Cortical Over-representation of Speech in Older Listeners Correlates with a Reduction in both Behavioral Inhibition and Speech Intelligibility

Peng Zan[1], Alessandro Presacco[2], 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, MD

Introduction

Aging is associated with an exaggerated representation of the speech envelope in the auditory cortex[1][2]. However, whether this over-representation is related to decreased speech intelligibility for older listeners is an open question. Source space analysis has shown the over-representation originates (at least) from early responses (~50 ms latency) in the auditory cortex[3].

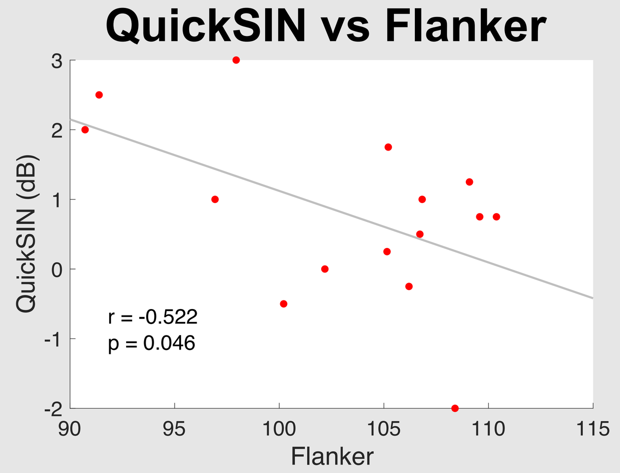
- The abnormally strong response to low-frequency speech envelope in older listeners may be related to speech processing problems
- Reanalysis of earlier experiment[1][2] using mutual information

Methods

Experiments

- 1-min speech segment (male speaker), both clean and masked with a competing female speaker, presented for 4 trials for each condition (quiet and 4 SNRs: 3, 0, -3 and -6 dB)
- Neural responses to continuous speech recorded by magnetoencephalography (MEG) at sampling frequency 1 kHz
- 17 younger adults (age: 18-27) and 15 older adults (age: 61-73), native English speakers with clinically normal hearing
- Visual inhibitory ability measured by Flanker score (higher is better)
- Listeners' ability to understand speech in noise measured by QuickSIN test (quick speech in noise, higher is worse)
- Weak negative correlation between the two behavioral scores seen for older listeners

QuickSIN vs Flanker



Temporal Mutual Information Function (TMIF)

- MEG recording denoised by TSPCA[4], and the first DSS component (1-8 Hz) extracted[5] from MEG signal as auditory response
- Low-frequency (1-8 Hz) speech envelope extracted
- Both response and speech envelope binned into 8 bins based on amplitude
- Temporal Mutual Information Function (TMIF) estimated by mutual information between speech envelope and response delayed at different time points
- At one time point t, the mutual information is estimated by

$$I_t(X;Y) = \sum_{i \in S, j \in S} p(x(\tau) = i, y(\tau + t) = j) \log \frac{p(x(\tau) = i, y(\tau + t) = j)}{p(x(\tau) = i) p(y(\tau + t) = j)}$$

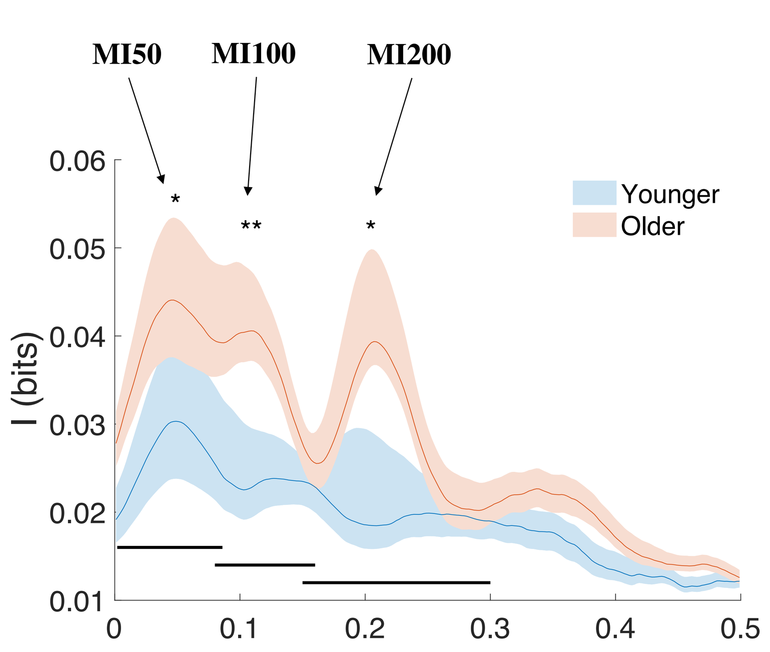
- $S = \{1,2,\ldots,8\}$, the set of amplitude bins from which i, j are drawn
- X and Y: random variables denoting stimulus and response. The joint probability distribution of X and Y estimated by amplitude of speech envelope and shifted response

