

Auditory Working Memory Explains Variance in Speech Recognition in Older Listeners Under Adverse Listening Conditions

This article was published in the following Dove Press journal:
Clinical Interventions in Aging

Subong Kim¹
Inyong Choi^{1,2}
Adam T Schwalje²
Kyoosang Kim³
Jae Hee Lee⁴

¹Department of Communication Sciences and Disorders, University of Iowa, Iowa City, IA 52242, USA; ²Department of Otolaryngology – Head and Neck Surgery, University of Iowa Hospitals and Clinics, Iowa City, IA 52242, USA;

³Department of Occupational Environmental Medicine, Seoul Medical Center, Seoul 02053, South Korea;

⁴Department of Audiology and Speech-Language Pathology, HUGS Center for Hearing and Speech Research, Hallym University of Graduate Studies, Seoul 06197, South Korea

Introduction: Older listeners have difficulty understanding speech in unfavorable listening conditions. To compensate for acoustic degradation, cognitive processing skills, such as working memory, need to be engaged. Despite prior findings on the association between working memory and speech recognition in various listening conditions, it is not yet clear whether the modality of stimuli presentation for working memory tasks should be auditory or visual. Given the modality-specific characteristics of working memory, we hypothesized that auditory working memory capacity could predict speech recognition performance in adverse listening conditions for older listeners and that the contribution of auditory working memory to speech recognition would depend on the task and listening condition.

Methods: Seventy-six older listeners and twenty younger listeners completed four kinds of auditory working memory tasks, including digit and speech span tasks, and sentence recognition tasks in four different listening conditions having multi-talker noise and time-compression. For older listeners, cognitive function was screened using the Mini-Mental Status Examination, and audibility was assured.

Results: Auditory working memory, as measured by listening span, significantly predicted speech recognition performance in adverse listening conditions for older listeners. A linear regression model showed speech recognition performance for older listeners could be explained by auditory working memory whilst controlling for the impact of age and hearing sensitivity.

Discussion: Measuring working memory in the auditory modality facilitated explaining the variance in speech recognition in adverse listening conditions for older listeners. The linguistic features and the complexity of the auditory stimuli may affect the association between working memory and speech recognition performance.

Conclusion: We demonstrated the contribution of auditory working memory to speech recognition in unfavorable listening conditions in older populations. Taking the modality-specific characteristics of working memory into account may be a key to better understand the difficulty in speech recognition in daily listening conditions for older listeners.

Keywords: auditory working memory, age, hearing loss, speech recognition

Correspondence: Jae Hee Lee
Department of Audiology and Speech-Language Pathology, HUGS Center for Hearing and Speech Research, Hallym University of Graduate Studies, Seoul 06197, South Korea
Tel +82-2-3453-9333
Fax +82-2-3453-6618
Email leejaehee@hallym.ac.kr

Introduction

Older listeners have difficulty understanding speech in unfavorable listening conditions. Presbycusis, age-related hearing loss, is a major cause of this difficulty. However, even older listeners with normal or near-normal hearing have difficulty understanding speech in these adverse listening conditions.^{1,2} Listening, especially in adverse conditions, places a heavy load on executive function.³⁻⁹ One component

of executive function, working memory, is especially crucial as listeners must retain speech information and relate it to the speech that follows while encoding the target signal^{10–12}. The concept of working memory that is highly involved in speech recognition when speech is degraded, due to hearing loss or masking noise, is well established in the Ease of Language Understanding model.^{13,14} In fact, older listeners' working memory capacity predicts their ability to recognize speech in various listening conditions.^{3,4} The reading span (RS) test of working memory¹² that asks participants to read multiple unrelated sentences and remember the last word of each sentence is often used to determine the relationship between working memory and speech recognition.^{4,7,8,15}

One underlying issue is whether the stimuli should be given with auditory or visual information, or both, for working memory tasks. For instance, the RS test has an auditory equivalent, the listening span (LS) test.¹⁶ Most of the auditory studies use visual stimuli for working memory tasks, considering that the audibility of the speech signal can affect performance on the working memory task.⁴ However, since Baddeley's working memory model introduced modality-specific subsystems for sensory information,^{10,11} it has been suggested that auditory working memory may be more relevant to speech recognition in various listening conditions than visual working memory.^{4,16–19} In other words, presenting the working memory task in the auditory modality may be more "ecologically valid" since speech recognition performance is also measured in the matched condition.^{17,18} Neuroimaging studies also reveal that cortical activity during working memory tasks depends on the sensory input.^{20,21} On the other hand, some behavioral studies do not support the necessity of using auditory working memory tasks, showing weak correlations between auditory working memory and speech recognition in unfavorable listening conditions.^{22,23} It should be noted that the tasks in those studies were administered in young normal hearing populations, thereby preventing any possible sensory deficiency from affecting the outcome. However, given that the predictive effect of working memory on speech recognition in adverse listening conditions may depend on listeners' age,²⁴ both age populations should be tested and compared in investigations about the association between auditory working memory and speech recognition scores.

