High Gamma Time-Locked Cortical Responses to Continuous Speech

Joshua P. Kulasingham[1], Vrishab Commuri[1] & Jonathan Z. Simon[1,2,3]

[1]Department of Electrical and Computer Engineering, [2]Institute for Systems Research, [3]Department of Biology, University of Maryland, College Park, Maryland

Introduction

- Recent studies have found cortical responses in the high gamma range (70-200 Hz) to continuous speech in MEG (Hertrich et. al.2009, Kulasingham et. al. 2020), and EEG (Canneyt et. al. 2021).
- This is similar to cortical Frequency Following Responses (FFRs) (Coffey et. al. 2016) which may originate from the thalamorecipient layers of cortex (Gnanateja et. al. 2021).
- Attentional modulation has been found for both high frequency cortical FFRs (Hartmann et. al. 2019) and low frequency cortical pitch tracking (Brodbeck and SImon 2022).

Here, we report our previous work on high gamma MEG Temporal Response Functions (TRFs) to continuous speech (Kulasingham et. al. 2020), and extend it to investigate high gamma TRFs to:

- Male vs. female speech
- Attended vs. unattended speech in a cocktail party paradigm

Methods

Study #1 (Kulasingham et. al. 2020)

- MEG responses of 17 younger and 23 older subjects listening to continuous speech by a male speaker.
- Previously published dataset - Presacco et. al. 2016a, 2016b, Kuchinsky et.al. 2017.

Study #2

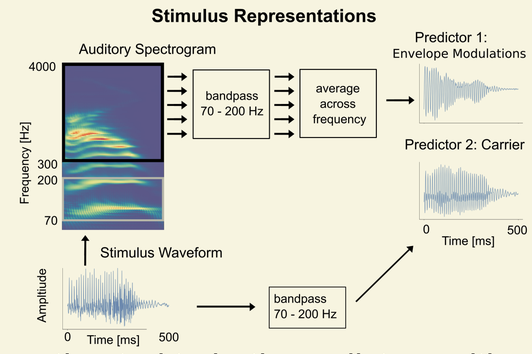
- MEG responses of 22 younger subjects listening to simple, continuous, rhythmic sentences and equations spoken by male and female speakers.
- Previously published dataset - Kulasingham et. al. 2021
- Subjects attended to one speaker during cocktail party conditions

Preprocessing and Source Localization

- MEG data was bandpassed 70-200 Hz
- Source localization using MNE onto a volume source space of 7mm grid spacing
- Temporal lobe and subcortical ROIs

TRF Estimation

- High frequency envelope modulations and carrier were used as predictors.



- TRFs were jointly estimated to both predictors with boosting (David et al. 2007) using Eelbrain (Brodbeck et. al. 2021).

$$\mathbf{y} = \boldsymbol{\tau}_e * \mathbf{e} + \boldsymbol{\tau}_c * \mathbf{c} + \mathbf{n}$$

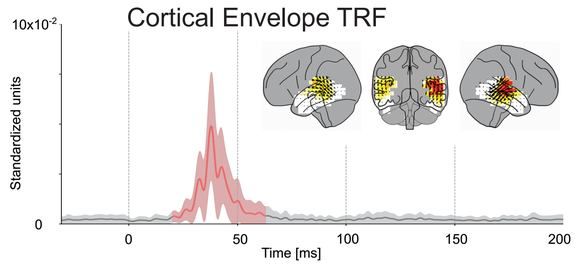
$\mathbf{y}$: MEG response, $\boldsymbol{\tau}_e$: env. TRF, $\boldsymbol{\tau}_c$: carrier TRF, $\mathbf{e}$: envelope, $\mathbf{c}$: carrier, $\mathbf{n}$: noise

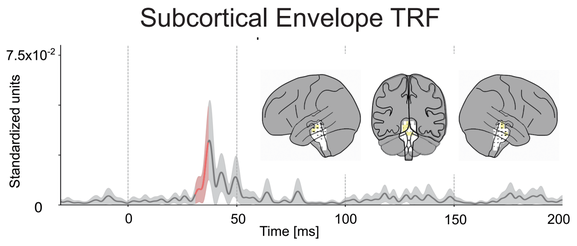
- Statistical tests with null models from circularly shifted predictors using TFCE.

