Journal of Neural Engineering

PAPER

Real-time control of a hearing instrument with EEG-based attention decoding

Jens Hjortkjær[1,2,*], Daniel D E Wong[3,4], Alessandro Catania[1], Jonatan Märcher-Rørsted[1], Enea Ceolini[5], Søren A Fuglsang[1,2], Ilya Kiselev[5], Giovanni Di Liberto[6], Shih-Chii Liu[5], Torsten Dau[1], Malcolm Slaney[7] and Alain de Cheveigné[3,4]

[1] Hearing Systems Section, Department of Health Technology, Technical University of Denmark, Kgs. Lyngby, Denmark
[2] Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital—Amager and Hvidovre, Copenhagen, Denmark
[3] Laboratoire des Systèmes Perceptifs, CNRS UMR, Paris 8248, France
[4] Département d'Etudes Cognitives, Ecole Normale Supérieure, PSL, Paris, France
[5] Institute of Neuroinformatics, University of Zurich and ETH Zurich, Zurich, Switzerland
[6] School of Computer Science and Statistics, Institute of Neuroscience, Trinity College, The University of Dublin, Dublin, Ireland
[7] Center for Computer Research in Music and Acoustics (CCRMA), Stanford University, Stanford, CA, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: jhjort@dtu.dk

Keywords: EEG, attention decoding, BCI, hearing

- Hearing aids: Which sound source to amplify?
- Cognitively-controlled hearing aids (decode attention via EEG)
  - Correlation between EEG & Speech envelope (Ding and Simon 2012a & O'Sullivan 2015) – Offline studies
- "hope" for a real-time implementation
- Challenges:
  1. BCI attenuates ignored speech – is it possible to switch back to "ignored" speech?
  2. What if the attenuation leads to inaudible speech that the user cannot hear anymore? (min. amplitude)
  3. Can visual feedback help in such situations?
  4. Are users tolerant for latency & BCI misclassifications?
  5. How about accuracy & speed of the BCI?
  6. How should the acoustic signal be rendered to the user?
  7. Implementable in a hearing aid?
  8. Does the benefit require training?

This paper presents a real-time implementation of a closed-loop AAD system.

