Effects of Aging on the Cortical Representation of Continuous Speech

I.M Dushyanthi Karunathilake[1], Jason L. Dunlap[2], Janani Perera[2], Alessandro Presacco[3], Lien Decruy[3], Samira Anderson[2], Stefanie E. Kuchinsky[4], Jonathan Z. Simon[1,3,5]

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

INTRODUCTION

Older adults often report difficulty in understanding speech in noise[1]. These difficulties may arise from age-related physiological changes and temporal processing deficits. Previous studies using M/EEG have reported that aging is associated with exaggerated representation of speech envelope in the auditory cortex[2][3]. This robust representation may relate to

- Low level age-related changes e.g., excitation/inhibition imbalance
- Recruitment of additional top-down processing
- Decreased connectivity and redundant local processing

Motivation

To further investigate age-related neuro-physiological differences

- At what stages (latencies) do age-related processing differences occur?
- How does the task difficulty change the neural response?
- How are the foreground (FG) and background (BG) speakers represented neurally?

METHODS

Participants

- 18 younger (17-26 years) and 17 older (65-78 years) adults
- Native English speakers w/ normal hearing (125-4000 Hz ≤25 dB HL)

Task

- Listening to 1-min long speech segments from an audio book
  - Clean speech
  - Mixed speech (Male vs female speaker) [0 dB, -6 dB]
  - Babble speech (Female speaker vs 3 speaker babble) [0 dB]

Data

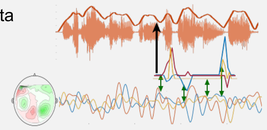
- MEG data

Analysis

- Denoising Source Separation (DSS)[4]
- Low frequency (1-10 Hz) log speech envelope and neural response
- Boosting algorithm with 10-fold cross validation[4]
- Statistical significance evaluated by Linear Mixed Effect Models in R

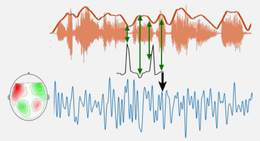
Stimulus reconstruction from response

- First 6 DSS components filtered data for the reconstruction
- Both foreground and background speaker envelopes reconstructed separately
- Reconstruction accuracy was estimated by the Pearson correlation coefficient between the true and reconstructed speech envelopes.



Temporal Response Function (TRF)

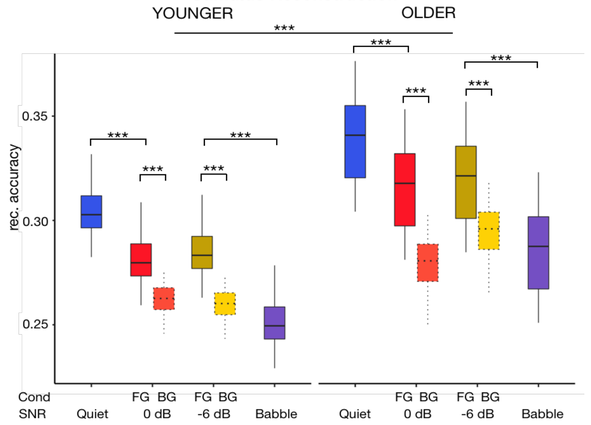
- Response prediction from stimulus
- First DSS component filtered data used as the auditory response
- TRF is estimated as the linear filter that transforms the speech envelope to the neural response



$$DSS_1(t) = \sum_{\tau} TRF_{FG}(\tau)ENV_{FG}(t-\tau) + TRF_{BG}(\tau)ENV_{BG}(t-\tau)$$

- TRF has 3 prominent peaks ~50 ms (M50) a positive peak, ~100 ms (M100) a negative peak and ~200 ms (M200) a positive peak

Stimulus Reconstruction



0.0001(***), 0.001(**), 0.01(*), 0.05(.)

Older adults exhibit better(!) stimulus reconstruction than younger adults

- Holds for all SNR levels and for both Foreground and background
- This could be due to age related changes e.g., excitation/ inhibition imbalance, recruitment of additional top-down resources and increased attention

Task difficulty significantly worsens foreground reconstruction in both groups

- Background noise significantly worsens the envelope representation in the cortical response

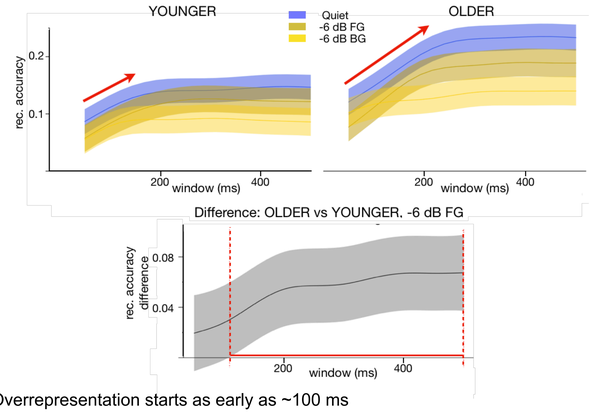
Foreground is better reconstructed than background in both groups

- Reconstruction accuracy is modulated by selective attention

At what latencies does overrepresentation occur? How long envelope is processed?

