typically expressed in [ms] account for delay between EEG and input sound

$$Y = \begin{bmatrix} r_1(0) & r_1(1) & \cdots & r_1(\tau_{\max}) & \cdots & r_1(T) \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & \cdots & r_1(0) & \cdots & r_1(T-\tau_{\max}) \\ \vdots & \vdots & & & & \vdots \\ r_n(0) & r_n(1) & \cdots & r_n(\tau_{\max}) & \cdots & r_n(T) \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & & r_n(\tau_{\max}) & & r_n(T-\tau_{\max}) \end{bmatrix}$$

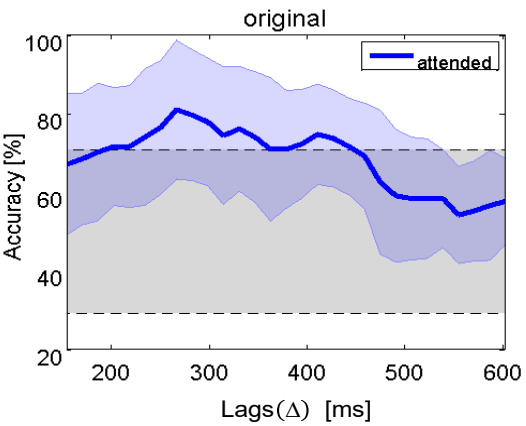


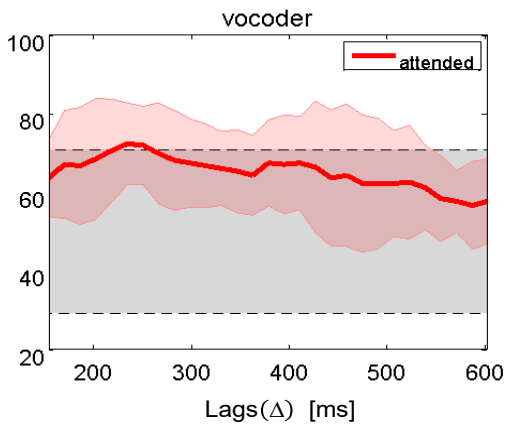
- Study 1: Investigate whether the degradation of spectral resolution as produced by a CI impacts the classification of selective attention

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)

(a) Attended Decoder





Lags: Delayed versions of the EEG

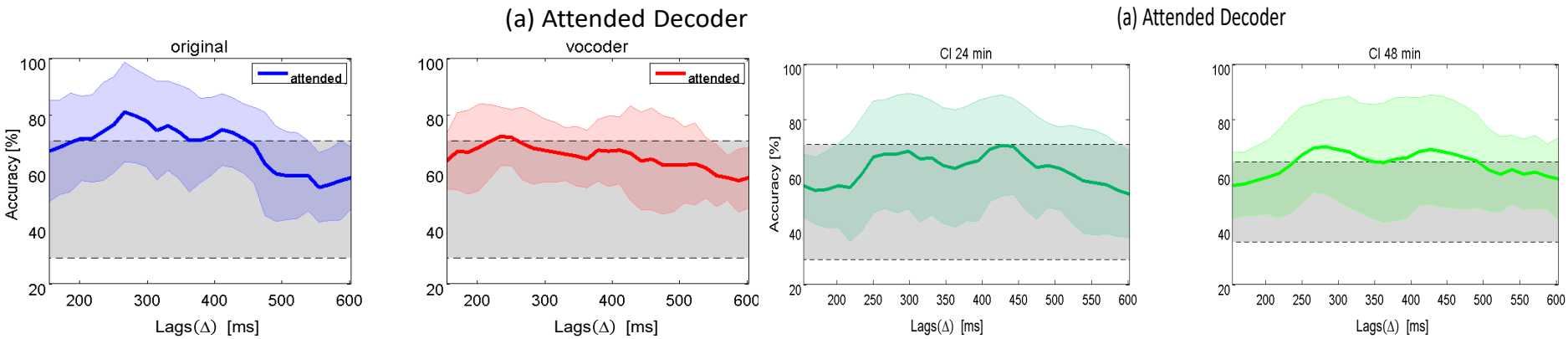
- No significant difference between Original speech or Vocoded Speech
- Envelope is preserved

Grey area: Chance Level

- Study 2: Investigate whether CI artifact impacts the classification of selective attention