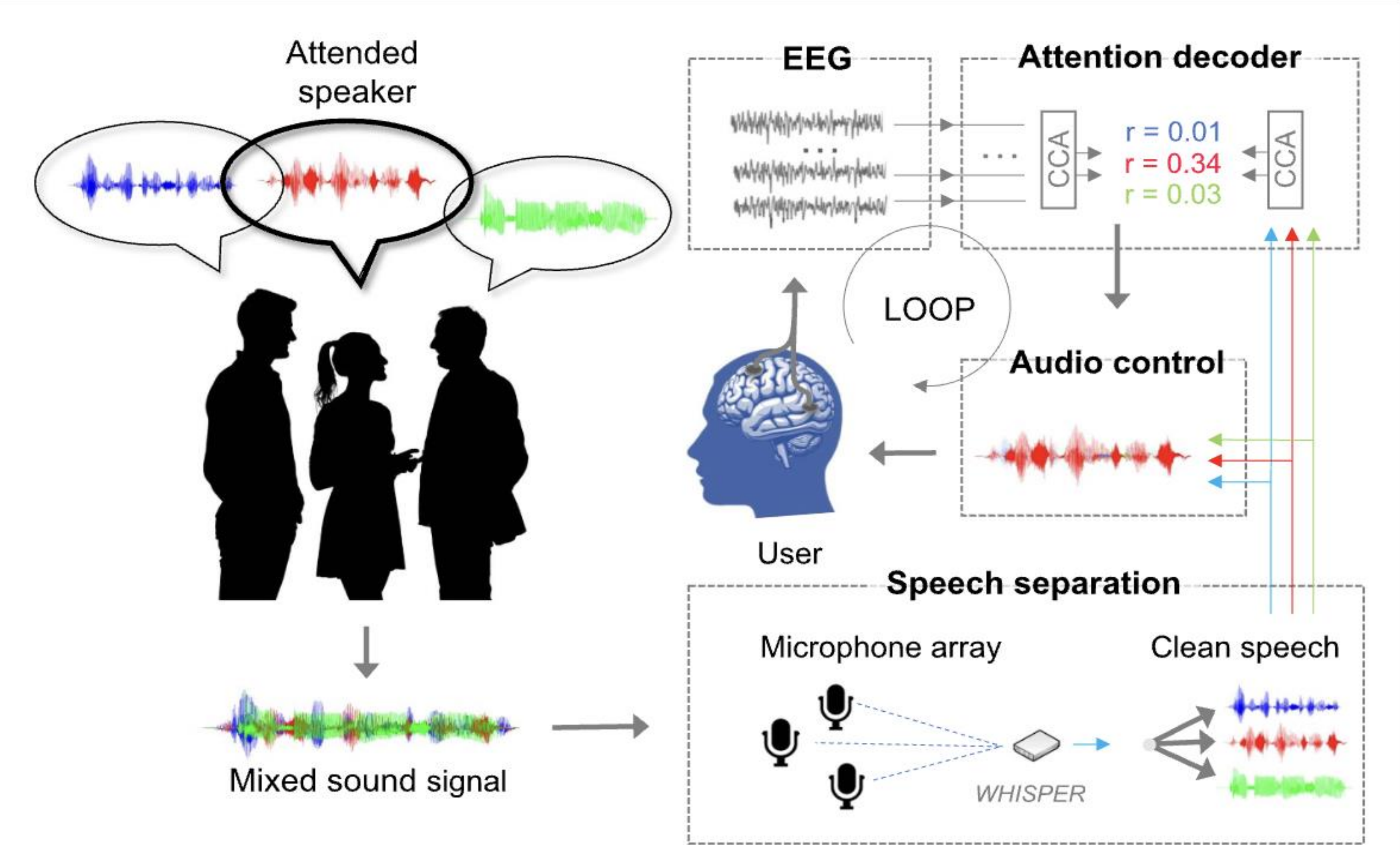


Figure 1. The cognitively controlled hearing aid in overview. Speech is picked up by an ad hoc array of microphones and processed to produce clean speech streams, one per speaker. Continuous EEG signals from the user are synchronized with the separated audio streams and transmitted to an attention-decoding module that classifies the attended speech stream based on correlations between audio and EEG features. The time-varying classification outputs are then used to control the relative gains of the separated audio streams.

Cognitive Control of a Hearing Aid

1. WHISPER (microphone array)
2. EEG + CCA (AAD decoding)
3. Audio Control

- WHISPER:
  - 4 microphones
  - Beamforming algorithm
  - Latency = algorithmic latency + wireless transmission delay + negative latency
  - 11 ms (only for the algorithm)