TMIF in neural source space

- Neural source space response via minimum norm estimation[6] (MNE)
- TMIF of each neural source is estimated

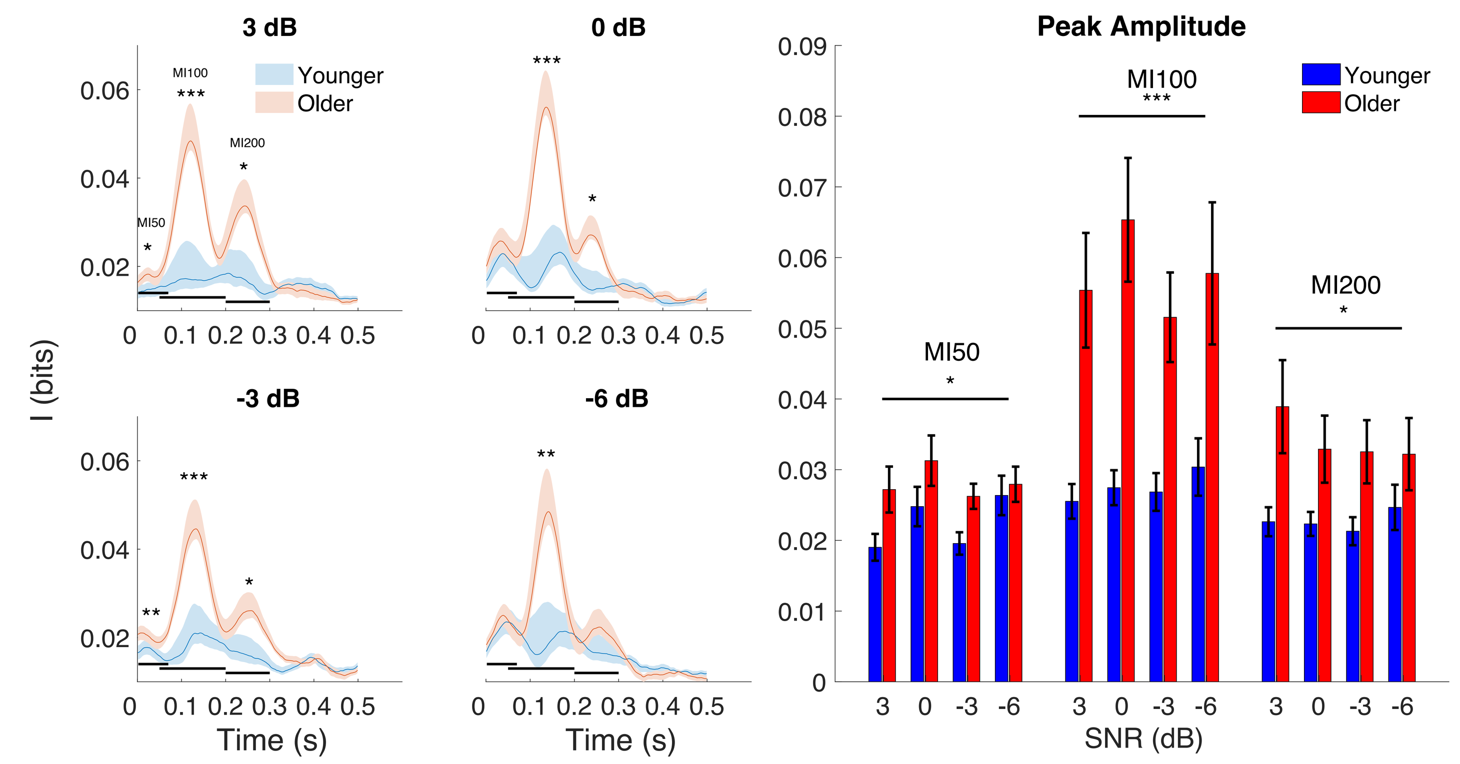
Results

TMIF of speech in quiet



- TMIF = information theory analog of linear Temporal Response Function (TRF)