Methods:
Study 3: Procedure

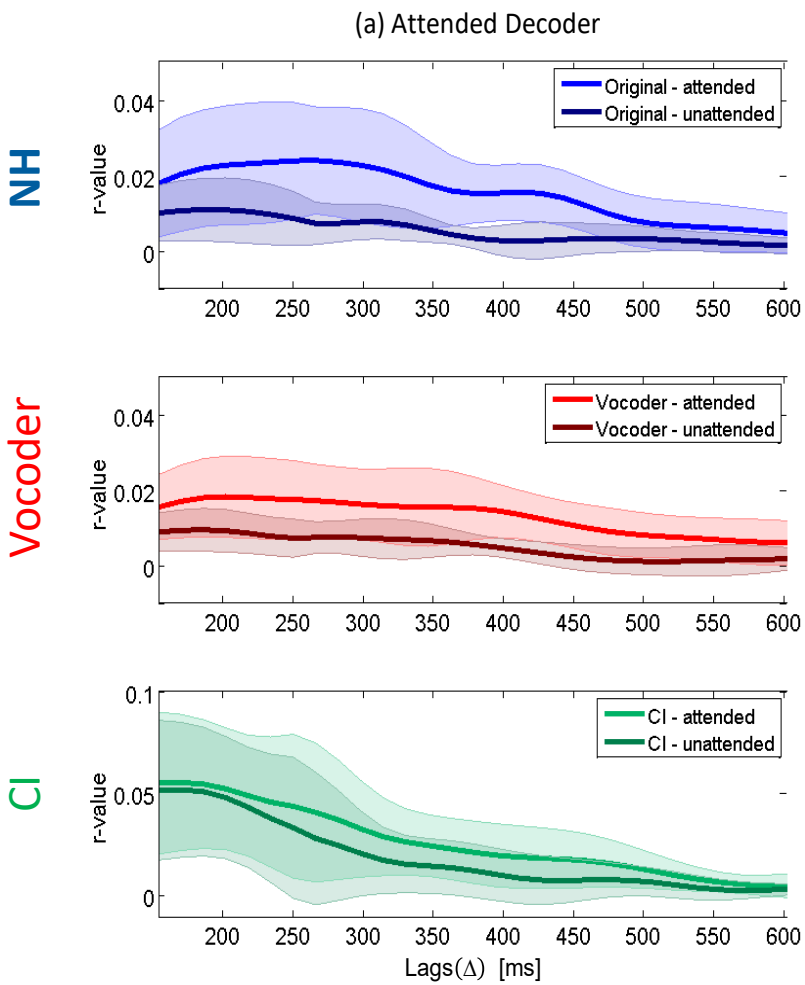
- N=12 CI users and 12 NH listeners
- Recording 48 minutes (CI), 24 minutes (NH)



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

Nogueira et al 2019, IEEE Transactions Biomedical Engineering.

Correlation Coefficient Analysis

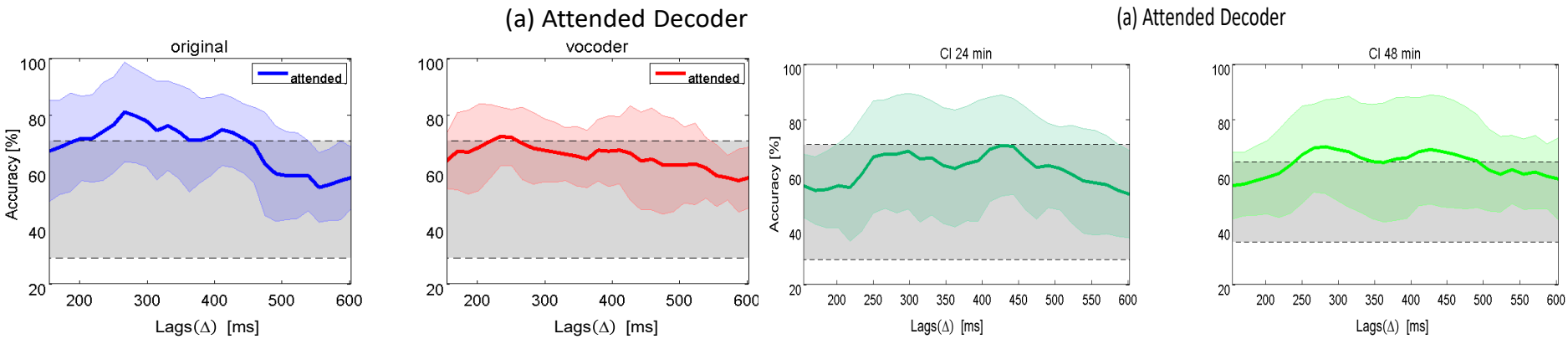


Nogueira et al 2019, IEEE Transactions Biomedical Engineering.