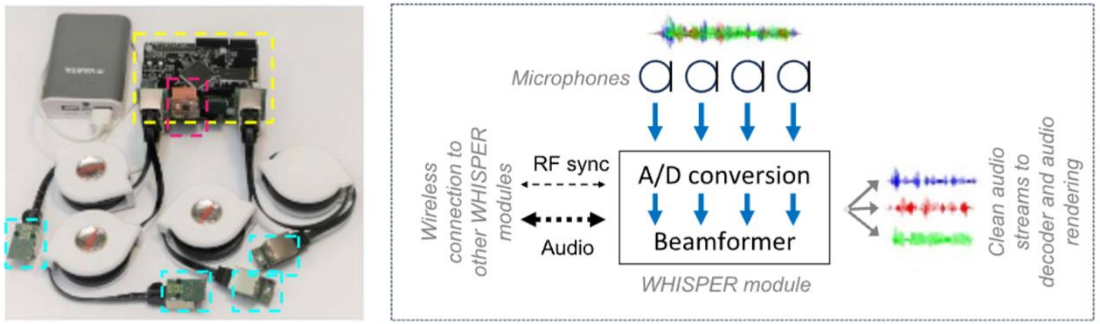
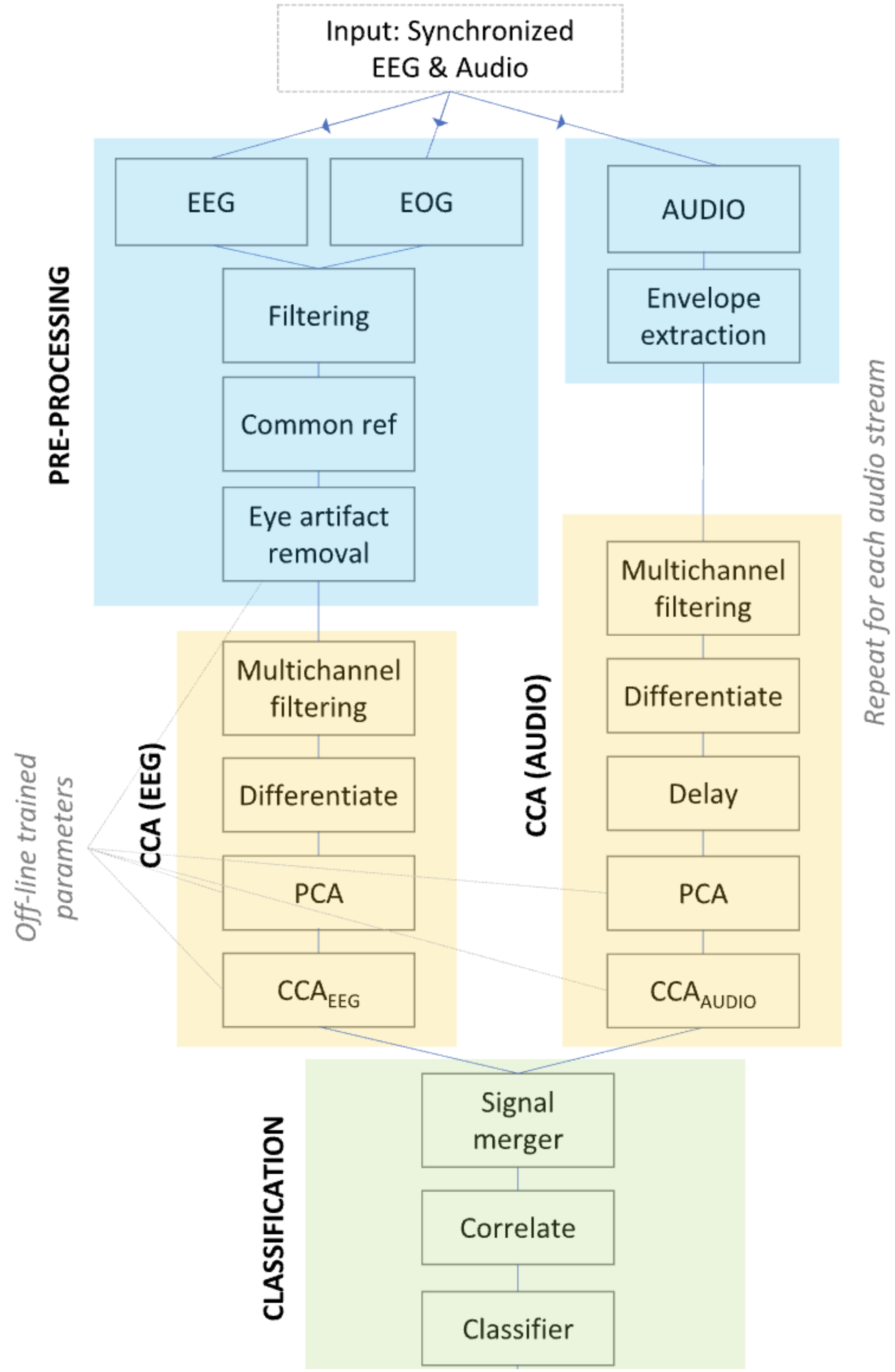
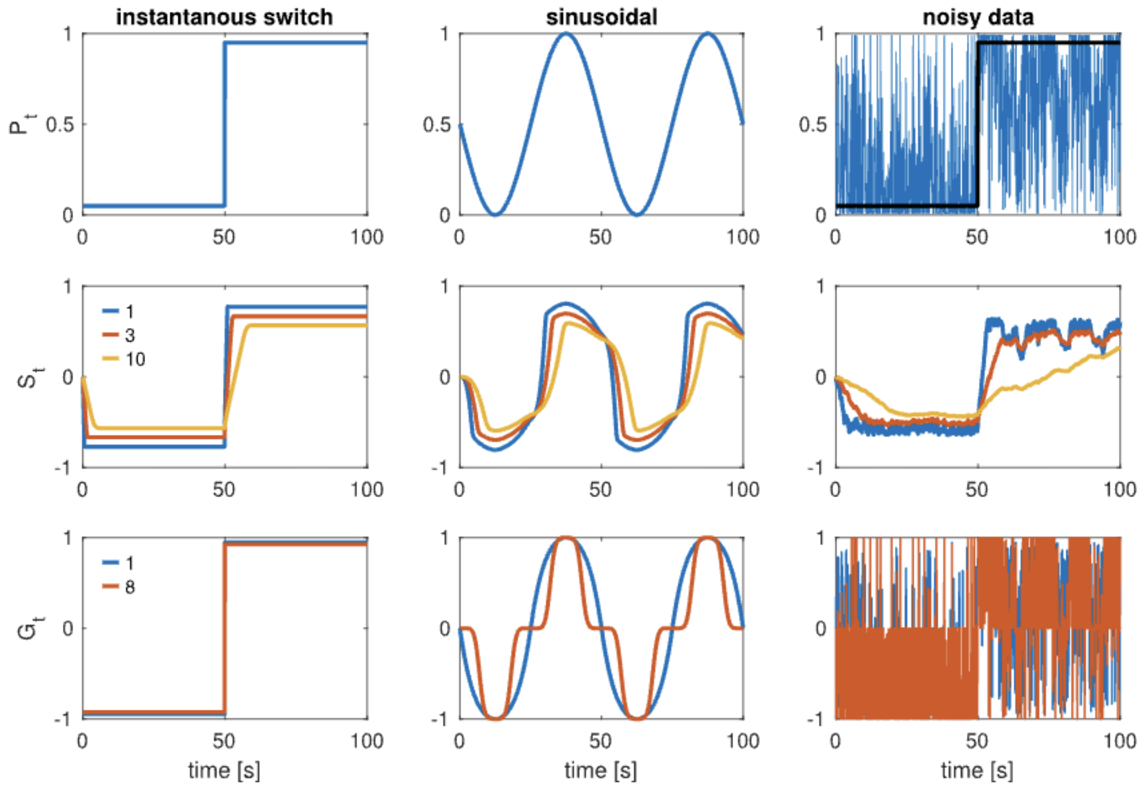


