Origins of cortical over-representation of speech in older adults

Computational Sensorimotor Systems Lab

Christian Brodbeck, Alessandro Presacco, Stefanie Kuchinsky, Samira Anderson & Jonathan Z. Simon
University of Maryland, College Park

Background

Previous research using magnetoencephalography (MEG) has found that older adults' cortical responses to speech track the envelope of the acoustic signal more robustly than younger adults' responses (Presacco et al., 2016a&b). This could have different reasons:

- Low level age-related change, e.g., excitation/inhibition imbalance
  - Decrease in cortical inhibition could lead to stronger evoked responses (e.g. Overton & Recanzone, 2016)
- Top-down/strategic processing
  - Higher level processes recruited to compensate for lower level deficits (e.g., degraded input from the periphery) lead to activation in additional brain regions (e.g., Peelle et al., 2010)
- Attention:
  - Increased sensory attention due to increased task demands is associated with stronger sensory responses (Woldorff et al., 1993)

Here we used MEG source localization to determine

- Which parts of the temporal lobe show increased phase-locked activity
- At what latency increased responses occur

Methods

Participants

- 17 younger (18-27 years) and 23 older (60+) adults

Procedure

- 157 axial gradiometer whole head MEG (KIT, Kanazawa, Japan)
- For source space analysis, MEG responses to 2 one-minute long segments of clean speech (The Legend of Sleepy Hollow); each segment repeated 3 times for a total of 6 minutes of data per subject
- For Decoding analysis, additional segments with two speakers at different signal to noise ratios, task to attend to one and ignore the other

Stimulus reconstruction (Presacco et al., 2016b)

- Speech stimulus represented as envelope of the analytic signal (1-8 Hz)
- Linear 500 ms kernel trained to predict stimulus from all MEG data (1-8 Hz)
  - Coordinate descent algorithm (David et al., 2007)
  - Early stopping based on cross-validation

Source localization

- Temporal signal space separation
- Zero-phase FIR filter (1-8 Hz)
- Average brain model, scaled to match each participant's head (FreeSurfer fsaverage)
- Minimum norm estimates at virtual source dipoles equally spaced across the white matter surface, oriented perpendicularly to the cortical surface
- Speech stimulus represented as envelope of the analytic signal
- Linear 500 ms kernel trained for each source dipole to predict estimated current time course from the stimulus
  - Basis of 50 ms Hamming windows
  - Coordinate descent algorithm (David et al., 2007)
  - Early stopping based on cross-validation

