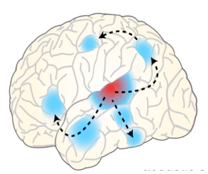
Time-Constrained Neural Decoding From Multiple Auditory Cortical Areas

Krishna Puvvada[1], Jonathan Z Simon[1,2,3]
[1]Dept of Electrical & Computer Engineering, [2]Dept of Biology, [3]Institute for Systems Research
University of Maryland College Park

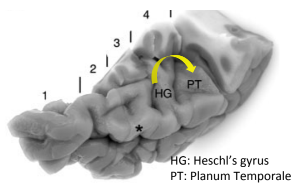
Computational Sensorimotor Systems Lab

Background:

Sensory Signals are processed hierarchically in neural domain.

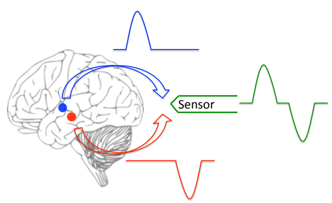


(Peelle et al. 2012)



(Sweet et al. 2012)

Electromagnetic responses from nearby neural sources add up in sensor domain when recorded using MEG/EEG.



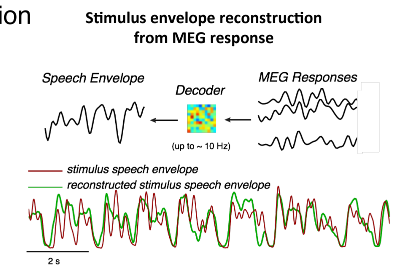
Experimental Design

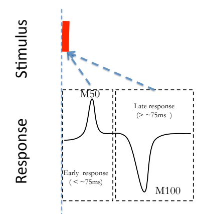
- 3 co-located speakers simultaneously narrating separate stories.
- 220 s stimulus duration, 3 attention conditions x 3 repetitions
- N = 9 Subjects.
- In each trial the subject counts the number of times a given keyword is heard in the attended story.
- At the end of 3 trials, the subject reports a summary of story.
- Magnetoencephalography (MEG), 157 channels.
- 1kHz sampling , Time-shifted PCA based de-noising.
- Spatial filtering used to reduce 157 channels to 8, more reliable, virtual channels.

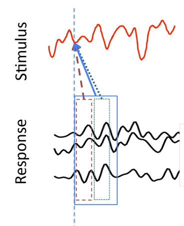
Method:

Stimulus reconstruction

- Temporal envelope of stimulus is reconstructed from cortical responses using optimum linear filters.
- Reconstruction based on integrating neural responses over a temporal window.
- Correlation between reconstructed and actual envelope is used as metric as how faithfully the foreground or background is represented in cortical responses.
- Differing latency ranges in the temporal integration window represent different neural areas, with different processing specializations and representations.

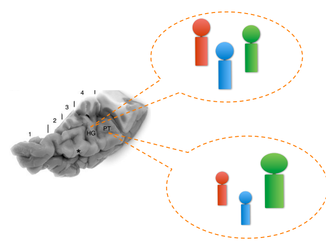






Hypothesis

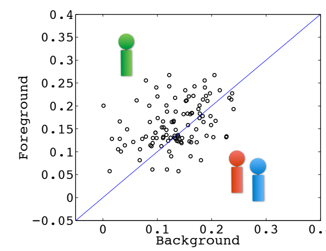
- Attended (foreground) speech is better represented in later neural sources.



Analysis & Results

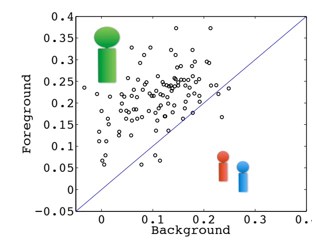
Stimulus reconstruction using only early responses

- HG (early neural source) represents attended and unattended speech with almost equal fidelity. ($P < 0.001$)



Stimulus reconstruction using only late responses

- PT (late neural source) is highly selective to attended speech. ($P < 1e\text{-}15$)



Discussion

- Hierarchical processing of sensory signals leads to time separability in sensory domain, which can be used to achieve functional separability of neural sources in high time-resolution imaging techniques such MEG/EEG.
- Demonstrated using selective listening paradigm in multi-speaker scenario in MEG imaging.

References: 1) Peelle, J. E., & Davis, M. H. (2012). Neural Oscillations Carry Speech Rhythm through to Comprehension. *Frontiers in Psychology*, *3*, 320.
2) Sweet et al (2005). Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *Journal of Comparative Neurology*, *491*, 3.
3) Ding N. and J. Z. Simon, (2012) *The Emergence of Neural Encoding of Auditory Objects While Listening to Competing Speakers*, PNAS, 109(29), 11854-11859.
Acknowledgements: Funding from NIH R01 DC 008342