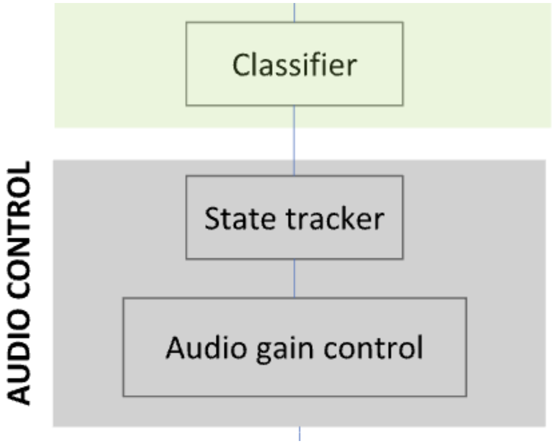
Figure 2. Left: prototype WHISPER module (yellow box) equipped with four microphones (cyan boxes) and a wireless synchronizer (red box). Right: a WHISPER module acquires audio from up to four microphones and exchanges sampled audio and synchronization signals with other WHISPER modules or a common laptop. Microphone streams are merged and processed by a beamformer to produce clean streams for the decoder and audio rendering modules.

- AAD:
  - OpenVibe platform; LSL
  - Canonical Correlation Analysis (CCA)
  - Better accuracy than forward/backward models (de Cheveigné 2018)
  - EEG and audio signals are transformed via multichannel filterbank
  - CCA weight matrices from “training” data are stored for online decoding
  - Offline studies: 70-80% with decoding windows of 6-8s.



- Audio Control:
  - Classifier output – Adjust the gain
  - Instantaneous vs state-tracker