- Amplitude response in older listeners is larger than younger listeners in quiet
- Older listeners have higher MI50, MI100 and MI200
- For older listeners, over-representation not only occurs early, but occurs for three different latencies
- *p<.05, **p<.01, ***p<0.001

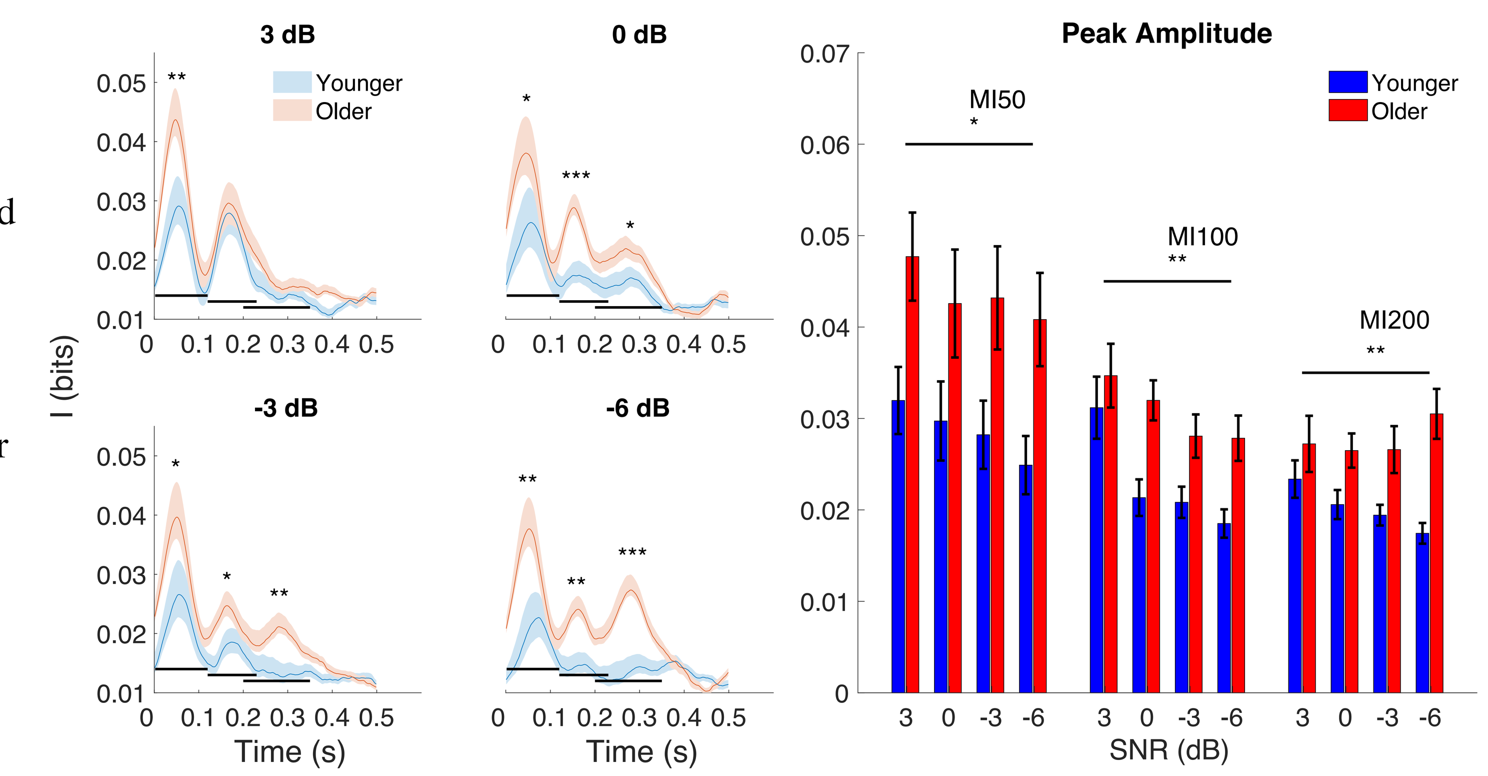
TMIF of foreground speech in noisy conditions