Results - Study #1

Cortical Origin of High Gamma TRFs

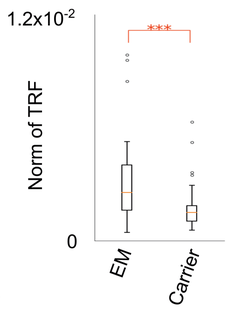
- Cortical ROI has stronger responses compared to subcortical ROI
- TRF responses are right lateralized, agreeing with cortical FFR studies

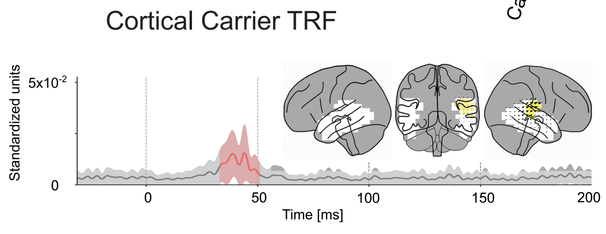


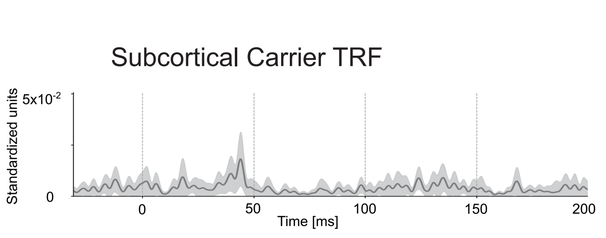


TRFs to speech envelope vs. carrier

- Envelope TRFs are signficantly stronger than carrier TRFs

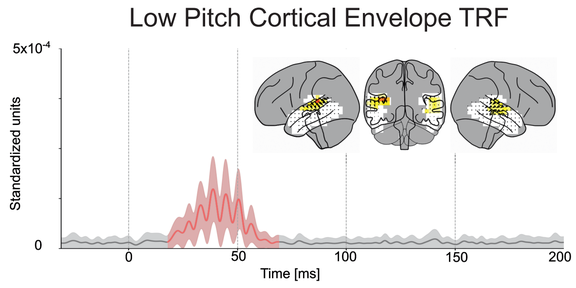


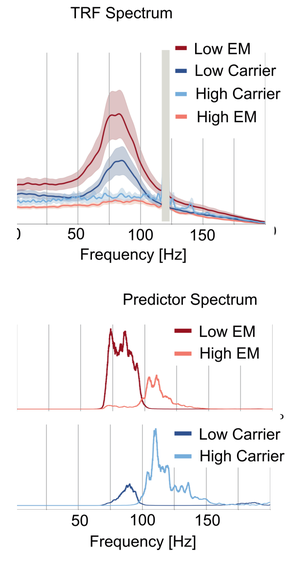


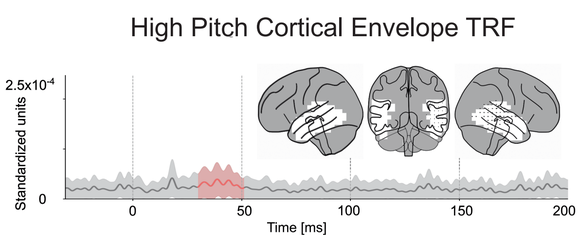


TRFs to low pitch vs. high pitch segments

- Speech seperated into low vs. high pitch segments around midpoint of 98 Hz
- TRFs jointly estimated for high and low pitch segments
- Low pitch TRFs are much stronger than high pitch TRFs
- TRF frequency spectrum peaks at 80-90 Hz



