Influence of aging on cortical auditory temporal processing of speech in noise

*Alessandro Presacco[1,2], Samira Anderson[1,2] Jonathan Z. Simon[2,3,4,5],

[1]Hearing and Speech Sciences Department, [2]Neuroscience and Cognitive Science Program, [3]Biology Department,
[4]Institute for Systems Research, [5]Department of Electrical & Computer Engineering
University of Maryland, College Park, MD
*apresacc@umd.edu

Background

Older adults often report that during a conversation they can hear what is said, but cannot understand the meaning, particularly in a noisy environment. These difficulties may arise from deficits in auditory temporal processing [1]. Recent results using magnetoencephalography (MEG) [2,3] have shown the feasibility of reconstructing the envelope of speech in noisy conditions by using low frequency oscillations of the brain in younger adults. Although the effects of neural speech processing has been investigated in quiet conditions [4], little is known about how noise impacts cortical speech processing in younger vs. older adults. Here, we compared the effects of noise on cortical responses in younger and older adults with normal hearing, hypothesizing that in favorable conditions (SNR ≅0 dB) differences in performance between the two age groups will be mainly linked to the fact that younger adults are better than older adults at suppressing the competing speech signal.

Materials and Method

Participants

- Participants were native speakers of English: 8 young adults (20 – 28 years old, mean ± SD, 23.8 ± 3.1 years) and 8 older adults (60 - 68 years old, mean ± SD, 63.3 ± 3 years).
- All participants had clinically normal hearing (≤ 25 dB HL at 125 – 4 kHz) and no history of neurological or middle ear disorders.
- Participants had normal IQ scores [mean ± SD, 112.5 ± 10.26 for younger adults, and mean ± SD, 123.14 ± 13.8 in older adults on the Wechsler Abbreviated Scale of Intelligence [5]].
- Older adults were also screened for dementia on the Montreal Cognitive Assessment (MOCA) [6] [mean ± SD, 25.875 ± 2.23].

Behavioral data

The Quick Speech-in-Noise test (QuickSIN) [7] was used to objectively measure the participant's sentence recognition in noise. Four lists were used for each participant and were averaged to produce a final score.