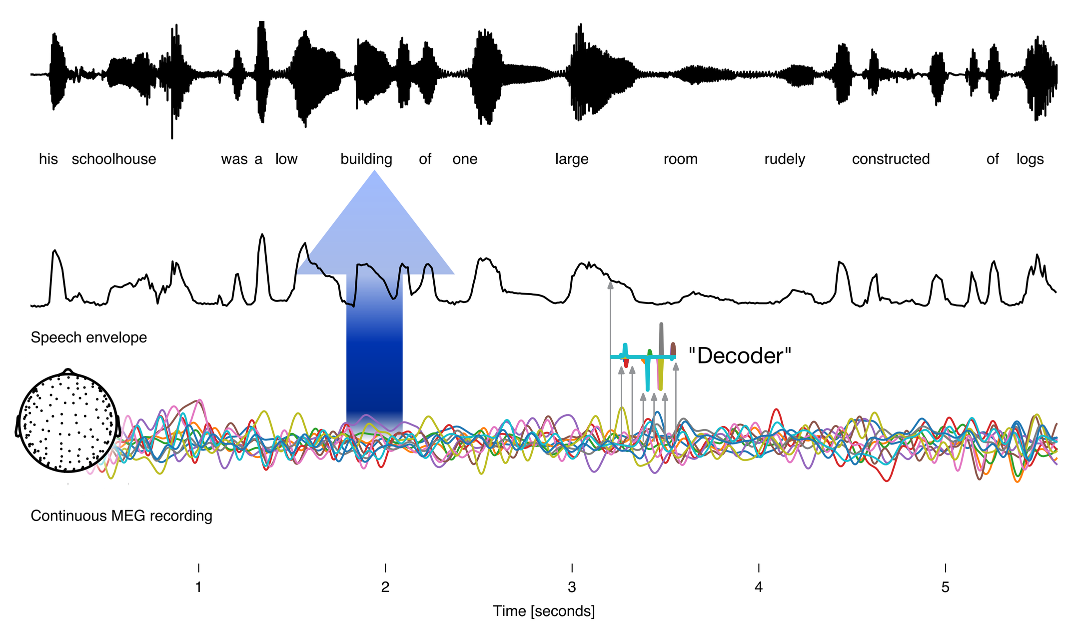
Statistical evaluation

- Overall model fit:
  - Correlation coefficient between predicted and actual source time course (Fisher z-transformed to correct distribution for fixed end-points at -1 and 1)
  - Bias corrected using model in which the predictor variable was temporally misaligned with the response
  - Age difference with repeated-measures t-test at each source dipole,
  - Threshold-free cluster enhancement (Smith and Nichols, 2009)
  - Estimation of the null distribution by permuting group membership 10.000 times
- TRF:
  - All values transformed to their absolute value, to prevent negative and positive currents from cancelling out
  - Weighted average in the region of significant difference in z-values
  - TRF amplitude time course analyzed with repeated-measures *t*-test, TFCE and permutation distribution as above
- TRF peaks
  - Peak windows determined based on inspection of TRF time course
  - Average of the absolute TRF in window for each participant
  - Smoothed with Gaussian kernel (STD = 5 mm)
  - Tested as above

Stimulus reconstruction

MEG responses to one minute long segments of continuous speech, under natural listening conditions (excerpts from audiobook)

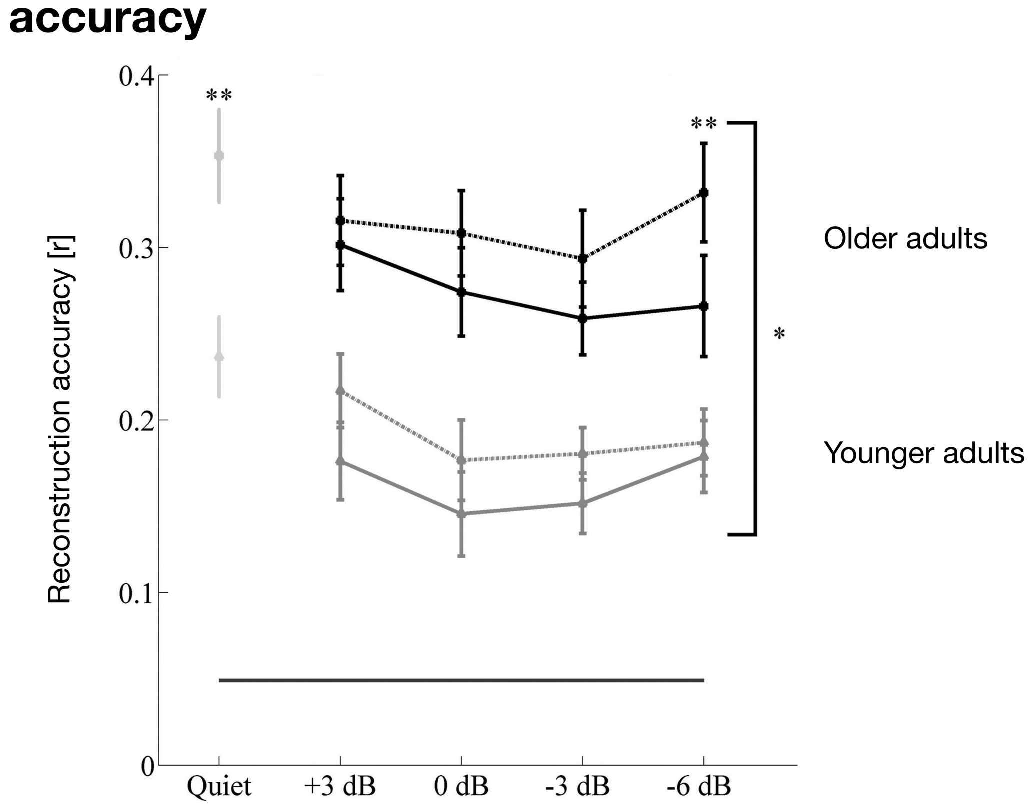
Method: Stimulus reconstruction



The signal from the MEG sensors is used jointly to reconstruct the acoustic envelope of the speech stimulus

