Inner Speech Decoding from EEG and MEG

Richard Csaky[1,2,5], Mats W.J. van Es[1,2], Oiwi Parker Jones[2,3,4], Mark Woolrich[1,2]

University of Oxford; [1]Oxford Centre for Human Brain Activity, Department of Psychiatry; [2]Wellcome Centre for Integrative Neuroimaging; [3]Department of Engineering Science; [4]Jesus College; [5]Christ Church

BACKGROUND

- Limited research on **inner speech** with **non-invasive** methods [Panachakel et al., 2021].
- Differences between **repetitive** (cued either with a text or audio of a word), and **self-generated inner speech** have not been analysed, particularly in terms of decoding.
- Such research can lead to **word-level** communication with **BCI**s [Metzger et al., 2022].

What inner speech decoding performance can be achieved in EEG and MEG with a large number of per-participant trials?

Can we transfer decoders across sessions and tasks?

DATA COLLECTION

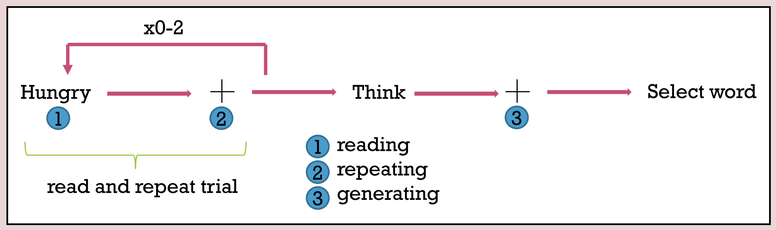

Figure 1: **Experimental protocol for version 1**. Words and cross cues appeared for 0.8-1.0 seconds, followed by 0.8-1.0s blank screen. (1) is the **silent reading** trial, while cross cue (2) is the **repetitive**, and (3) is the **self-generated inner speech** trial. Participants indicated their self-generated word (from the set of 5 words) at the end of the trial.

Version 1

	EEG	MEG
P1 sessions	6	6
P2 sessions	2	2
P3 sessions	2	2
total **reading** trials	3250	3250
total **inner speech** trials	5750	5750

- We collected both **MEG** (Elekta Neuromag 306-channel) and **EEG** (Easycap 64-channel) data, as well as **ECG, EOG, EMG** (on the jaw), and **eye-tracking** data.
- The 5 words used in the experiment are ***hungry, tired, pain, thirsty, toilet***.

Version 2

	EEG	MEG
P1 sessions	1	1
P2 sessions	1	1
P3 sessions	10	1
total **reading** trials	2080	520
total **inner speech** trials	16000	4000

- Instead of a single cross cue, **four consecutive crosses** were shown, spaced at **1-second intervals** so that participants repeated the word 4 times.
- Each cross was shown for 0.3 seconds followed by 0.7 seconds of blank screen.

Version 3

- **3750 silent reading** trials across 3 participants with combined MEG and EEG.
- No inner speech task in this version.
- We also collected **CTF** and **OPM** data for the same participants to compare across modalities, however, we do not have the results yet.

20K inner speech trials (EEG)

10K inner speech trials (MEG)

9K silent reading trials (EEG)

7.5K silent reading trials (MEG)

across **5 words**

Near chance inner speech decoding

30-40% silent reading

decoding accuracy

Take a picture for main author contact details

RESULTS

Preprocessing

- Bandpass filter between 1-40Hz -> bad segment and channel removal -> ICA with 64 components for MEG only

MEG inner speech decoding at chance level

Methods tried:

- Fully-connected NN; CNN; Linear Discriminant Analysis (LDA); Logistic Regression
- Channel selection; Using the covariance matrix of the trial as features
- Concatenating the 4 consecutive trials or averaging them
- Per-session decoding or using trials from all sessions

EEG inner speech decoding above chance in 3/10 sessions (P3 from version 2)

- **25%** cross-validated **accuracy** with method 1: **covariance matrix** features + **LDA** model
- **33%** cross-validated **accuracy** with method 2 (Danger of overfitting)
  a **single LDA** model trained on all trials from the **3 good sessions**
  concatenating the **4 consecutive trials** into a single epoch
  subtracting mean **session-level evoked** response and **covariance** from each epoch

Methods tried (other than the ones for MEG data):

- Trial-level normalization; temporal alignment of trials; denoising with PCA, Xdawn classifier with riemannian features; baseline correction; laplace denoising

Silent reading decoding above chance in both MEG and EEG

- Cross-validated 2-layer **linear neural network** trained on each participant separately
- Channel by timesteps **1-second epoch** flattened to a feature vector

	MEG version 1	MEG version 3	EEG version 3
Average accuracy across participants	39.33%	35%	30%

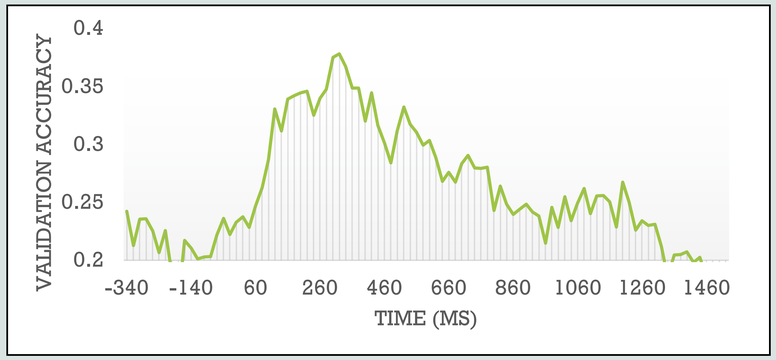

Figure 2: **Silent reading** validation accuracy **time course** from a **sliding window** (100ms) **LDA** model trained on the MEG data of 1 participant from version 1. The word is shown at 0ms and stays on for 0.8-1.0 seconds.

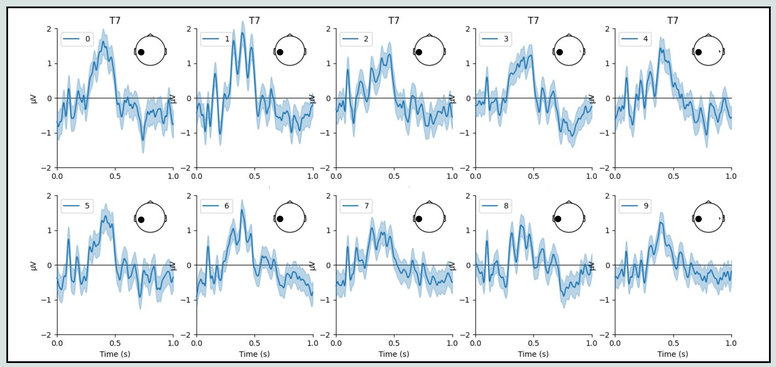

Figure 3: **Inner speech evoked** responses for the **T7 EEG** channel of participant 3 in version 2 of the experiment. Each plot shows a different session.