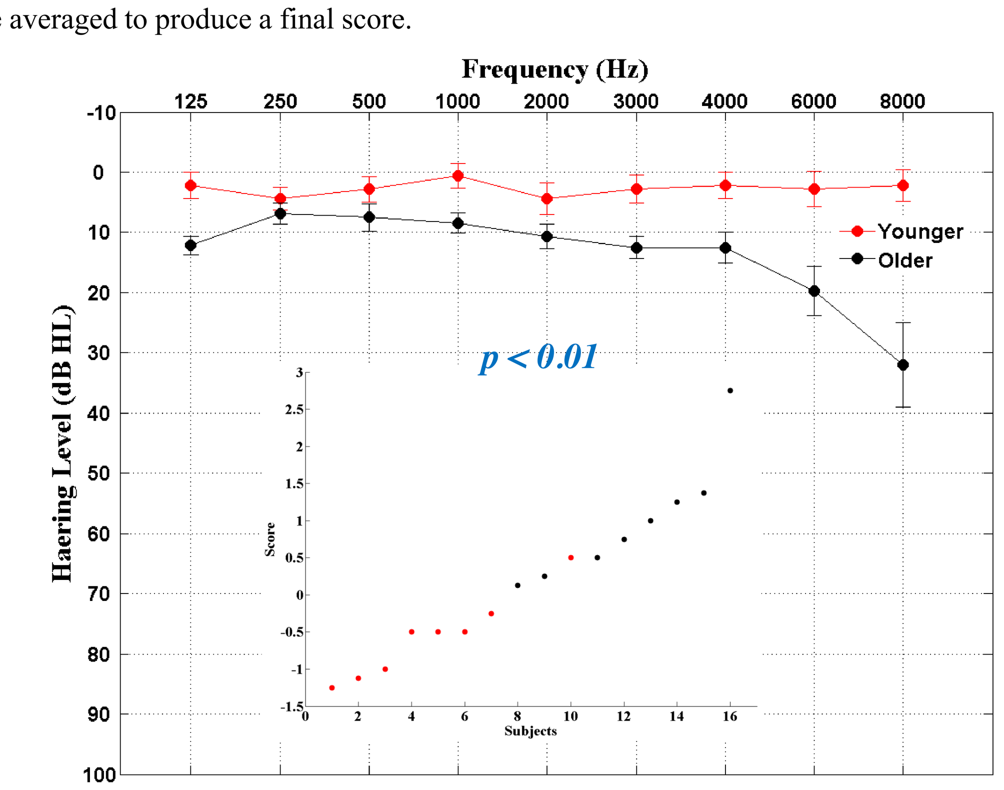
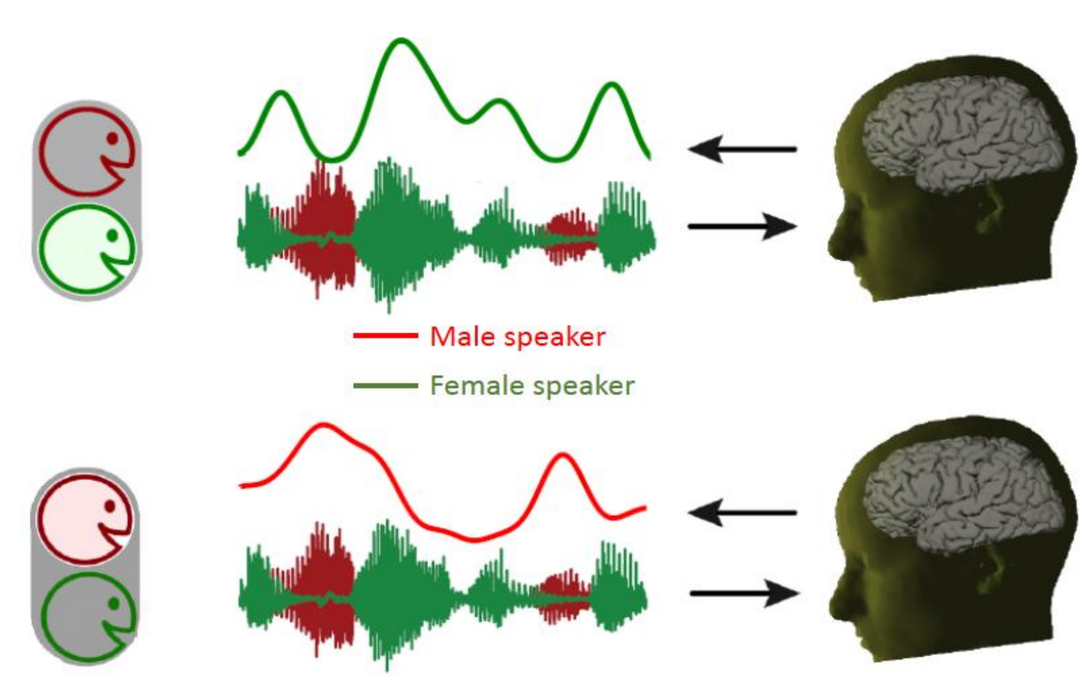


Fig. 1 Audiogram (mean ± 1SE) for younger (red) and older (black) adults. The inset shows the results of the QuickSIN for each participant in ascending order (the lower the score, the better the understanding of speech in noise).

MEG recordings

- The speech presented at 62 dB SPL and low-pass filtered below 4 kHz.
- Participants were asked to attend to one of two stories presented diotically while ignoring the other one.
- One story was spoken by a male and the other by a female. Three trials, each one approximately 1 minute in duration were recorded.
- Neuromagnetic signals were recorded using a 157-signal whole head MEG system (Kanazawa Institute of Technology, Kanazawa, Japan) in a magnetically shielded room, with a 1 kHz sampling rate. A 200 Hz low-pass filter and a notch filter at 60 Hz were applied online.

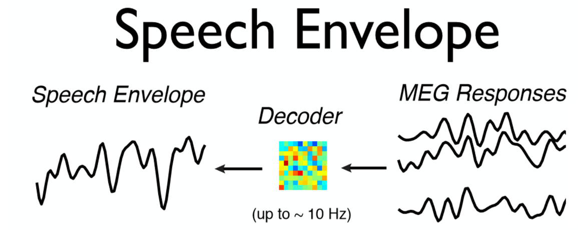


Graphical representation of the MEG task. Subjects were instructed to attend to either the male speaker (red) or to the female speaker (green), while trying to ignore the competing talker. The MEG response was used to reconstruct the envelope of the speech stimulus to which the participant was instructed to attend.

