High Frequency Cortical Processing of Continuous Speech in Younger and Older Listeners

Joshua P. Kulasingham[1], Christian Brodbeck[2], Alessandro Presacco[2] , Stefanie E. Kuchinsky[5], Samira Anderson[3] & Jonathan Z. Simon[1,2,4]

[1]Department of Electrical and Computer Engineering, [2]Institute for Systems Research, [3]Department of Hearing and Speech Sciences, [4]Department of Biology, University of Maryland, College Park, Maryland,
[5]Audiology and Speech Pathology Center, Walter Reed National Military Medical Center

Introduction

The neural processing of natural sounds, such as speech, changes along the ascending auditory pathway, and is often characterized by a progressive reduction in representative frequencies. This is observed in two well known neural responses:

1. **The Frequency Following Response (FFR):**
   – Thought to originate in the auditory brainstem and midbrain.
   – Frequencies of ~100 Hz to several hundred Hz.
   – Time-locks to fast acoustic envelope and waveform.
2. **The cortical low frequency Temporal Response Function (TRF):**
   – Frequencies of around 1-20 Hz.
   – Time-locks to slow acoustic envelope of continuous speech.

Recent studies have shown that the FFR is driven by cortical sources in addition to subcortical sources (Coffey et al. 2016).

However it is unclear if this is due to the MEG being biased towards cortical sources (Bidelman 2018).

Recent studies have also shown a continuous speech generated ABR-like response in brainstem (Maddox and Lee 2018) and that such responses are modulated by attention (Forte et al. 2017).

High frequency responses to speech have also been found in MEG (Hertrich et. al.2009)

Age-related differences, have been observed for FFR and the slow cortical TRF, of opposite direction:
– FFR responses are stronger in younger subjects (Anderson et al. 2012).
– Cortical low frequency responses are stronger in older adults (Brodbeck et al. 2018, Decruy et al. 2019).

How do cortical and subcortical areas each contribute to high frequency responses to speech?

Are these responses to the envelope or to the carrier?

Are high frequency responses in older listeners overrepresented as in the low frequency TRF or weakened as in the EEG FFR?

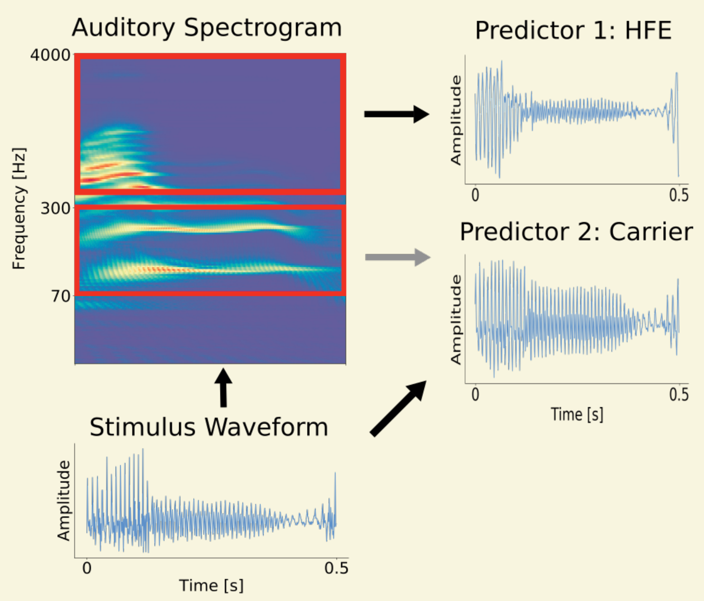
Methods

MEG data was collected from 17 younger and 23 older subjects while they listened to 6 minutes of continuous speech. Earlier analyses of the data were published in Presacco et. al. 2016a, 2016b, Kuchinsky et.al. 2017.

Stimulus Representation

Two separate high frequency aspects of the speech stimuli were considered.

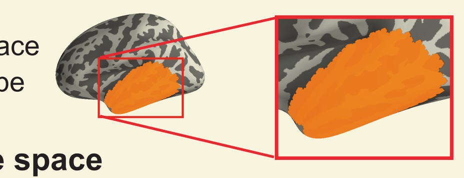
1. **Carrier:** speech waveform bandpassed 70-300 Hz.
2. **High Frequency Envelope (HFE)**: The 300-4000 Hz envelope of the speech, bandpassed 70-300 Hz:
   – Compute the envelope of the auditory spectrogram (Yang et al. 1991).
   – Filter the 300-4000 Hz components of the spectrogram envelope at 70-300 Hz.
   – Average those filtered spectrogram envelope components across frequency.