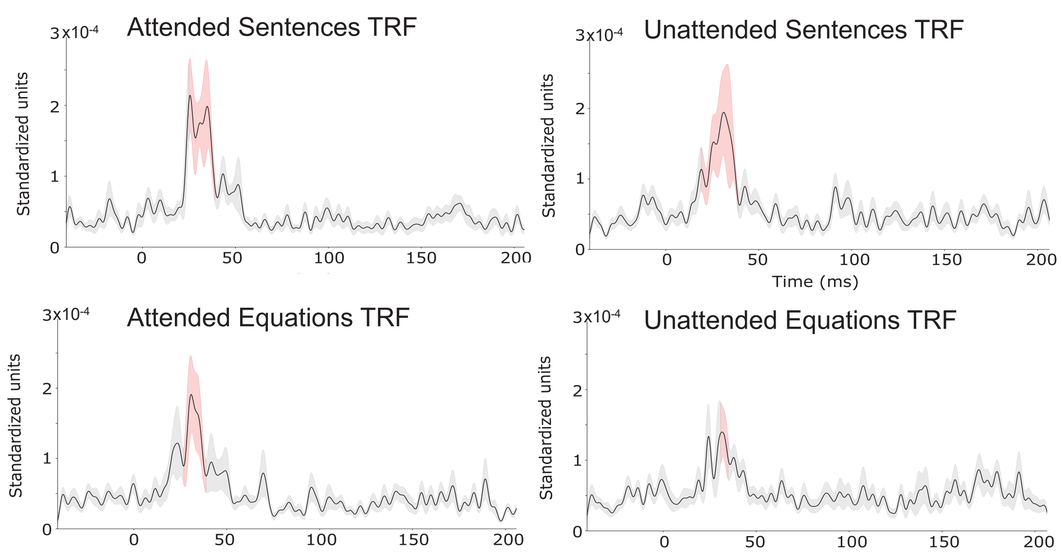




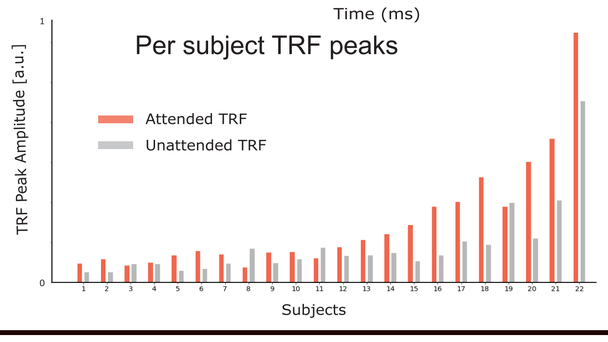
Results - Study #2

Attentional Modulation of High Gamma TRFs

- Attentional effects were investigated in the high gamma TRFs to the envelope in the cortical ROI for subjects attending to the male speaker.
- Significant high gamma TRFs are present for both attended and unattended speech

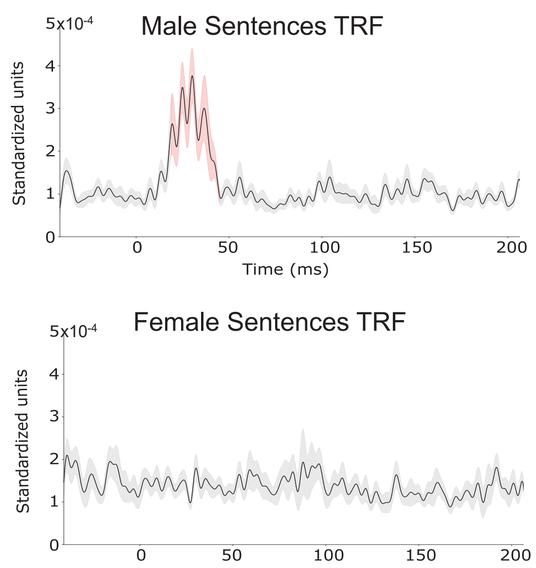


- TRFs were averaged across sentences and equations and peaks in the 20-50 ms range were extracted per subject
- TRF peak amplitudes are significantly larger for attended compared to unattended (paired t-test $t_{21} = 3.85$, $p < 0.001$)



High Gamma TRFs to male vs. female speakers

- High gamma TRFs to the envelope in the cortical ROI of subjects listening to a male speaker (top) or a female speaker (bottom)
- Significant high gamma TRFs only for the male speaker



Discussion

- This work investigates high gamma cortical responses to continuous speech using MEG.
- High gamma responses to the high frequency modulations of the speech envelope are stronger than to the carrier.
- High gamma responses are strongest to the low pitch segments of speech below 100 Hz.
- Accordingly, high gamma responses are present for male speech, but not for female speech.
- High gamma responses exist for both attended and unattended speech, consistent with low-level pre-attentive processes.
- TRF peak amplitudes are modulated by attention, perhaps indicating top-down attentional mechanisms.
- Further work must be done to investigate these early (<50 ms) attention modulated responses.

References

Kulasingham, J.P., et. al. High gamma cortical processing of continuous speech in younger and older listeners. NeuroImage (2020)

Presacco A., et. al. Effect of informational content of noise on speech representation in the aging midbrain and cortex. Journal of Neurophysiology (2016)

Presacco A., et. al. Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. Journal of Neurophysiology (2016)

Kuchinsky, S., et al. "Auditory-Cognitive Training Alters the Neural Encoding of Speech for Older Adults with Normal Hearing" ARO (2017)

Kulasingham, J.P., et. al. Cortical Processing of Arithmetic and Simple Sentences in an Auditory Attention Task. Journal of Neuroscience (2021)

Hertrich, I., et. al., Magnetic brain activity phase-locked to the envelope, the syllable onsets, and the fundamental frequency of a perceived speech signal. Psychophysiology (2012)

Van Canneyt, J., Neural tracking of the fundamental frequency of the voice: The effect of voice characteristics. European Journal of Neuroscience (2021)

Coffey, E.B.J., et. al., Cortical contributions to the auditory frequency-following response revealed by MEG. Nature Communications (2016)

Gnanateja, G.N., et. al., Frequency-Following Responses to Speech Sounds Are Highly Conserved across Species and Contain Cortical Contributions. eNeuro (2021)

Brodbeck, C., Simon, J.Z., Cortical tracking of voice pitch in the presence of multiple speakers depends on selective attention. biorxiv (2021)

Kegler, M., et. al., The Neural Response at the Fundamental Frequency of Speech is Modulated by Word-level Acoustic and Linguistic Information. biorxiv (2022)

David, S.V., et. al., Estimating sparse spectro-temporal receptive fields with natural stimuli. Network: Computation in Neural Systems (2007)

Brodbeck, C., et. al., Eelbrain: A Python toolkit for time-continuous analysis with temporal response functions. biorxiv (2021)

This work was supported by the National Institute on Deafness and Other Communication Disorders (R01 DC-014085); the National Institute of Aging (P01 AG-055365); and the National Science Foundation (SMA1734892).