MEG analysis

- Data were denoised using Time-shifted PCA [8].
- Denoised data were filtered between 2 – 8 Hz and separated into components via the DSS algorithm [9].
- Only the first 6 DSS components were retained, and then filtered between 1 - 8 Hz.
- A linear model [2,3] used these filtered responses to reconstruct the envelope of the foreground and background. Success in this prediction is measured by the linear correlation between the predicted and actual speech envelope and of the contrast between foreground and background. The first measurement will ensure us that the stimulus was properly encoded, while the latter measurement will give us an indication of how efficient the subject was in suppressing the background noise.

Neural Reconstruction of Speech Envelope



Results

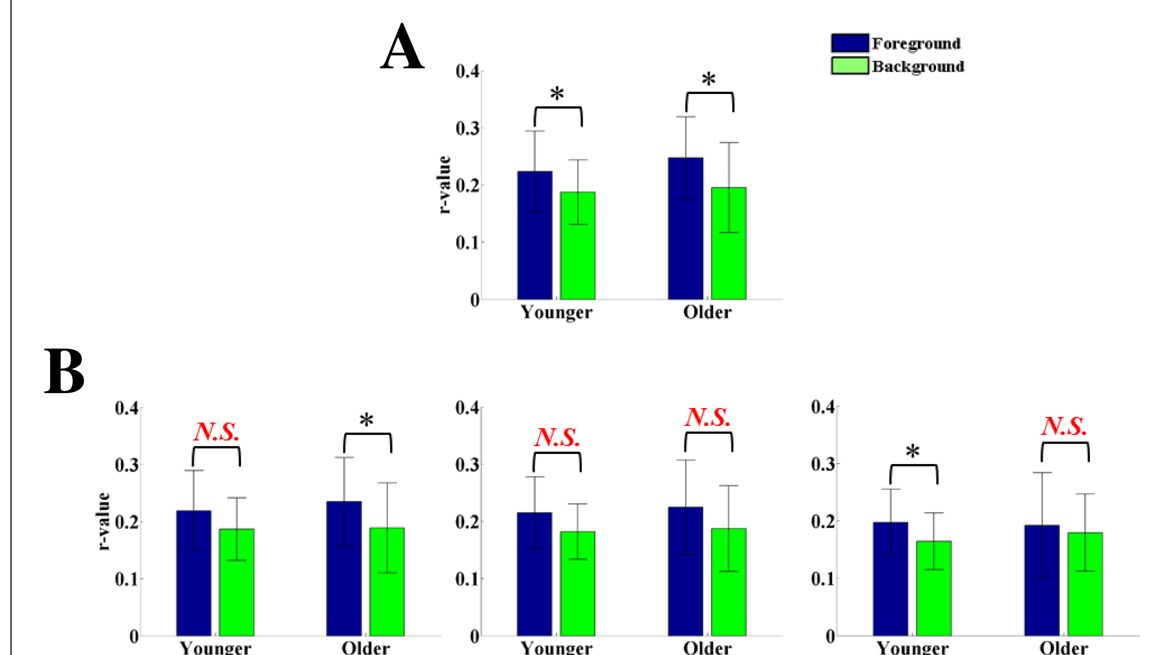


Fig. 2 *Top* Decoding accuracy of the average of the three trials for younger and older adults for foreground (blue), and background (green) calculated in a 450 ms integration window. Both younger and older adults reconstruct the foreground significantly better than the background and the noise floor. *Bottom* Examples of how the integration windows (350, 250 and 150ms are represented) significantly affects older adults more than younger adults. Each subplot represents a different integration window used to reconstruct the speech envelope. For all the integration windows utilized, foreground and background were significantly different from noise floor ($p < 0.001$). ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$; N.S. = Non significant (paired t-test)

Integration window (ms)	500	450	400	350	300	250	200	150	100
Younger	***	***	***	***	***	***	***	**	N.S.
Older	***	***	**	**	*	*	N.S.	N.S.	N.S.

Table 1 Significance values (paired t-test) for the contrast between foreground and background when the speech envelope was reconstructed using single trials. ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$; N.S. = Non significant