- 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)

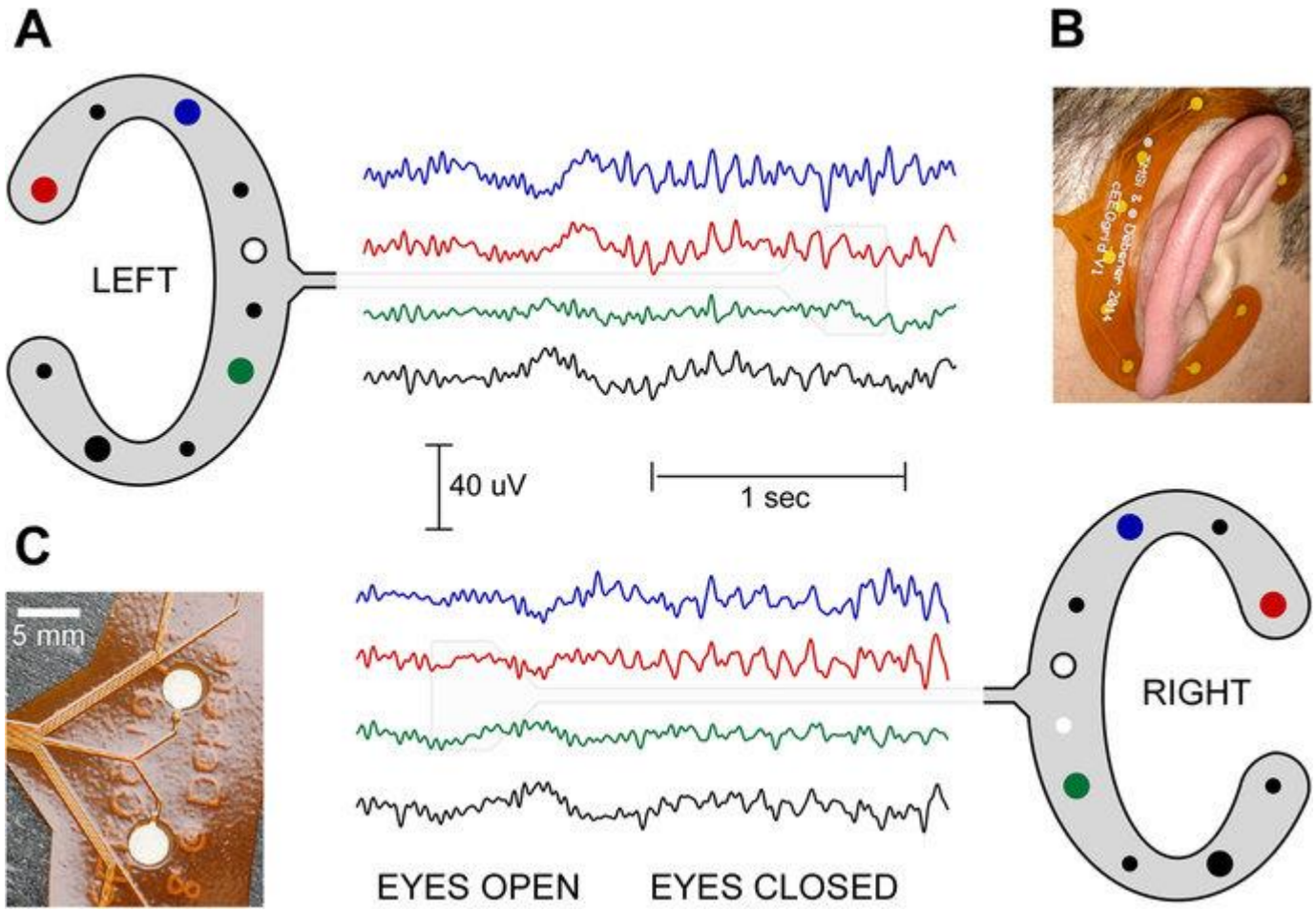


- Decoding accuracy in CIs is above chance level
  → Less influence from the artifact than expected

Nogueira et al 2019, IEEE Transactions Biomedical Engineering.