- Speech envelope reconstructed using 50-500 ms Integration windows
- Generalized Additive Mixed Models applied to the resulting time series data

Integration Window Analysis

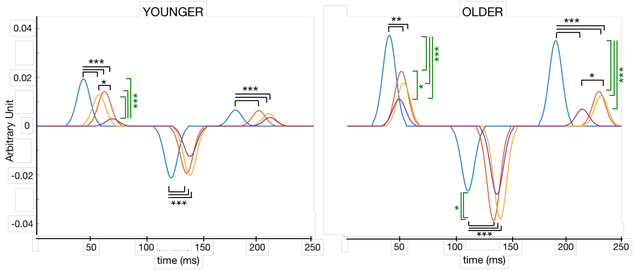


Overrepresentation starts as early as ~100 ms

Reconstruction takes more time for older adults

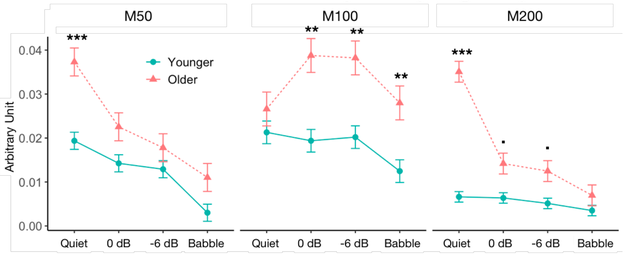
- Late processing ~200 ms to compensate the temporal processing deficits

Temporal Response Function – TRF (Foreground)



Amplitude and latencies are analyzed separately

TRF – Peak Amplitudes (Foreground)



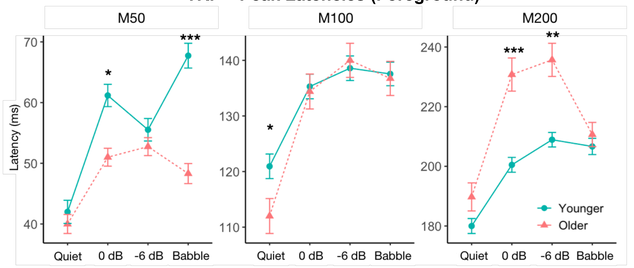
Older adults exhibit enlarged peak amplitudes

- M50 (Only quiet) : Excitation and inhibition imbalance
- M100 (except quiet) : Increased attention, left hemisphere recruitment[5]
- M200 : Recruitment of additional late resources

With the task difficulty,

- M50 decreases - M50 is shared between FG and BG
- In older adults M100 increases except babble - greater attention, giving up in babble
- In older adults M200 decreases - possibly a negative polarity source turning on

TRF – Peak Latencies (Foreground)

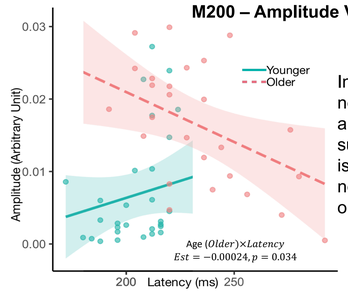


In both groups when the task gets harder, all peaks are typically delayed, harder the task, takes more time to process

M200 is significantly delayed in older adults except in quiet and babble. Late neural mechanisms are not involved in quiet, whereas for babble subjects may have given up the task.

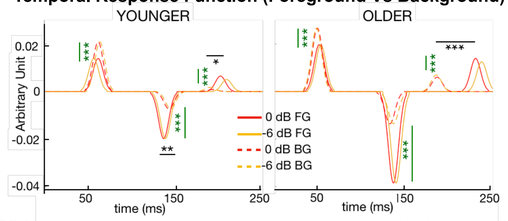
Age×SNR interaction effect indicated that peaks are delayed from quiet to 0 dB significantly more than the younger group for both M100 and M200

M200 – Amplitude Vs Latency



In older adults significant negative slope between the amplitude and latency may suggest that delayed M200 is modulated by a second neural mechanism with opposite polarity

Temporal Response Function (Foreground Vs Background)



FG is stronger compared to BG for both M100 and M200, suggesting middle and late peaks are modulated by attention

Late processing of the BG terminates before that of the FG

CONCLUSION

- Older adults' neural response robustly tracks the speech envelope, and to a greater extent than younger adults, possibly due to several mechanisms, e.g., excitation/ inhibition imbalance, recruitment of additional top-down resources, redundant local processing
- M200 peak is late enough to be modulated by many compensatory mechanisms
- Early activity, i.e., the M50, is not modulated by attention, while late activity, M100 and M200, is
- More difficult tasks produce longer latencies
- Altogether, despite impaired speech intelligibility in noise, time locked speech responses are exaggerated in older adults compared to younger

Acknowledgements

This work was supported by the National Institutes of Health grants P01-AG055365 and R01-DC014085.

References

[1]. Gordon-Salant, S., Fitzgibbons, P. J., & Yeni-Komshian, G. H. (2011). **Auditory temporal processing and aging: Implications for speech understanding of older people**. Audiology Research, 1(1S), e4.

[2]. Presacco, A., Simon, J. Z., & Anderson, S. (2016). **Evidence of degraded representation of speech in noise, in the aging midbrain and cortex**. Journal of Neurophysiology, 116(5), 2346–2355.

[3]. Decruy, L., Vanthornhout, J., & Francart, T. (2019). **Evidence for enhanced neural tracking of the speech envelope underlying age-related speech-in-noise difficulties**. Journal of Neurophysiology, 122(2), 601–615.

[4]. Ding, N., & Simon, J. Z. (2012). **Emergence of neural encoding of auditory objects while listening to competing speakers**. Proceedings of the National Academy of Sciences, 109(29), 11854–11859.

[5]. 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.