Source Localization

The 157 MEG sensors were transformed to current dipole sources distributed on the brain using MNE source localization (Gramfort et al. 2014).

Surface Source Space
ROI: Temporal Lobe



1. Volume source space
– Brain volume is divided into a 7 mm voxel grid.
– Current dipole vector with 3D orientation and magnitude placed at each voxel.
– ROIs: voxels closest to a) the temporal lobe and b) the brainstem.

2. Surface source space
– Neural sources are distributed on the white matter surface of cortex.
– Current dipoles fixed at an orientation perpendicular to the cortical surface.
– ROI: neural sources in the temporal lobe.

TRF Estimation and Statistical Tests

Joint Estimation of Carrier and HFE TRFs
Envelope and waveform TRFs were estimated with the boosting algorithm (David et al. 2007) using Eelbrain (Brodbeck 2018).

$$r^s = \tau^s_{hfe} * hfe + \tau^s_{carr} * carr$$

Statistical Tests
– The TRF model was compared to noise models formed by circularly time-shifting each predictor.
– Pearson correlation coefficients between the actual response and the predicted response were used as a measure of model fit. t-values were used with TFCE (Smith and Nichols 2009) to test for significant increase in model fit of the true model over the noise model.
– The norm of the TRF vectors of the volume source space and the rectified TRFs of the surface source space were spatially smoothed using a Gaussian window of $5mm^2$ / $5mm^3$ and tested for significance between the true model and the noise model.

Acknowledgments

This work was funded 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).

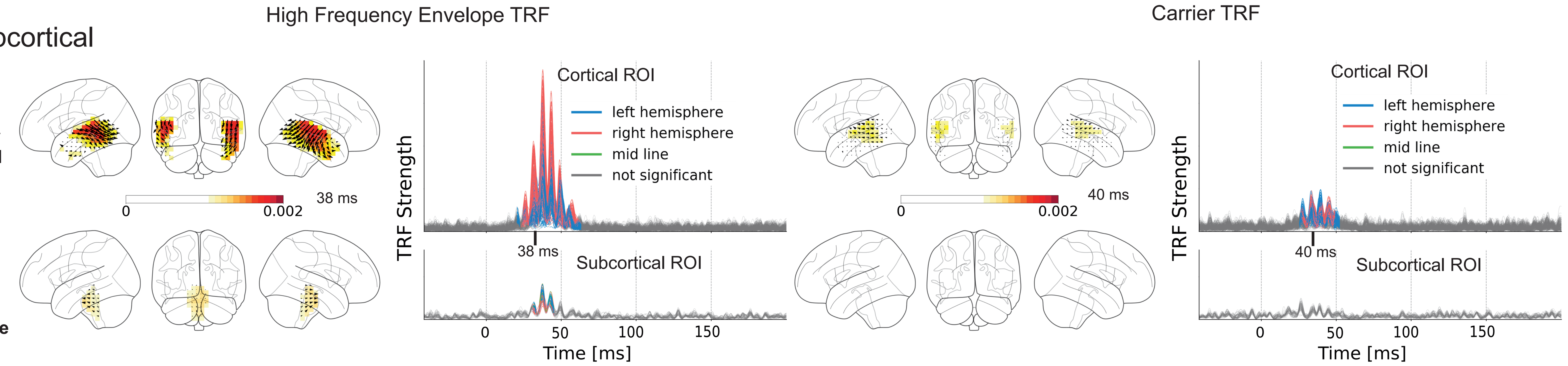
Poster and preprint available at http://ter.ps/simonpubs

Results

Origin: Cortical vs Subcortical

TRF estimation using volume source space. Significant TRFs observed for all relevant voxels. The subcortical responses appear as artifactual leakage from cortical sources:
– low amplitudes with cortical latencies (~40 ms).
– consistent with simulations of purely cortical sources.