- The reconstructed stimulus is the convolution of the kernel ("decoder") with the MEG signal
- Reconstruction accuracy is an estimate of how much information the responses contain about the stimulus
  - Measured as correlation between actual and reconstructed stimulus
- The convolution model used for reconstruction is primarily sensitive to phase-locked brain activity

Older adults: higher stimulus reconstruction accuracy



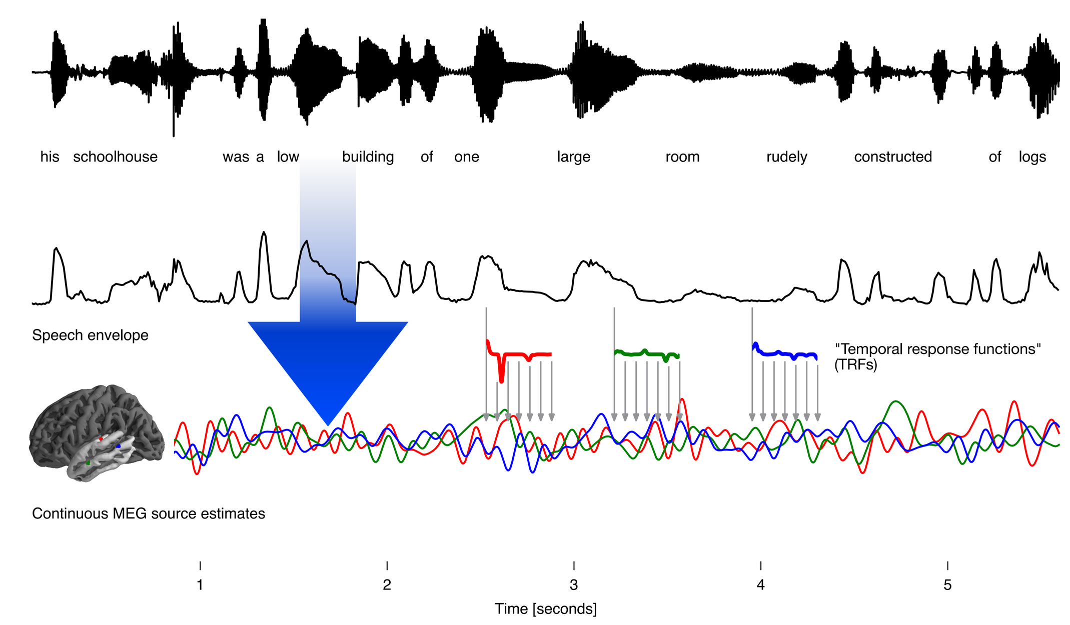
Older adults' cortical responses allow more accurate reconstruction of the speech envelope than younger adults' (from Presacco et al., 2016b)

- Holds across different listening conditions (clean speech and speech with background speaker at difference SNRs)
- Suggests that older adults' cortical responses carry more information about the speech envelope
- *Where in the cortex and at which latencies are older adults' responses amplified?*

Brain responses

Distributed minimum norm estimates (MNE) used to estimate electrical activity at virtual current source dipoles across the temporal lobes. Activity at these source dipoles was modeled as a response to the acoustic envelope of speech using a linear convolution model (David et al., 2007; Brodbeck et al., 2018).

