Cortical Responses Time-Locked to Continuous Speech in the High-Gamma Band Depend on Selective Attention

Vrishab Commuri[1], Joshua P. Kulasingham[2], Jonathan Z. Simon[1]

[1]University of Maryland, College Park; [2]Linköping University

Computational Sensorimotor Systems Lab

Introduction

Auditory **cortical** responses to speech obtained by magnetoencephalography (MEG) show robust speech tracking in the **high-gamma band (70-200 Hz)** with latency ~40 ms [1], but little is currently known about whether such responses depend at all on the focus of selective attention. In this study we address:

- What differences are there in high-gamma cortical responses between the cases of male (F0 ~100 Hz) vs female (F0 ~200 Hz) speech?
- Do these early high-gamma cortical responses to speech, thought to originate from primary auditory cortex [2], depend on selective attention?

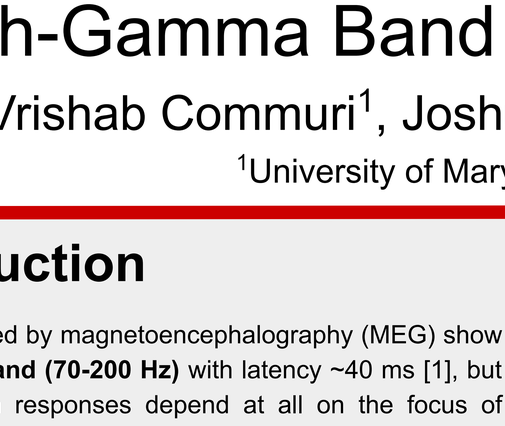
Methods

Experiment Design

MEG data were collected from 22 subjects (10 female, average age 22.6 years). Subjects listened to isochronous (fixed-rate) speech from two synthesized speakers in single-speaker and cocktail-party paradigms. These data were obtained in a previously published study [1].

Stimulus Representation

1. Carrier is derived from bandpass filtering the data from 70-200 Hz.
2. High-Frequency Envelope
   - Compute auditory spectrogram [3]
   - Filter spectrogram from 70-200 Hz
   - Average across frequency bins to obtain the envelope time series.



Source Localization

Volume source space comprising a grid of 7mm voxels. Sources localized using minimum norm estimation to a temporal lobe region of interest. Current dipole vectors were normed and averaged such that peaks and troughs were both represented as peaks.

TRF Estimation and Statistical Tests

- High frequency carrier and envelope used as predictors, with Temporal Response Functions (TRFs) estimated jointly using boosting [4] via eelbrain [5].

$$y = \tau_e * e + \tau_c * c + n$$

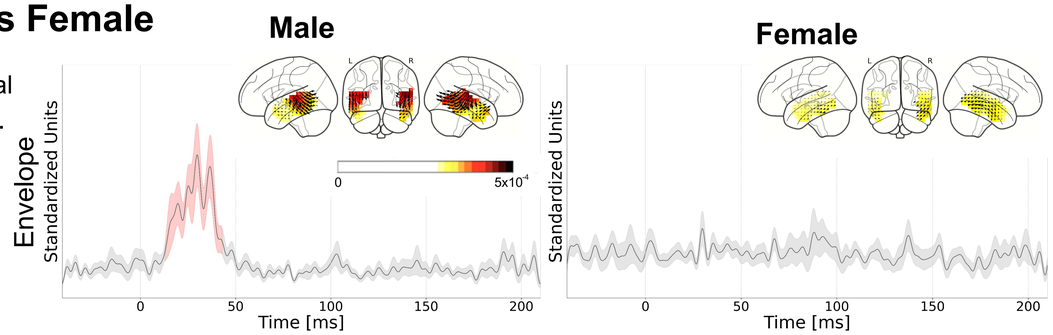
y: output (MEG)
τ: TRF
e: envelope
c: carrier
n: noise

- Statistical tests with null models from circularly shifted predictors using Threshold-Free Cluster Enhancement (TFCE) [6].

Results

Responses: Male vs Female

- Male speech: strong cortical responses with ~40 ms lag.
- Female speech: no significant response.

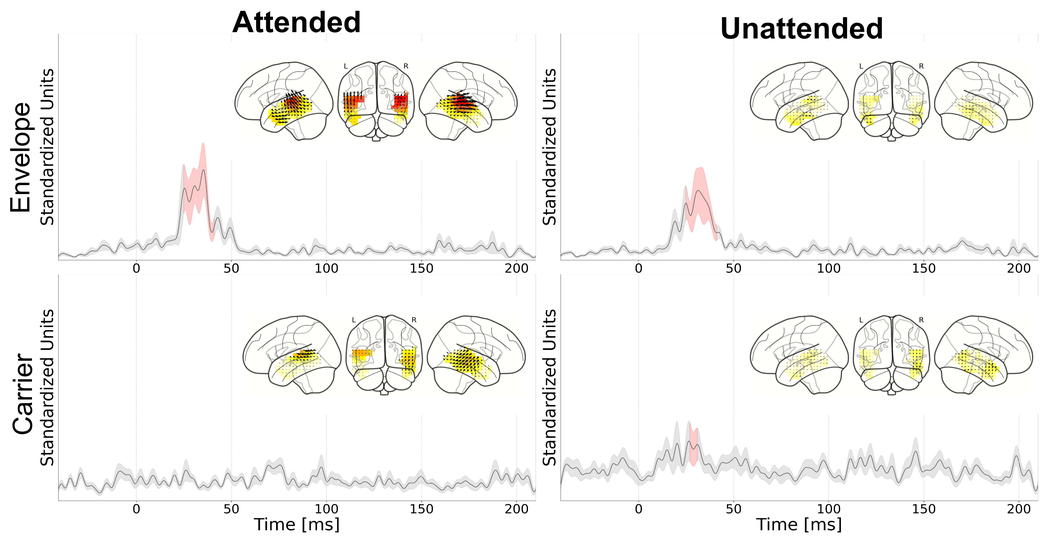


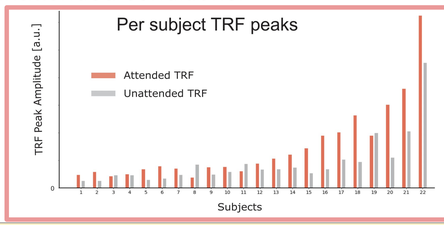
Responses: Attended vs Unattended

- Male speech: still robust despite concurrent female speech.
- Significant cortical responses observed for both attended and unattended male speech.

Further investigation:

- Peaks in the 20-50 ms range were extracted from each subject TRF.
- Per-subject comparison between attended and unattended conditions:





- TRF peak amplitudes are significantly larger for attended compared to unattended speech
- paired t-test $t_{21} = 3.85$, $p < 0.001$

Discussion

- High-gamma cortical responses are stronger to high freq. envelope than to carrier.
- High-gamma cortical responses are strongest to low-pitch portions of speech (≲100 Hz).
- Significant cortical responses exist for both attended and unattended speech streams, consistent with low-level (pre-attentive) representations from subcortical regions.
- The magnitude of these responses is modulated by selective attention, indicating top-down involvement in processing within primary auditory cortex.

References

[1] Kulasingham et al. (2021), J Neurosci.
[2] Simon et al. (2022), Front. Neurosci.
[3] Yang et al. (1992), Auditory IEEE Trans. Inf. Theory
[4] David et al. (2007), Network: Computation in Neural Systems
[5] Brodbeck et al. (2021), biorxiv
[6] Smith & Nichols (2009), Neuroimage