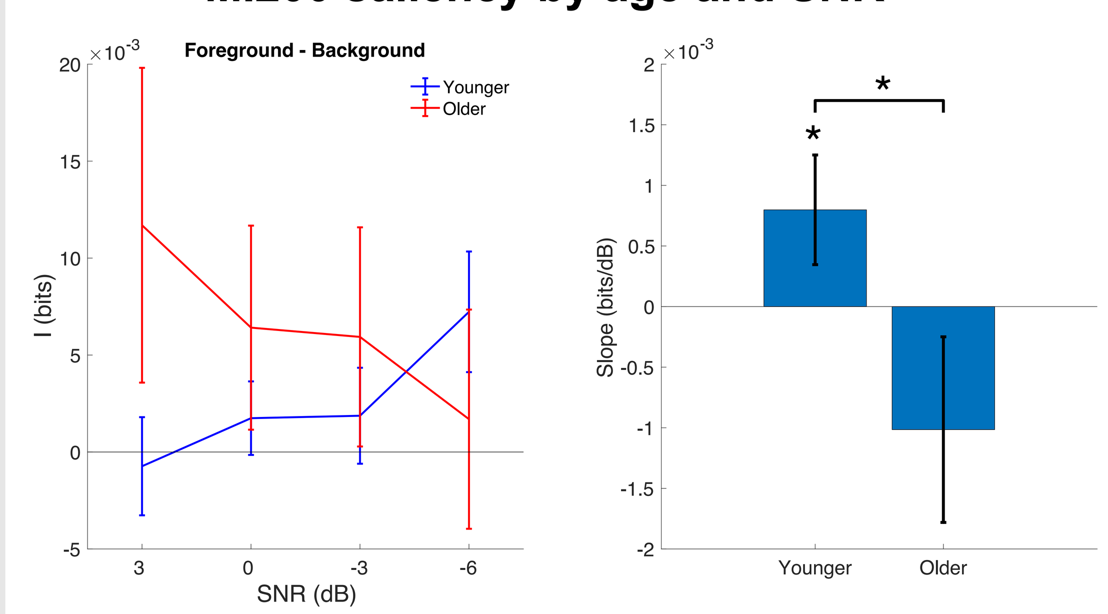
- The TMIFs of foreground contain larger peaks of MI50, MI100 and MI200 for older listeners measured by average across SNR conditions
- Over-representation at ~50, ~100 and ~200 ms also maintained for speech in noise

TMIF of background speech in noisy conditions

- Over-representation also occurs for background speech at ~50, ~100 and ~200 ms for older listeners measured by average across SNRs
- MI200 shows changes with worsening SNR: increasing for older listeners but decreasing for younger