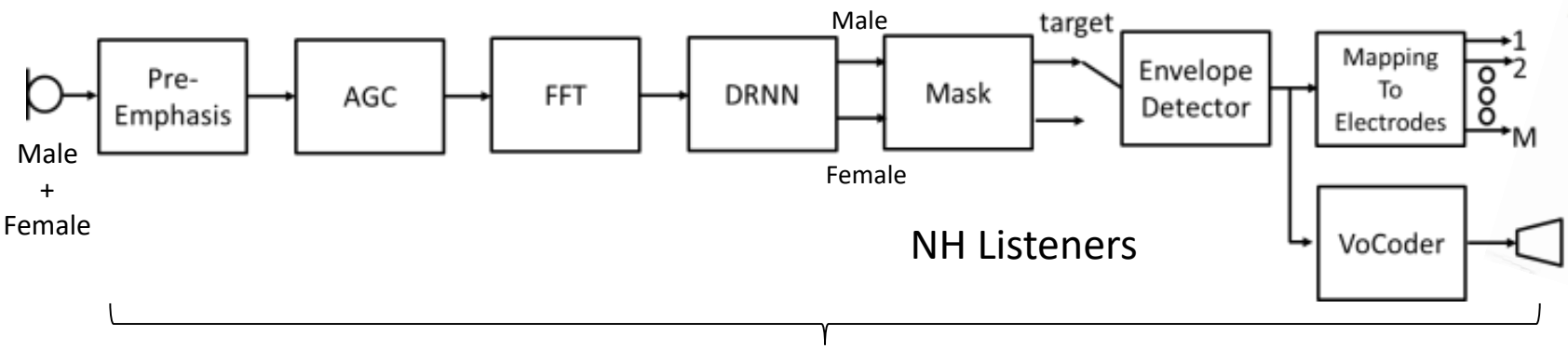
Study 3: cEEGrid electrode array

Debener et. al. 2015



Cochlear Implants
Blind Source Separation

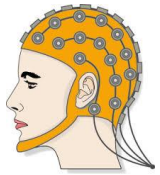
Front-end signal processing to address the “cocktail party” scenario

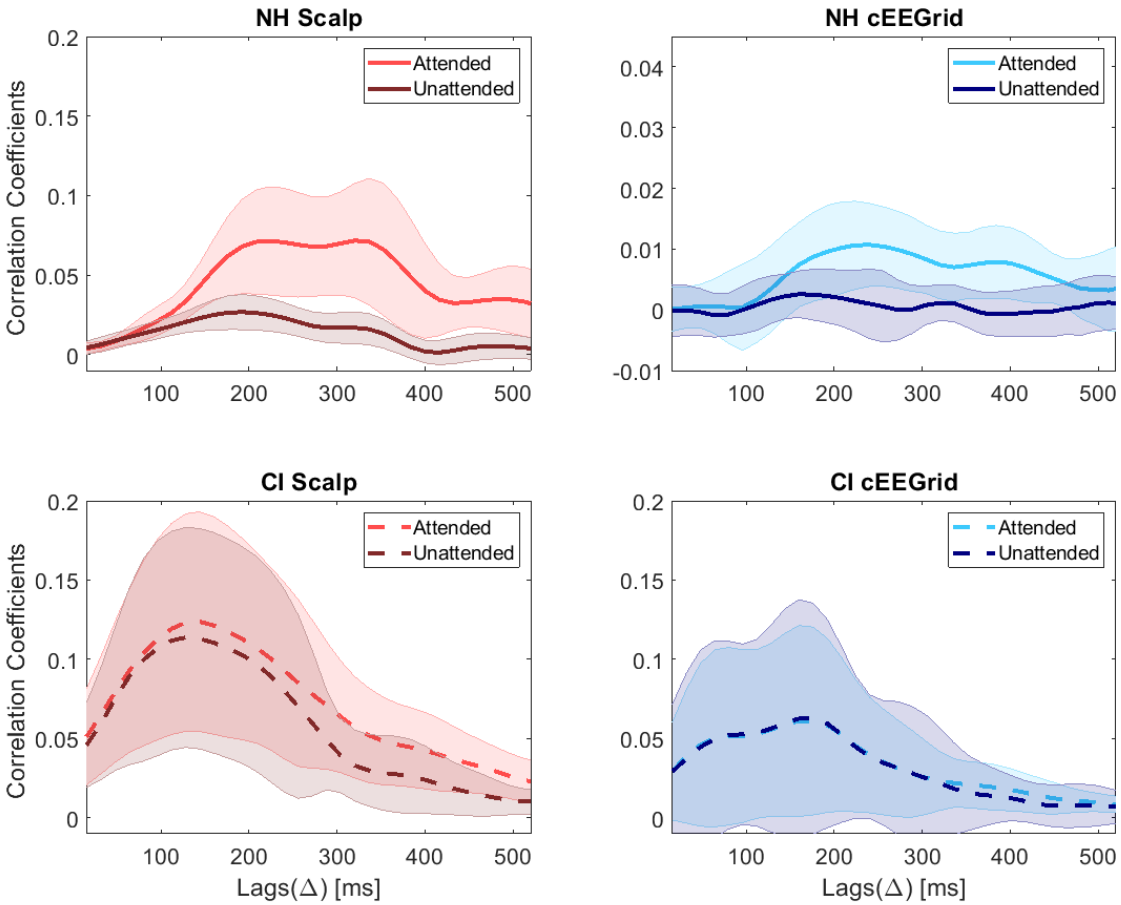


Speech in Speech Task:

Target: Male Speaker (HSM)
Interference: Female Speaker

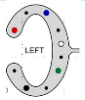
Selective Attention Accuracy
High Density Scalp Electrodes





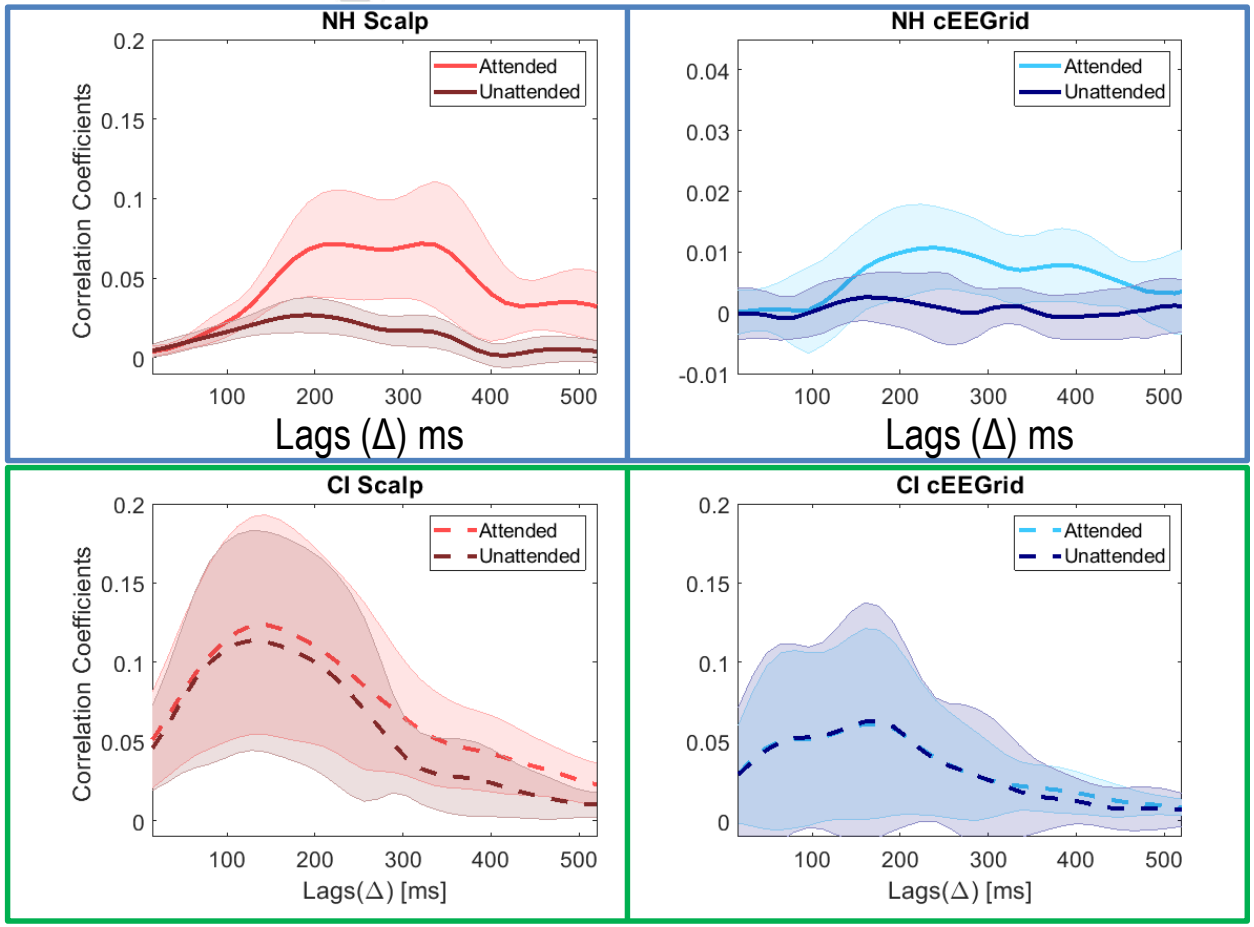
Nogueira et al 2019, Frontiers Neuroscience.

Selective Attention Accuracy cEEGgrid vs Scalp Electrodes



NH

CI



Nogueira et al 2019, Frontiers Neuroscience.