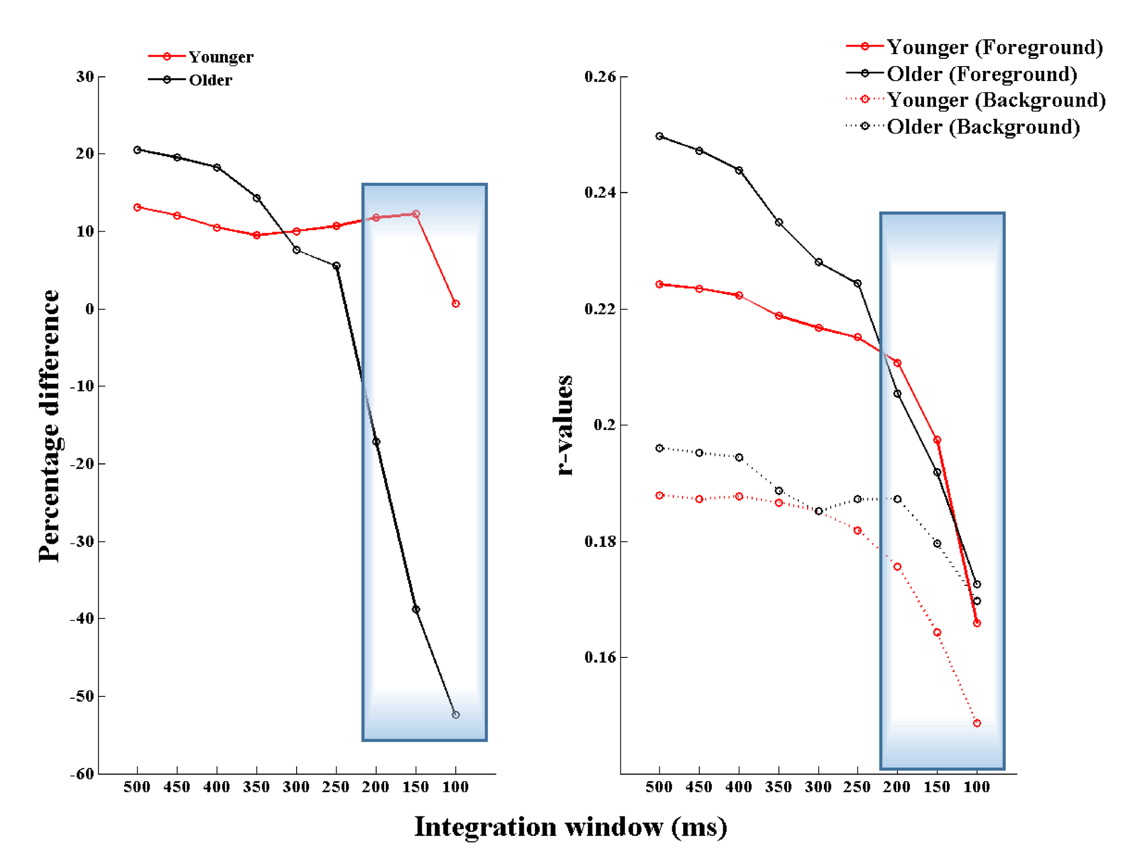


Fig. 3 *Left*: Percentage difference between foreground and background at each integration window for younger (red) and older (black) adults. *Right*: Foreground (solid line) and background (dashed line) for younger (red) and older (black) adults at each integration window. In both figures the shadow area represents the region where the performance of older adults becomes significantly worse with respect to younger adults. Note how the contrast (foreground vs. background) in younger adults stays stable up to 150 ms, while below 250 ms older adults experience problems suppressing the competing talker.