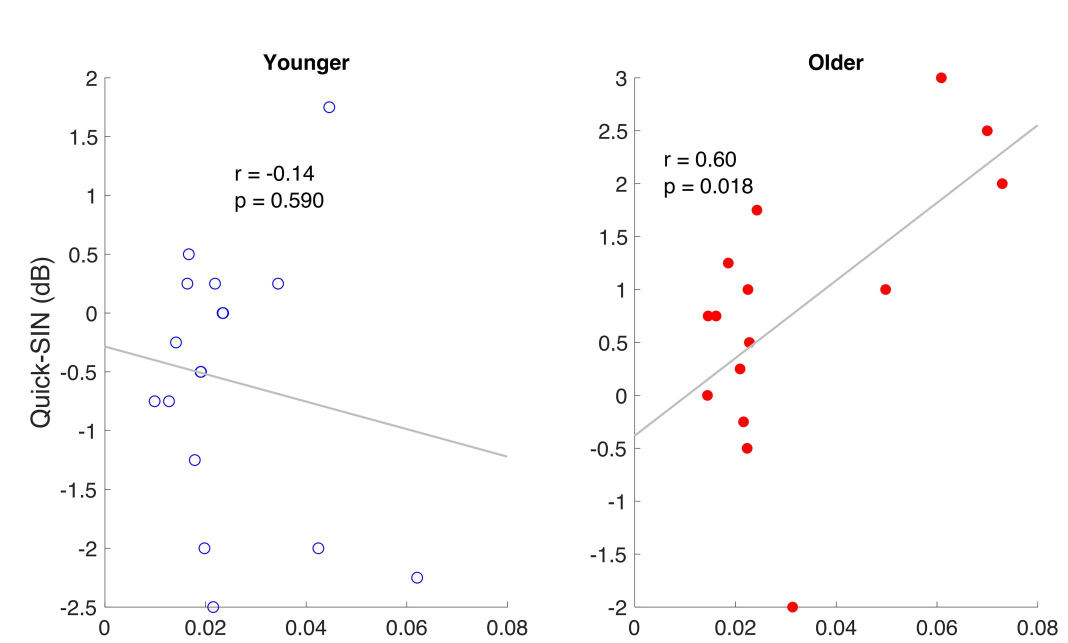


MI200 saliency by age and SNR

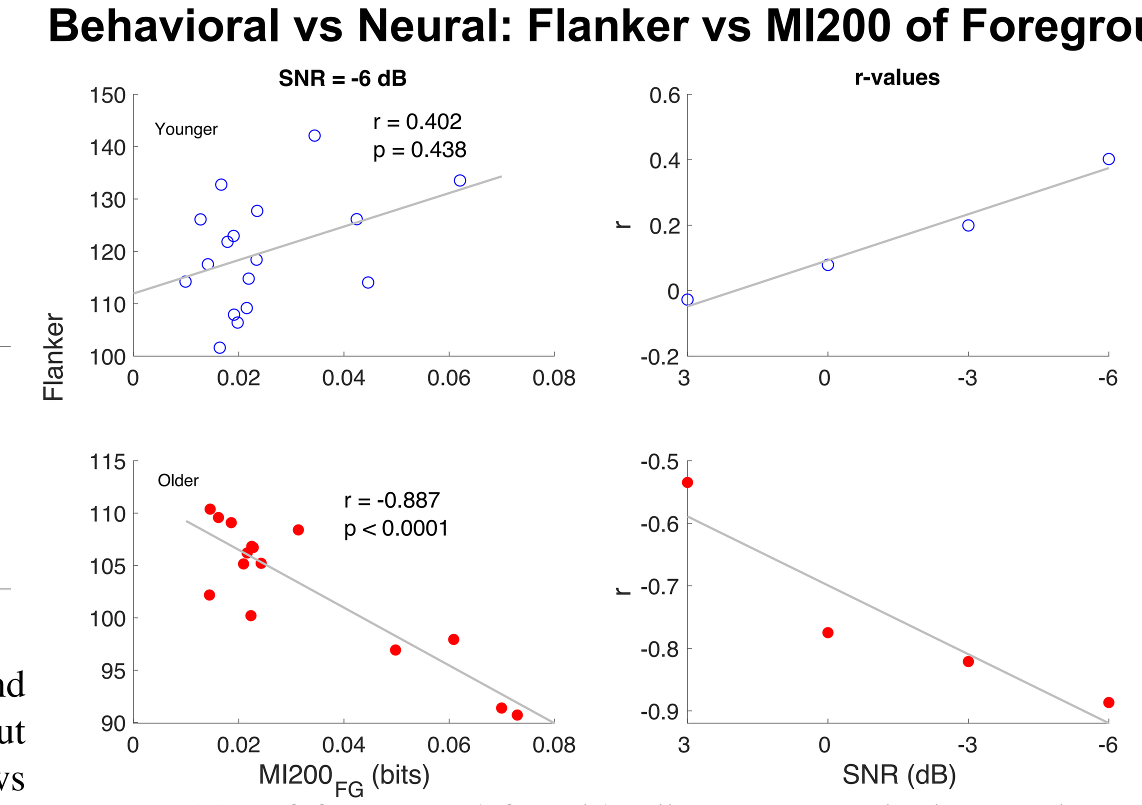


- MI200 saliency, measured by MI200 foreground and background difference, shows decrease as noise worsens for older listeners but increase for younger (asterisk above younger errorbar shows significance)
- The MI200-by-SNR slope is significantly larger for younger listeners than for older, illustrated by the top asterisk

