Sound, Sight, Age: Speech Recognition Processes in Noise

Kenneth I. Vaden Jr., James W. Dias

Department of Otolaryngology - Head & Neck Surgery, Medical University of South Carolina, Charleston, SC, USA

Aging slows perceptual decision-making during speech recognition in noise

Introduction

Older adults often show slower and poorer speech recognition in noise than younger adults (Dubno et al., 1984; Humes & Christopherson, 1991; Sommers et al., 2011).

Visual information can substantially improve speech recognition in noise (e.g., Sumby & Pollack, 1954) and may support similar performance for older and younger adults despite poorer auditory-only recognition in older adults (Dias et al., 2021).

It remains unclear to what extent processes supporting recognition of noisy speech are affected by older age and intelligibility-related factors. Perceptual decision-making (PDM) models use response latencies to estimate evidence accumulation rate, response criterion, and nondecision time (e.g., Anders et al., 2016; Hanks and Summerfield, 2017).

The current study tested whether age, modality, and speech signal-to-noise ratio (SNR) influence PDM processes for words in noise.

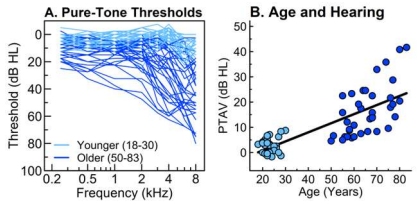
Older age was predicted to produce slower accumulation, more cautious criteria, and longer nondecision times (Starns & Ratcliff, 2010; Dully et al., 2018). Greater speech intelligibility was predicted to produce faster accumulation, more cautious criteria, and shorter nondecision times (Ratcliff & McKoon, 2008; Vaden et al., 2022).

Methods: Data and Participants

Retrospective analyses were performed with data from a speech recognition in noise experiment (Dias et al., 2021; 2026).

Participants (N = 65; 44 women; age: 47.4 ± 22.2 years, mean ± SD) ranged from 18 to 83 years of age, with the participant samples concentrated in younger (18–30) and older (50–83) age ranges.

Normal hearing to moderate hearing loss, with average pure-tone thresholds (PTA) from -1.7 to 41.6 dB HL (250–4000 Hz).



Above. A range of normal hearing to moderate hearing loss was observed for the younger and older adult study participants. Average audiograms for each of the participants are shown in (A). The PTA was significantly correlated with participant age [r = 0.77, p < 0.001] (B).

Methods: Word Identification in Noise Task

Participants identified 192 bi-syllabic words pseudo-randomly presented in a Gaussian noise at -5, 0, +5 dB SNR, in audio-only (AO) or audiovisual conditions (AV; articulating face visible).

Participants were instructed to type the perceived word on each trial or guess if they were not sure. Response latencies were collected at the onset of typed responses.

A Shifted-Wald PDM model (Anders et al., 2016) was fitted to the correct-trial response latencies to estimate three model parameters per participant and condition. A linear mixed model regression was used to assess the effects of age, PTA, SNR, and AO/AV on each of the PDM parameters.

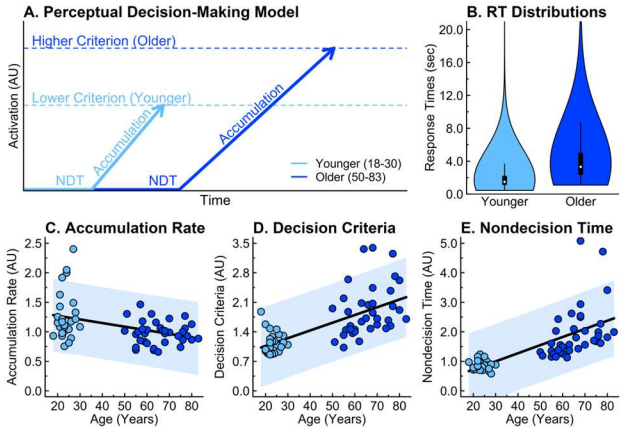
Results: Performance and Perceptual Decision-Making

Word Identification Accuracy

Accuracy was significantly better with higher SNR and AV compared to AO [p < 0.001]. Significant interactions included a larger AV benefit with lower SNR [p < 0.02] and for older compared to younger adults [p = 0.02].

Age-Related Differences in Perceptual Decision-Making Parameters

Older adults exhibited significantly slower evidence accumulation rates [$t(63)$ = -3.61, p < 0.001], higher decision thresholds [$t(63)$ = 7.09, p < 0.001], and longer nondecision times than younger adults [$t(63)$ = 7.81, p < 0.001]. Hearing loss (PTA) was not associated with PDM parameters [p > 0.50], nor did it account for age effects in control analyses.