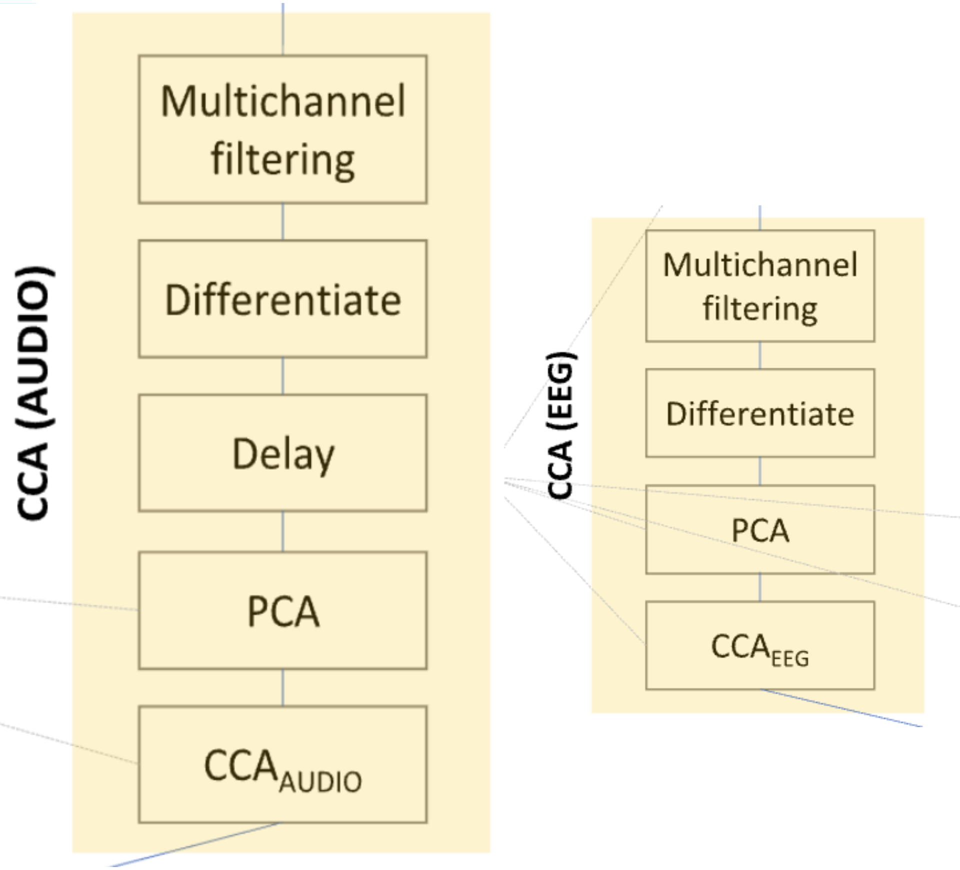




- EEG systems:
  - Smarting 24-channel (real-time demo)
  - Biosemi ActiveTwo 64-channel (2 experiments)

- EEG training data:
  - Offline EEG acquired per subject
  - Dichotic presentation with 1 male and 1 female audiobooks for 30-40 mins; 50s long trials
  - Subjects were asked to attend to one side/trial
  - Switch in stream per subjects
  - Extract CCA weights and store!

- Online audio processing:
  - LPF at 20Hz
  - Down-sampled to 64 Hz
  - Power of 0.3 (compression)



- Online audio-EEG classification
  - CCA-transformed EEG and audio
  - Each CC-pair was z-scored based on the training data
  - SVM classifier – probabilistic outputs

- Audio gain control:

Different gain control functions were investigated. In experiment I, classification probabilities were mapped to a time-varying gain control signal G_t using a non-linearity corresponding to a modified Weibull cumulative distribution function:

$$G_t = \begin{cases} 1 - e^{-\left(\frac{2P_t - 1}{\lambda}\right)^k} & \text{for } P_t \geqslant 0.5 \\ e^{-\left(\frac{1-2P_t}{\lambda}\right)^k} - 1 & \text{for } P_t < 0.5 \end{cases} \quad (1)$$

The state at time step t is computed as:

$$S_t = S_{t-1} + C\frac{2}{K \cdot F_s}\mathrm{ITR}_t - \mathrm{sign}\left(S_{t-1}\right)\left|S_{t-1}\right|^{\alpha} \quad (2)$$

where C is the attention state corresponding to the classified speaker:

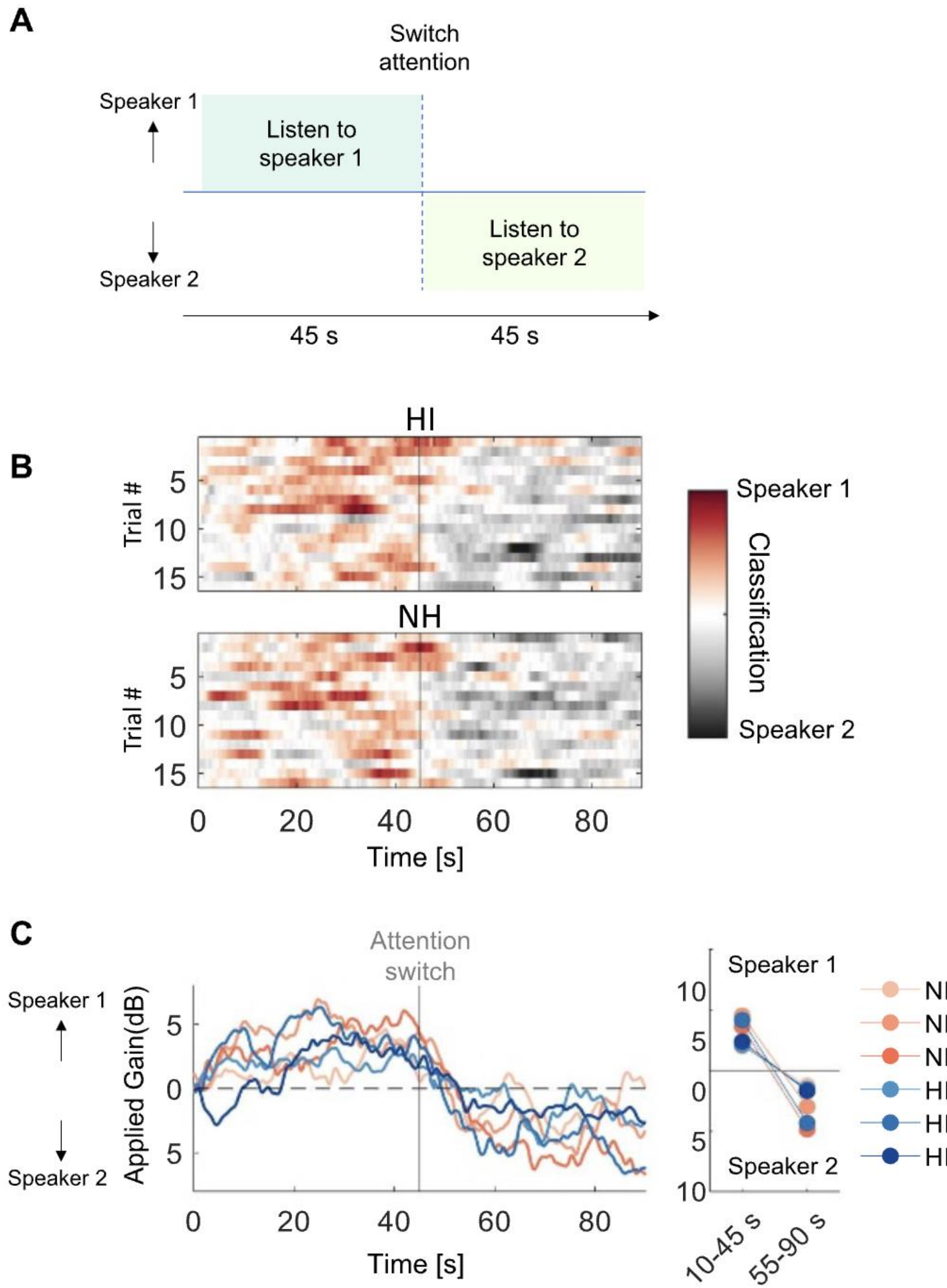
$$C = \begin{cases} 1 & \text{for } P_t \geqslant 0.5 \\ -1 & \text{for } P_t < 0.5 \end{cases} \quad (3)$$

Demonstration with real talkers:

https://iopscience.iop.org/article/10.1088/1741-2552/ad867c

Experiment I: 3 young normal-hearing; 3 older hearing-impaired without hearing aids.