QuickSIN vs MI200



Behavioral vs Neural: Flanker vs MI200 of Foreground

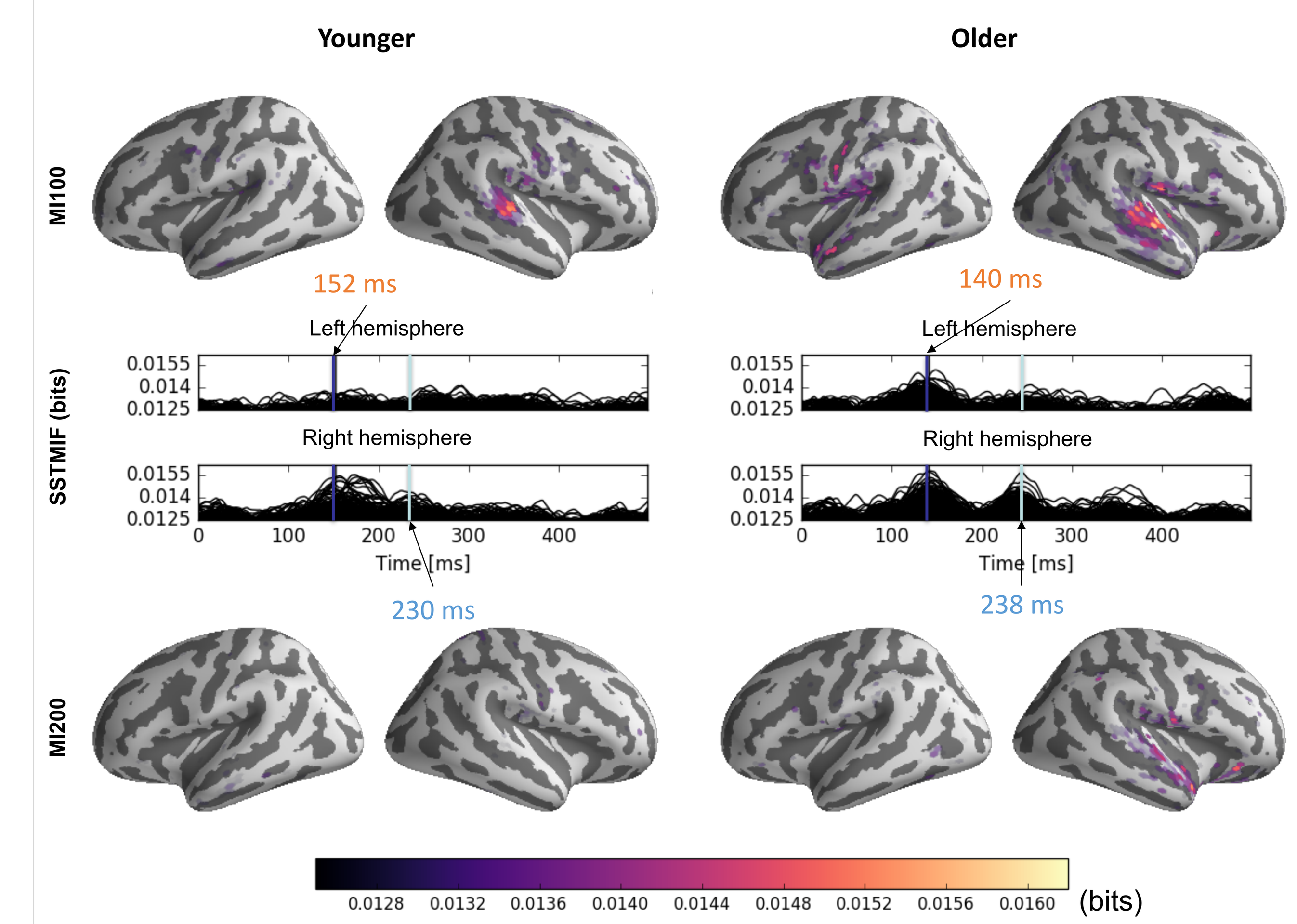


- MI200 of foreground for older listeners negatively correlates with Flanker score, and correlation grows more negative with more noise
- Even though flanker test measures behavioral visual inhibition, it reflects domain-general inhibition and correlates strongly with an auditory neural biomarker

- MI200 for older listeners is positively correlated to QuickSIN speech intelligibility score, while no correlation is seen for younger listeners
- Larger MI200 corresponds to worse speech intelligibility for older listeners
- Linear mixed effect model of MI200 ~ Flanker * QuickSIN shows significant effect from QuickSIN after ruling out interaction from Flanker

For mutual information analysis of midbrain response, see poster PS 236.

TMIF in neural source space



- Older listeners show a right-lateralized response in auditory cortex for MI200, while no significant response is seen for younger listeners
- Younger listeners show right-lateralized MI100 response in auditory cortex, while older listeners' response is bilateral
- Neural sources for MI200 localize to auditory cortex (despite correlation with visual Flanker score)

Conclusions

- An over-representation of low-frequency speech envelope is observed for older listeners illustrated by peaks in TMIF.
- At ~100 ms latency, older listeners engage additional areas (e.g., left hemisphere) over younger listeners; at ~200 ms latency, older listeners show new response (MI200) (dominantly right hemisphere) not shown by younger listeners.