The association between working memory and speech recognition depends on the type of working memory task, speech recognition measure, masking noise, and any other acoustic distortion.²⁵ Although previous studies explore

the association between working memory, mostly measured by RS, and unaided speech recognition in noise, findings are variable.^{3,4} Even when the results show a significant association, the predictive effect of working memory on speech recognition is often secondary to hearing loss or predicted by age. This inconsistency could be attributed to the use of less informative visual working memory measures. To clarify the modality-specific association between auditory working memory and speech recognition in adverse listening conditions, we need to examine this association systematically along with different types of auditory working memory tasks and various listening conditions for speech recognition tests. Gordon-Salant and Cole²⁶ found that LS, compared to RS or other cognitive measures, was the greatest contributor to speech recognition in noise and that the linguistic complexity of speech tests mediated the effect of working memory and age on the speech-in-noise performance. Based on these findings, the present study further investigated how the linguistic features and the complexity of auditory working memory can affect the association between auditory working memory and speech recognition in noise. In addition, we presented sentences as the speech signal with several variations of listening conditions.

In the present study, we investigated the association between auditory working memory and speech recognition in adverse listening conditions, in younger adults with normal hearing and older adults with up to mild hearing loss. In older listeners, we screened for both audibility and cognitive functions, so that they would be capable of completing auditory working memory tasks. The LS, as well as digit forward/backward span (DFS/DBS) and word span (WS) tasks, were used for auditory working memory measures. A sentence recognition task was conducted in four listening conditions manipulated by the addition of babble noise and rate changes of the target speech signal. The present study was designed to address three essential questions. First, is the auditory working memory capacity of older listeners significantly associated with their speech recognition performance in adverse listening conditions? Second, does the association between auditory working memory and speech recognition depend on the type of working memory tasks and the given listening conditions for speech recognition tests? And third, how is the association in younger listeners different from that in older listeners? We hypothesized that older listeners' auditory working memory capacity would predict their speech

recognition performance independent of age and hearing thresholds. This association would differ among working memory tasks and significantly increase when comparing an easier listening condition to a harder one that would consist of babble noise and faster speech. However, younger listeners with normal hearing were expected to show less contribution of auditory working memory to speech recognition performance in any listening conditions.

Materials and Methods

All procedures were reviewed and approved by the Seoul Medical Center's Institutional Review Board. Written informed consent was obtained for every participant, and all work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Participants

Older Listeners

Native Korean speakers 60 years of age or older, who had no experience with hearing aids, were recruited. All participants had no history of neurological disorders, middle ear pathology, or any prolonged exposure to high-level environmental noise. Inclusion criteria for this older population were passing score ($\geq 27/30$) on the Korean version of Mini-Mental Status Examination (MMSE-KC; Lee et al²⁷); this cutoff score was set to rule out any possible mild cognitive impairment while balancing between sensitivity and specificity of the MMSE-KC test.²⁸ In addition, participants needed 90% or better word recognition score (WRS) using the Korean Standard Monosyllabic Word Lists for Adults (KS-MWL-A)²⁹ at the better ear, and pure-tone hearing thresholds of the better ear were required to be within the 95th percentile of hearing threshold distributions obtained from otologically screened Korean older population according to ISO 8253-1 (2010) across all frequencies at octave-scale.³⁰ Of the 111 adults 60 years of age or older who participated in research at the Seoul medical center, 76 participants (68%) met inclusion criteria. Data from twenty young listeners (range 23 to 29 years, mean 26.95 years, standard deviation (SD) 1.82 years, median 26.50 years, five females (25%)) and 76 older listeners (range 60 to 83 years, mean 68.91 years, SD 5.62 years, median 68 years, 60 females (79%)) were analyzed. Total years of formal education in younger and older listeners, on average, were 15.80 and 12.92 years, respectively.

To isolate the effect of auditory working memory on speech recognition in older listeners, the participant population and experimental settings were controlled as above, and the presentation level was adjusted to the preferred listening level (range 65 to 80 dB SPL, median 65 dB SPL) for each participant to ensure audibility.

Younger Listeners

Young native Korean speakers with normal hearing were selected from the student population of a local university. Normal hearing thresholds were verified using pure-tone audiometry before completing the tasks described below. The presentation level of the stimuli for younger listeners was 65 dB SPL.

