High Frequency Time-Locking in Human Auditory Cortex to Continuous Speech

Joshua P. Kulasingham1, Christian Brodbeck2, Alessandro Presacco3, Stefanie E. Kuchinsky4, Samira Anderson3 & Jonathan Z. Simon1,2,4

1Department of Electrical and Computer Engineering, 2Institute for Systems Research, 3Department of Hearing and Speech Sciences, 4Department of Biology, University of Maryland, College Park, Maryland, 5Auditory 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-lags to fast acoustic envelope and waveform.

2. The cortical low frequency Temporal Response Function (TRF): – Frequencies of ~1-20 Hz. – Time-lags 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 et al. 2016). Recent studies have also shown a common speech-gated ABR-like response in brainstem (Maddox and Lee 2016) and that such responses are modulated by attention (Port et al. 2017).

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

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

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 bandpass 70-300 Hz.
2. High Frequency Envelope (HFE): The 300-4000 Hz envelope of the speech, bandpass 70-300 Hz.

– Compute the envelope of the spectrogram (Yang et al. 1991).
– Filter the 300-4000 Hz components of the spectrogram envelope at 70-300 Hz.
– Average these 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).

1. Volume source space – The brain volume is divided into 7 mm^3 voxels. – Current dipole vector with 3D orientation and magnitude placed at each voxel. – ROIs: voxels closest to the temporal lobe and/or the brainstem.

TRF Estimation and Statistical Tests

Joint Estimation of Carrier and HFE TRFs

Envelope and waveform TRFs were estimated with the boosting Joint Estimation of Carrier and HFE TRFs (2).

TRF estimation using volume source space. Significant TRFs were observed for all relevant voxels, but the subcortical responses appear as artifactual leakage from cortical sources. low amplitudes with cortical latencies.

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

TRF Power

– ROI: neural sources in the temporal lobe.
– ROI: Temporal Lobe

Cortical Responses: Younger vs. Older

Surface source space data was analyzed for a ROI that included the temporal lobe. TRF analysis reveals significant predictive power in younger and older groups for neural sources in 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.

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.

Both peak latency (40 ms) and source localization 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 the FFR, which is weaker for older adults, and the low frequency cortical responses 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.

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 and the frequency domain must be considered when investigating age related changes in the auditory system.

References


Acknowledgments

This work was funded by the National Institute of Health (NIH) grants R01 EY025985 and P30 AG032935. Poster available at http://psych.umaryland.edu.