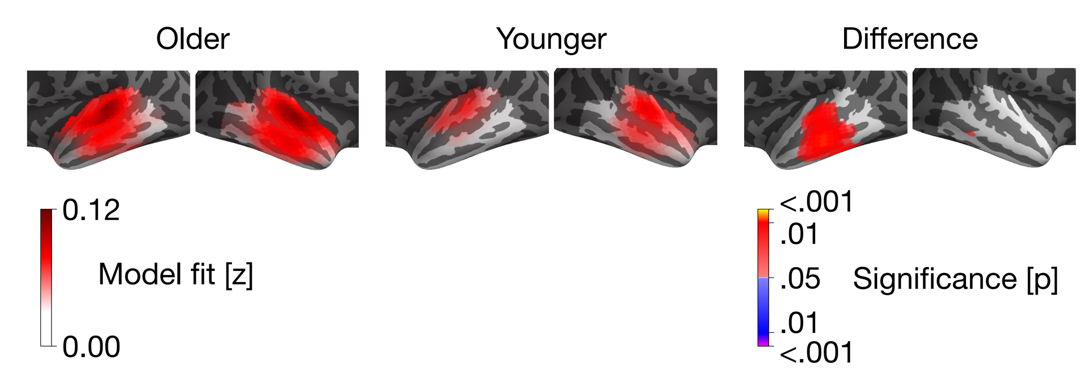
Method: Temporal response functions



The signal at each virtual source dipole (illustrated as red/green/blue lines) is modeled as linear convolution of the speech envelope with a temporal response function (TRF)

- Model fit is evaluated by how well the signal at each dipole can be modeled (correlation coefficient)

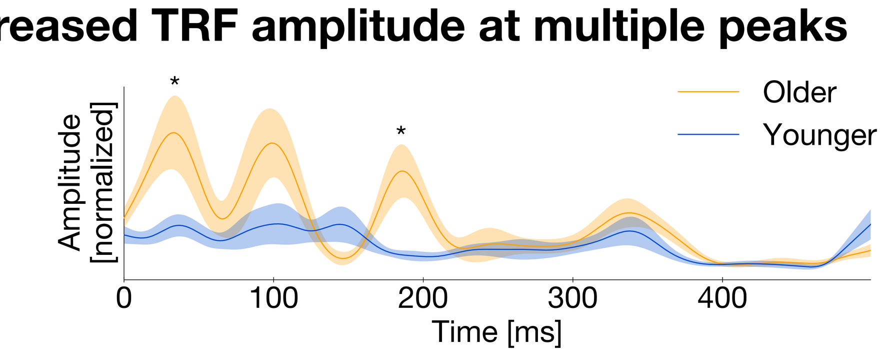
Localized response prediction accuracy



Older adults exhibit stronger responses to clean speech in non-primary auditory cortex

- Older adults' MEG responses reflected the acoustic envelope more strongly in a region of the left temporal lobe
- Localization consistent with a region outside of core auditory cortex
- Lateralization was not significant (p = .285)
- The amplitude of the temporal response functions (TRFs) in the significant region was analyzed for a better understanding of the timing of the effects

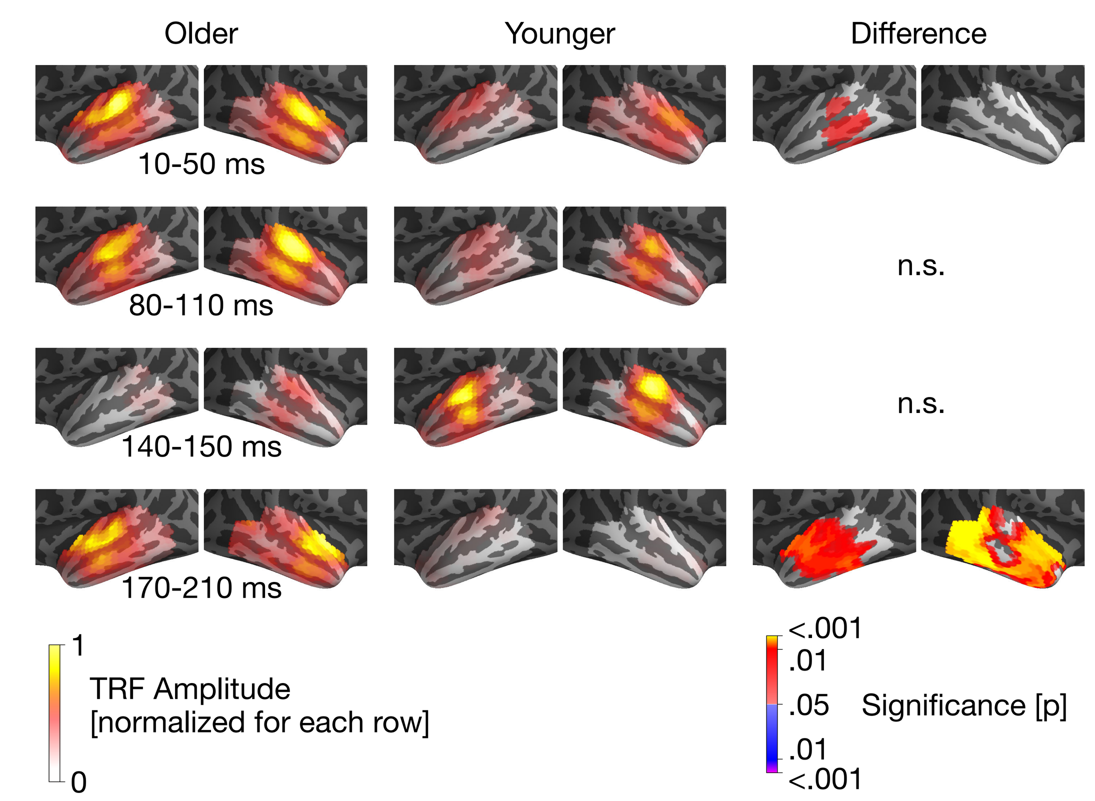
Increased TRF amplitude at multiple peaks