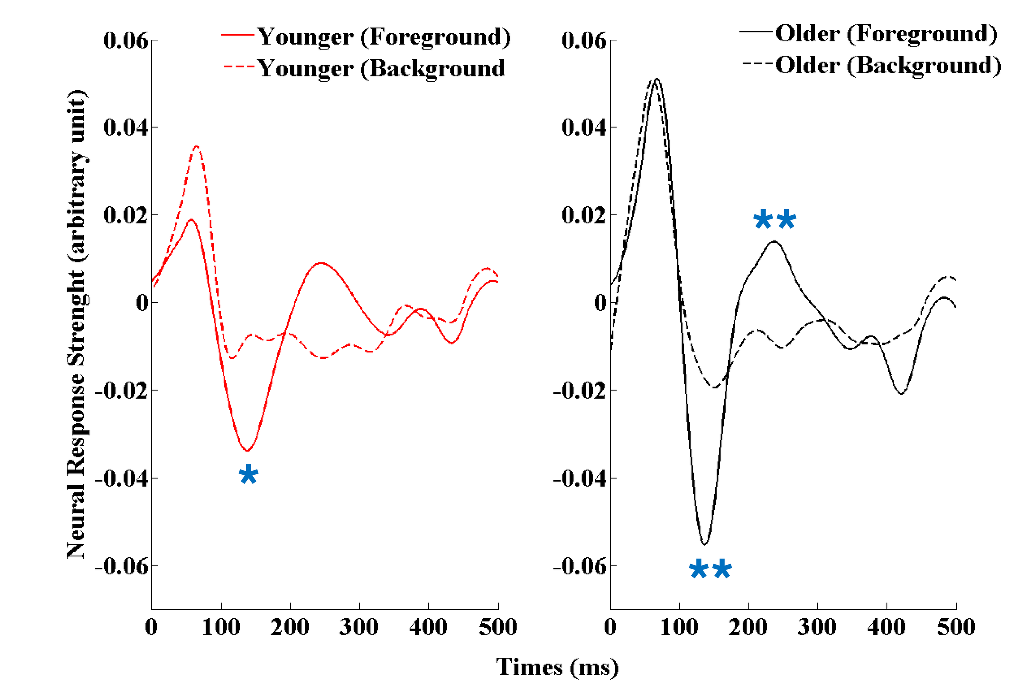


Fig. 4 Temporal Response Field (TRF) of the younger adults (left panel) and older adults (right panel) for the first DSS. Solid line represents the TRF of the foreground, while dashed line represents the TRF of the background. Three major peaks are identified: $M50_{TRF}$ and $M100_{TRF}$ and $M200_{TRF}$. Note that only $M100_{TRF}$ is significant different (paired-t test; *$p < 0.05$, **$p < 0.01$) between foreground and background in both age groups, while $M50_{TRF}$ does not significantly change in either group. However, the difference at $M200_{TRF}$ is only significant in older adults (**$p < 0.01$)

Conclusions

- Older adults performed more poorly than younger adults while listening to speech in noise, despite having a normal audiograms. Results from the QuickSIN showed significant differences between younger and older adults in sentence recognition in noise.
- Older adults were able to neurally reconstruct the target speech and filter out the competing talker as long as the integration window was long enough to allow them to process the speech.
- As the integration window was shortened, the contrast between foreground and background decreased in older adults, mainly due to the inability to neurally suppress the competing talker.
- The contrast remained stable in younger adults up to 150 ms.
- The forward model showed that the TRF has two main peaks at ~50 ms and ~100 ms.
- The $M100_{TRF}$ of the foreground is significantly different from the background in both age groups, in agreement with the general notion that attention modulates late activity [10].
- The $M50_{TRF}$ of the foreground is not significantly different from the background in both age groups, in agreement with the general notion that early activity is not modulated by attention [11].
- The $M200_{TRF}$ is significant different between foreground and background only in older adults; interestingly, the contrast in older adults significantly drops below ~200 ms.
- Altogether our findings support the idea that loss of temporal precision in older adults may, at least in part, account for their difficulties in understanding speech in noise.

References

1. Gordon-Salant et al. (2006). **Age-related differences in identification and discrimination of temporal cues in speech segments**. *J. Acoust. Soc. Am., 119(4), 2455-2466*.
2. Ding and Simon (2012a) **Emergence of neural encoding of auditory objects while listening to competing speakers** *PNAS 11854-11859*
3. Ding and Simon (2012b) **Neural coding of continuous speech in auditory cortex during monaural and dichotic listening** *J Neurophysiol* 107**:** 78-89
4. Anderson et al (2012) **Aging affects neural precision of speech encoding** *J Neurosci, 32*, 14156-14164
5. Zhu, J., Garcia, E. (1999). **The Wechsler Abbreviated Scale of Intelligence (WASI)**. *New York: Psychological Corporation*
6. Nasreddine, Z. S., Phillips, N. A., Bedirian, V., et al. (2005). **The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment**. *J Am Geriatr Soc, 53*, 695-699
7. Killion et al. (2004). **Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners**. *J Acoust Soc Am, 116*, 2395-2405
8. De Cheveigné, A and Simon, J. Z. (2007). **Denoising based on time-shift PCA**. *J Neurosci Methods, 165*, 297-305
9. De Cheveigne and Simon (2008) **Denoising based on spatial filtering** *J Neurosci Methods* 171**:** 331-339
10. Okamoto et al. (2011) **Sound processing hierarchy within human auditory cortex** *J Cogn Neur, 23(8),1855-1863*
11. Ross et al (2000) **A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones** *J Acoust Soc Am, 108*(2), 679-691

Acknowledgements

This study has been funded by UMD ADVANCE Program for Inclusive Excellence, National Science Foundation and by and by NIH grant R01DC008342