TRF peak latency of 40 ms and spatial location strongly indicate a purely cortical source



Cortical Responses: Younger vs. Older

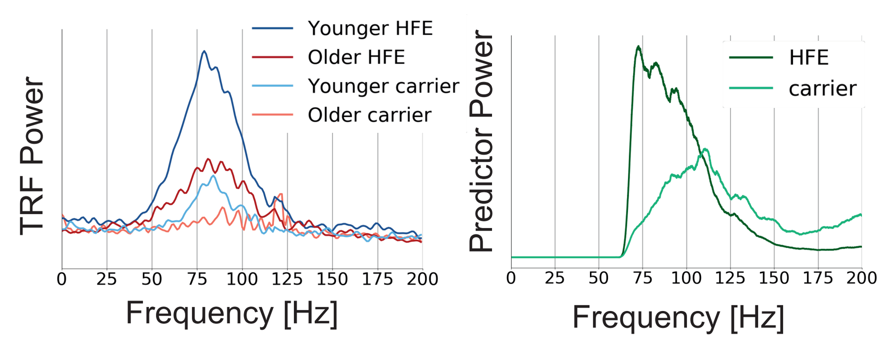
Surface source space data was analyzed for a ROI that included the temporal lobe.

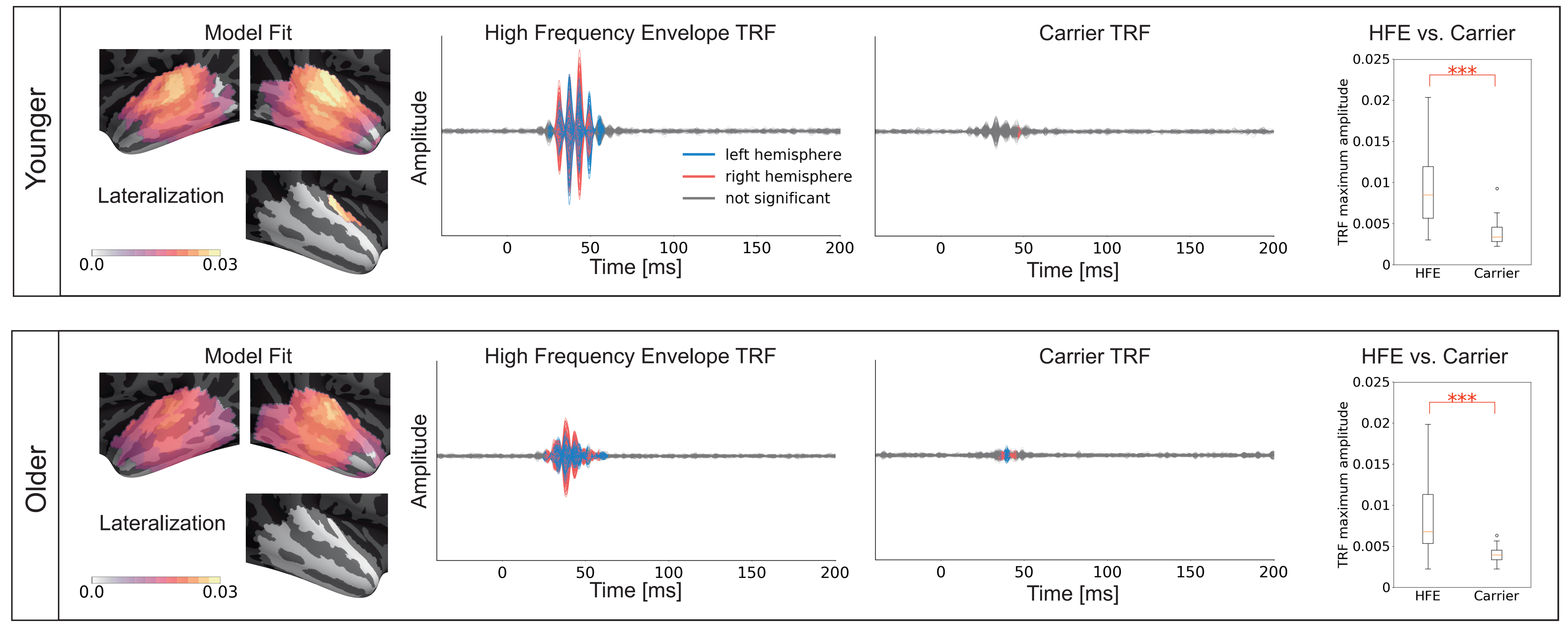
Predictive power is **right lateralized** in younger but not in older subjects. However the difference in lateralization is not significant.

There were **no significant differences** between age groups.
– ANOVA on low frequency and high frequency TRFs. Significant main effect of TRF frequency (F(1,38) = 125.94, p < 0.001) and **TRF frequency x age** (F(1,38) = 7.52, p = 0.009) but no main effect of age (F(1,38) = 2.26, p = 0.141).