Experimental Measures

Auditory Working Memory Tasks

In the present study, two types of digit span task, DFS and DBS, were conducted in quiet using Korean monosyllabic digits from 1 to 9. Digit span tasks started with two sequential digits and incremented by one. Each participant repeated the stimuli in the order for DFS task or in reverse for DBS task. Listener's digit span was determined at the maximum string where the participant gave the right answer on two out of three trials. WS¹¹ and LS¹² were conducted in quiet as speech span tasks. Bisyllabic words from the Korean Standard Bisyllabic Word List for Adults (KS-BWL-A)³¹ were used, and the participant recalled the word in the order. The WS task began with two sequential words and incremented by one. Final word span was determined using the same procedure as for the digit span tasks. For the LS task, each participant recalled the last word of each sequential sentence, regardless of the order. We used the sentences from Lee,³² the Korean version of LS/RS task developed by Daneman and others. The LS task began with two sentences and increased incrementally by one. Final LS was determined in the same way as the other working memory tasks in the present study, except that credit of 0.5 was given when the participant was correct on one out of three trials. For all auditory working memory tasks, if the participants got two correct answers consecutively, they did not need to complete the third trial; that is, two consecutive correct trials out of three earned a full 1 point of credit.

Speech Recognition in Four Listening Conditions

Sentences from the Korean Standard Sentence Lists for Adults (KS-SL-A)³³ were used to evaluate speech

recognition in four listening conditions: (1) natural sentences, (2) sentences with a 30% time-compression (TC), (3) sentences with multi-talker noise that had a 0 dB signal-to-noise ratio (SNR), and (4) sentences with multi-talker noise and a 30% TC. Each listening condition included 20 sentences. The sentence recognition score (SRS) was calculated in percent correct based on keyword scoring. Research has shown that the association between working memory and speech recognition in adverse listening conditions depends on target speech signal and masking noise.²⁵ Listening conditions with 8-talker babble noise, from a collection of multiple passages spoken by a female and a male talker, were included to ensure enough lexical complexity and informational masking with no floor effect. TC by 30% was determined based on previous studies³⁴ and pilot study results that showed a significant drop in speech recognition performance while the target speech signal still sounded natural with little distortion in pitch with this level.

Equipment and Stimuli

In the present study, we used auditory stimuli for both working memory tasks and sentence recognition tasks. All stimuli were recorded by a male and a female native Korean speaker, using the Computerized Speech Lab (Kay Elemetrics Co., Model 4500) with a Shure SM58 microphone located 10 cm from the speaker. All the recorded stimuli within a task were equalized in the root-mean-squared (RMS) level using Adobe Audition version 3.0 (Adobe Systems Incorporated, San Jose, CA, USA). The RMS level of target speech signal and 8-talker babble noise was matched to generate the 0 dB SNR condition. A speech synthesizer, STRAIGHT,³⁵ was implemented in MATLAB (The Mathworks, Inc., R2016b) to apply TC to the target speech signal by 30% (compress 1-second sentence to 0.7-second sentence) with pitch correction.

Study Protocol

All experimental tasks, as well as hearing assessments, including basic audiometry, tympanometry, and WRS, were performed in a double-walled audiometric sound booth. The participants completed the working memory tasks and speech recognition tests with sound stimuli presented by a single loudspeaker placed at 0 degrees and 45 cm from the participant. Stimuli were presented using a clinical audiometer (GSI 61, VIASYS Healthcare, Inc., Madison, WI, USA), and statistics were calculated with MATLAB software (The Mathworks, Inc., R2016b).

Statistical Analysis

To examine the relationship between auditory working memory and speech recognition in adverse listening conditions, we used a stepwise multiple regression model, and the results were compared between both age groups. Listeners' auditory working memory scores, as well as age and hearing sensitivity, were included in the model to explore which variable can significantly predict speech recognition performance in each listening condition, respectively. The unique contribution of each significant predictor to speech recognition performance was examined by running a partial correlation.

Results

Participant Characteristics

Older listeners' mean pure-tone average (PTA; 0.5, 1, 2, and 4 kHz) was 24.42 and 22.75 dB HL for right and left ears, respectively. Figure 1 illustrates the average hearing thresholds across two ears in both younger and older participants. Younger participants were all normal hearing listeners with pure-tone thresholds at or less than 20 dB HL across all the octave-scale frequencies. All participants presented with type A tympanogram in their test ear. The mean WRS measured at the better ear was 100% and 97.39% for the younger and older groups, respectively.