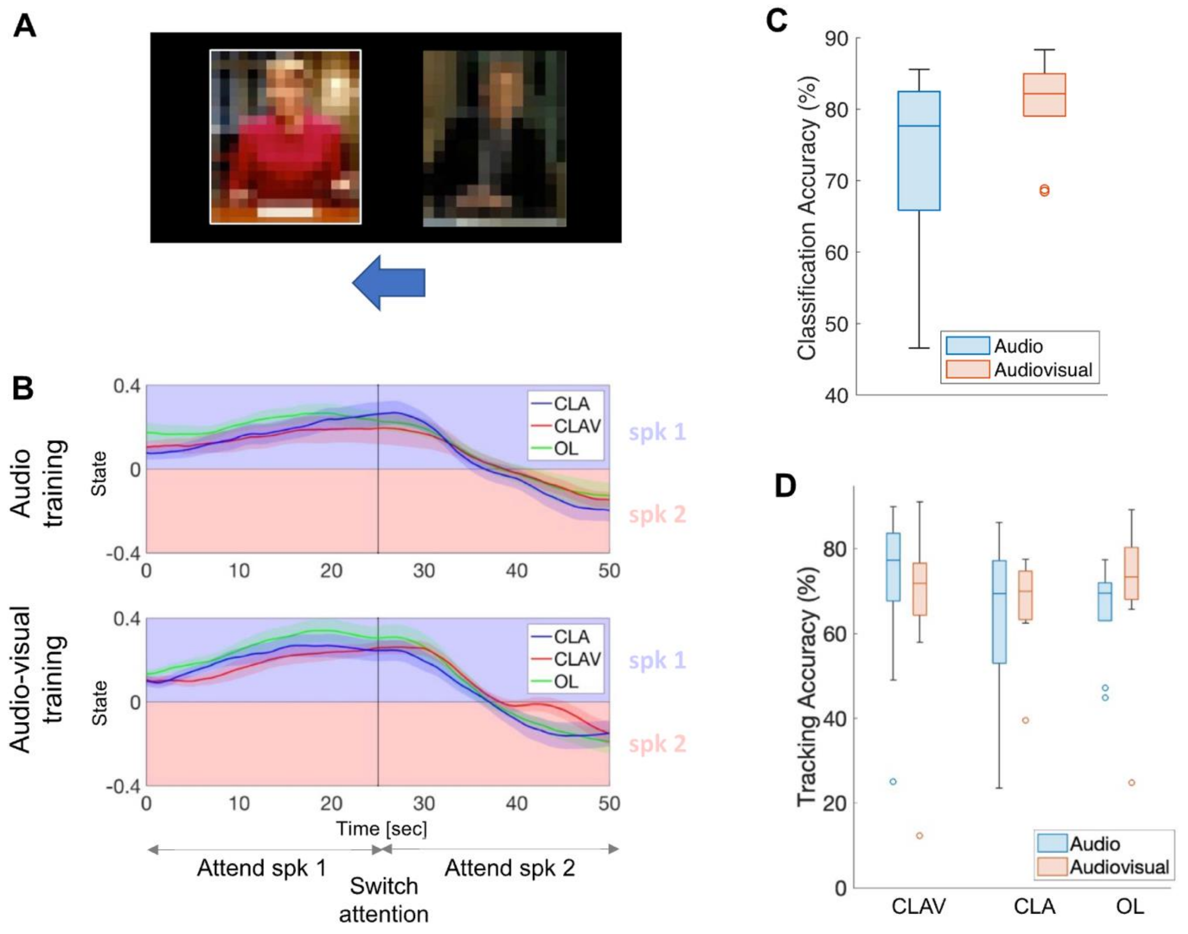
Switch audio - 4.7s



- **Experiment II:** 20 young normal-hearing listeners with audiovisual speech.

- Closed-loop AV
- Closed-loop Audio
- Open-loop

Switch audio – 12.6 s



Discussion:

1. Seamless BCI control based on attention rather than user inputs
2. Stimulus-driven vs generic EEG responses (e.g., alpha)
   - Recover the previously attenuated sources by up to 10 dB
   - But the boundary conditions yet to be explored
   - At very low SNRs
3. Inaccurate decoding can be detrimental
4. Better decoding with Neural Networks?
5. Ground truth for AAD – listen to a single speaker and train a model
6. Speed vs dynamicity of conversations
7. Alternative EEG solutions: Ear-EEG, ECoG
8. Speech separation tech:
   - Nodes in changing environments
   - Synch with wireless nodes

Discussion:

9. Audio control strategy – re-render the audio in different settings
10. Attention training:
    - Reward - enhanced attended speech
    - Task difficulty

11. The glass is half full, and yet half empty.