The response is **dominated by the HFE predictor.**

TRFs have a **spectral peak at ~80 Hz.**





No significant differences between age groups. Significantly right lateralized in early auditory cortex only for younger subjects.

Discussion & Conclusions

This work confirms the existence of high frequency cortical responses to continuous speech that are measurable using MEG.

Peak latency (40 ms), source localization, and simulations of cortical source leakage, strongly indicate a purely cortical origin for this response. This suggests that MEG and EEG are sensitive to different structures along the auditory pathway

The response is predominantly to the high frequency envelope of the stimulus and not to the carrier.

The responses are right lateralized in younger subjects, which agrees with previous work on MEG FFR (Coffey et. al. 2016).

Surprisingly, there are no age related differences in response magnitude, or latency. This is unlike both the FFR, which is weaker for older adults, and the low frequency cortical response to speech, which is stronger for older adults.

The responses oscillate around 80 Hz, although the stimulus has a broad spectrum around 70-120 Hz. This could reflect an intrinsic bias in cortical responses.

Electrocorticography (ECoG) studies have seen cortical phase-locked activity at these high rates (Nourski et al., 2014; Steinschneider et al., 2013). However, such rates are rarely seen with MEG.

These responses could reflect cortical input from the Medial Geniculate Body (MGB), since the frequency of these responses is consistent with the intermediate rate of thalamic auditory neurons (faster than cortical; slower than midbrain) (Miller et. al. 2002).

These fast responses might provide the substrate that allows precise spike timing (a few ms) in primary auditory cortex (Elhilali et. al. 2004).

EEG and MEG can be used as complementary techniques to analyze the processing of natural sounds along the auditory pathway.

Both the neural origin and the frequency domain must be considered when investigating age related changes in the auditory system

References

Anderson, S. et al. 2012. *Aging affects neural precision of speech encoding.* Journal of Neuroscience 32.41: 14156-14164.

Bidelman, G.M. 2018. *Subcortical sources dominate the neuroelectric auditory frequency-following response to speech.* NeuroImage 175, 56–69.

Brodbeck, C. 2019. : Eelbrain 0.30. Zenodo, 2019.

Brodbeck, C. et al. 2018. *Over-representation of speech in older adults originates from early response in higher order auditory cortex.* Acta Acust. United Acust. 104.5: 774-777.

Coffey, E. B. J. et al. 2016. *Cortical contributions to the auditory frequency-following response revealed by MEG.* Nature communications 7: 11070.

David, S. V. et al. 2007. *Estimating sparse spectro-temporal receptive fields with natural stimuli.* Network: Computation in Neural Systems 18.3: 191-212.

Elhilali, M. et al. 2004. *Dynamics of Precise Spike Timing in Primary Auditory Cortex.* Journal of Neuroscience 24, 1159–1172.

Forte, A. E., et al. 2017. *The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention.* elife 6: e27203.

Gramfort, A. et al. 2014. *MNE software for processing MEG and EEG data*. Neuroimage 86: 446-460.

Hertrich, I. et al. , 2012. *Magnetic brain activity phase-locked to the envelope, the syllable onsets, and the fundamental frequency of a perceived speech signal.* Psychophysiology 49, 322–334.

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

Maddox, R. K. et al. 2018 *Auditory brainstem responses to continuous natural speech in human listeners.* ENeuro 5.1.

Miller, L.M. et al. 2002. *Spectrotemporal Receptive Fields in the Lemniscal Auditory Thalamus and Cortex.* Journal of Neurophysiology 87, 516–527.

Nourski, K.V. et al. 2014. *Functional organization of human auditory cortex: Investigation of response latencies through direct recordings.* NeuroImage 101, 598–609.

Presacco, A. et al. 2016a. *Effect of informational content of noise on speech representation in the aging midbrain and cortex.* Journal of Neurophysiology 116.5: 2356-2367.

Presacco, A. et al. 2016b. *Evidence of degraded representation of speech in noise, in the aging midbrain and cortex*. Journal of Neurophysiology 116.5: 2346-2355.

Smith, S.M. et al. 2009. *Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference.* NeuroImage 44, 83-98.

Steinschneider, M. et al. 2013. *Representation of speech in human auditory cortex: Is it special?* Hearing Research 305.

Yang, X. et al. 1992. *Auditory representations of acoustic signals.* IEEE Trans. Inf. Theory 38, 824–839.