TRF amplitude (averaged in significant region shown above) is significantly larger in older adults at early (~30 ms) and late (~200 ms) peaks

- Younger adults seem to have similar but weaker peaks at ~30 and ~100 ms
- A third peak occurs in younger adults at ~150 ms already; older adults' ~200 ms peak could be an enlarged and delayed version

Response function peaks



TRF amplitudes, averaged in time windows around prominent peaks, suggest different anatomical origins

- ~30 ms: older adults' response significantly enlarged; region of significant group difference consistent with main difference outside core auditory cortex
- ~100 ms: group difference not significant
- ~150 ms: non-significantly enhanced response peak in younger adults
- ~200 ms: additional peak in older adults' TRFs with wide-spread distribution

Conclusions

Compared to younger adults, older adults' cortical responses track the acoustic envelope of speech more robustly.

- Older adults' responses to clean speech differ from younger adults' responses at different TRF peaks with different latencies, suggesting multiple reasons for increased tracking
- An early ~30 ms difference is consistent with a low-level processing change
  - Localization difference suggests non-primary auditory cortex involvement
  - Consistent with excitatory-inhibitory imbalance, leading to rapid activation in a larger area
- A late ~200 ms difference is consistent with recruitment of additional processing resources

References

Brodbeck, C., Presacco, A., & Simon, J. Z. (2018). Neural source dynamics of brain responses to continuous stimuli: Speech processing from acoustics to comprehension. NeuroImage, 172, 162–174. https://doi.org/10.1016/j.neuroimage.2018.01.042

David, S. V., Mesgarani, N., & Shamma, S. A. (2007). Estimating sparse spectro-temporal receptive fields with natural stimuli. Network: Computation in Neural Systems, 18(3), 191–212. https://doi.org/10.1080/09548980701609235

Overton, J. A., & Recanzone, G. H. (2016). Effects of aging on the response of single neurons to amplitude-modulated noise in primary auditory cortex of rhesus macaque. Journal of Neurophysiology, 115(6), 2911–2923. https://doi.org/10.1152/jn.01098.2015

Peelle, J. E., Troiani, V., Wingfield, A., & Grossman, M. (2010). Neural Processing during Older Adults' Comprehension of Spoken Sentences: Age Differences in Resource Allocation and Connectivity. Cerebral Cortex, 20(4), 773–782. https://doi.org/10.1093/cercor/bhp142

Presacco, A., Simon, J. Z., & Anderson, S. (2016a). Evidence of degraded representation of speech in noise, in the aging midbrain and cortex. Journal of Neurophysiology, 116(5), 2346–2355. https://doi.org/10.1152/jn.00372.2016

Presacco, A., Simon, J. Z., & Anderson, S. (2016b). Effect of informational content of noise on speech representation in the aging midbrain and cortex. Journal of Neurophysiology, 116(5), 2356–2367. https://doi.org/10.1152/jn.00373.2016

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. NeuroImage, 44(1), 83–98. https://doi.org/10.1016/j.neuroimage.2008.03.061

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proceedings of the National Academy of Sciences, 90(18), 8722–8726.

Acknowledgements

This research was funded by NIH grant P01-AG055365.

Some of this work has been accepted for publication in Acta Acustica united with Acustica as part of a the ISH 2018 special issue.