- The over-representation in older listeners may be due to the loss of cortical synaptic inhibition, exaggerated attentional efforts, and processing of contextual or redundant speech information.
- The neural mechanism behind the exaggerated information representation may relate to the loss of behavioral inhibitory control, which affects speech intelligibility in challenging environments for older listeners.

Acknowledgments

This work was funded by the National Institutes of Health (R01-DC014085 and P01-AG055365). We thank support for Peng Zan from the NSF funded COMBINE program.

References

[1] Presacco A, Simon JZ, Anderson S (2016) ***Evidence of degraded representations of speech in noise in the aging midbrain and cortex***, J Neurophysiol 116(5): 2346-2355

[2] Presacco A, Simon JZ, Anderson S (2016) ***Effect of informational content of noise on speech representation in the aging midbrain and cortex***, J Neurophysiol 116(5): 2356-2367

[3] Brodbeck C., Presacco A., Anderson, S., Simon J.Z., (2018) ***Over-Representation of Speech in Older Adults Originates from Early Response in Higher Order Auditory Cortex***, Acta Acust. United Acust. 104, 774–777.

[4] de Cheveigné A., Roux J.L., Simon J.Z. (2007) ***MEG Signal Denoising Based on Time-Shift PCA, 2007 IEEE International Conference on Acoustics, Speech and Signal Processing***, ICASSP '07, pp. I-317-I–320.

[5] de Cheveigné A., Simon J.Z. (2008) ***Denoising based on spatial filtering***, J. Neurosci. Methods 171, 331–339.

[6] Gramfort A., Luessi M., Larson E., Engemann D.A., Strohmeier D., Brodbeck C., Goj R., Jas M., Brooks T., Parkkonen L., Hämäläinen M. (2013) ***MEG and EEG data analysis with MNE-Python***, Front. Neurosci. 7.

Find poster at http://ter.ps/simonpubs or