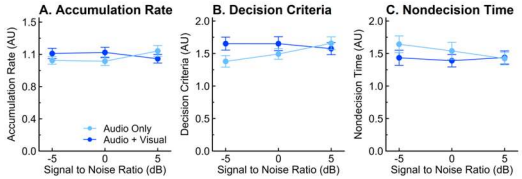


Above. PDM model schematic based on parameter estimates for the younger and older study participants (A), which were derived from their RT distributions (B). Older age was associated with significantly slower evidence accumulation (C), higher / more cautious decision criteria (D), and slower nondecision times (E).

Results: SNR × Modality Interactions

Significant SNR-modality interactions were observed, such that higher SNR was associated with significantly faster evidence accumulation [$t(63)$ = 2.10, p = 0.03], higher decision thresholds [$t(63)$ = 2.87, p = 0.004], and shorter nondecision times [$t(63)$ = -3.19, p = 0.002] for AO compared to AV speech.

The interaction effects also reflected faster evidence accumulation, higher decision thresholds, and shorter nondecision times for AV versus AO speech, in the lower SNR conditions.



Above. Significant interactions between presentation modality and speech SNR were observed for the PDM processes (A to C), with significant SNR effects in the AO condition and AV-AO differences in the lower SNR conditions.

Conclusions

Older adults exhibited more cautious decision criteria, longer nondecision times, and slower evidence accumulation than younger adults, all contributing to slower responses. The long task response intervals used here may be important for detecting some age effects.

Factors that improved speech intelligibility and recognition accuracy (e.g., higher SNR, audiovisual modality) were also associated with significant PDM benefits, consistent with prior diffusion-model findings for higher-quality sensory information (e.g., Ratcliff & McKoon, 2008).

Together, these findings support the view that multisensory input enhances perceptual recognition processes and benefits communication for older adults in challenging listening environments.

References. [1] Anders R, Alario FX, Van Maanen L, 2016. The shifted Wald distribution for response time data analysis. Psychol Methods 21, 309–327. [2] Dias JW, McClaskey CM, Harris KC, 2021. Audiovisual speech is more than the sum of its parts: Auditory-visual superadditivity compensates for age-related declines in audible and lipread speech intelligibility. Psychol Aging 36, 520–530. [3] Dias JW, Schwartz-Leyzac KC, Harris KC, 2026. Audiovisual speech perception in aging cochlear implant users and age-matched normalplanted adults. Ear Hear 47, 737–747. [4] Dubno JR, Dirks DD, Morgan DE, 1984. Effects of age and mild hearing loss on speech recognition in noise. J Acoust Soc Am 76, 87–96. [5] Dully J, McGovern DP, O'Connell RG, 2018. The impact of natural aging on computational and neural indices of perceptual decision making: A review. Behav Brain Res 355, 48–55. [6] Hanks TD, Summerfield C, 2017. Perceptual decision making in rodents, monkeys, and humans. Neuron 93, 15–31. [7] Humes LE, Christopherson L, 1991. Speech identification difficulties of hearing-impaired elderly persons: The contributions of auditory processing deficits. J Speech Hear Res 34, 686–693. [8] Ratcliff R, McKoon G, 2008. The diffusion decision model: Theory and data for two-choice decision tasks. Neural Comput 20, 873–922. [9] Sommers MS, Hale S, Myerson J, Rose N, Tye-Murray N, Spehar B, 2011. Listening comprehension across the adult lifespan. Ear Hear 32, 775–781. [10] Spaniol J, Madden DJ, Voss A, 2006. A diffusion model analysis of adult age differences in episodic and semantic long-term memory retrieval. J Exp Psychol Learn Mem Cogn 32, 101–117. [11] Sumby W, Pollack I, 1954. Visual contribution to speech intelligibility in noise. J Acoust Soc Am 26, 212–215. [12] Vaden KI, Teubner-Rhodes S, Ahlstrom JB, Dubno JR, Eckert MA, 2022. Evidence for cortical adjustments to perceptual decision criteria during word recognition in noise. NeuroImage 253, 119042.

Acknowledgements. This work at Medical University of South Carolina was supported in part by National Institutes of Health / National Institute on Deafness and Other Communication Disorders R01 DC021423 (Vaden). Original data collection was supported by NIH/NIDCD grants R01 DC014467, R01 DC017619, and P50 DC000422 (Harris), Institutional Training Grant T32 DC014435 (Dubno), and National Center for Advancing Translational Sciences UL1 TR001450. Research was conducted in facilities supported by C06 RR014516. We thank the original study participants.

Poster PDF