Auditory Working Memory Capacity

Both digit (DFS, DBS) and speech span (WS, LS) were evaluated for younger and older listeners. Auditory working memory capacity measured from those tasks was used as a dependent variable in a mixed analysis of variance (ANOVA) with a between-subject (age group) and a within-subject (DFS vs DBS or WS vs LS) factor. Table 1 and Figure 2 shows auditory working memory capacity for younger and older listeners. The effect for age group ($F_{1,94} = 43.092, p < 0.001$) and working memory condition ($F_{1,94} = 105.218, p < 0.001$, Greenhouse-Geisser correction) were significant for digit span. Speech span also showed a significant effect for age group ($F_{1,94} = 55.964, p < 0.001$) and working memory condition ($F_{1,94} = 76.749, p < 0.001$, Greenhouse-Geisser correction). The interaction between age group and working memory condition was not significant for either digit span ($F_{1,94} = 0.352, p = 0.554$, Greenhouse-Geisser correction) or speech span tasks ($F_{1,94} = 2.522, p = 0.116$, Greenhouse-Geisser correction).

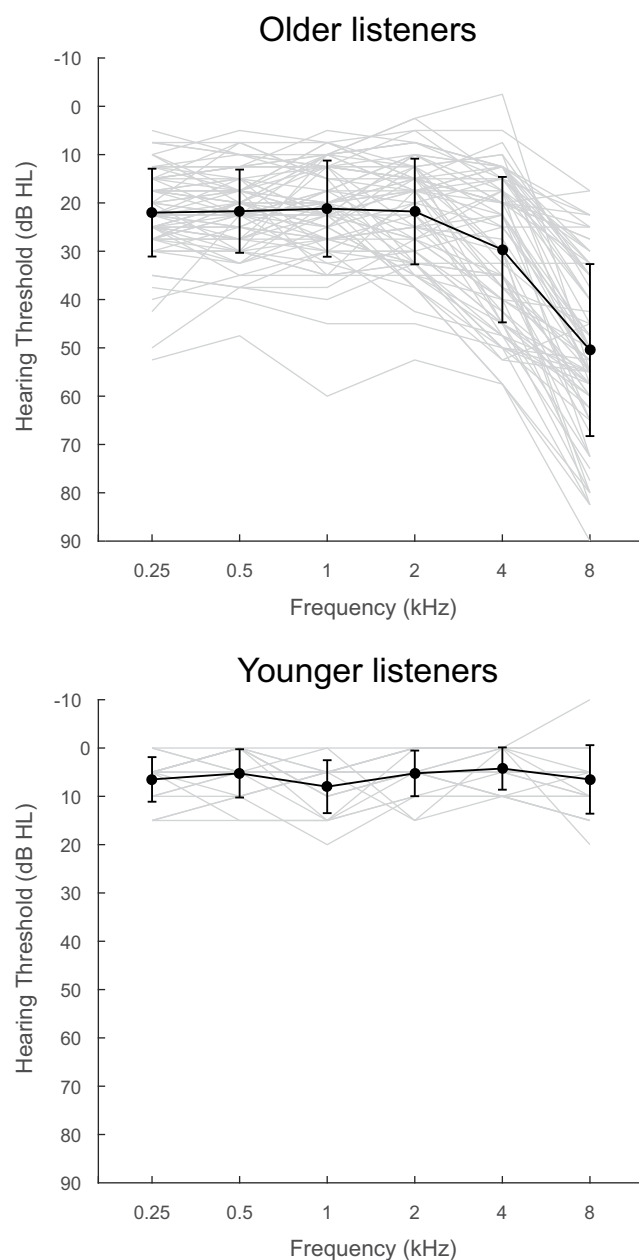


Figure 1 Hearing thresholds averaged across both ears in older and younger listeners. Grey lines represent each individual's hearing thresholds, and black dots indicate the mean hearing thresholds with error bars reflecting ± 1 standard deviation at octave-scale frequencies.

Speech Recognition in Adverse Listening Conditions

Table 2 and Figure 3 show the SRSs in four listening conditions for younger and older listeners. To evaluate speech recognition in given listening conditions between two age groups, the speech recognition performance was used as a dependent variable in a mixed ANOVA with a between-subject (age group) and a within-subject (listening condition) factor. Since the SRSs of younger

Table 1 Mean Working Memory Task Scores and Standard Deviation for Younger and Older Listeners

	DFS	DBS	WS	LS
Younger listeners	7.15 (1.39)	5.65 (1.18)	4.80 (0.70)	4.08 (0.41)
Older listeners	5.61 (1.26)	3.92 (1.00)	3.74 (0.68)	2.69 (0.92)

Abbreviations: DFS, Digit Forward Span; DBS, Digit Backward Span; WS, Word Span; LS, Listening Span.

participants were at the ceiling for the natural sentences and the sentences with a 30% TC, both speech recognition scores were excluded from the ANOVA. The main effect for age group ($F_{1,94} = 29.639$, $p < 0.001$) and listening condition ($F_{1,94} = 155.304$, $p < 0.001$, Greenhouse-Geisser correction), were significant. The interaction between age group and the listening condition was also significant ($F_{1,94} = 5.114$, $p = 0.026$, Greenhouse-Geisser correction).

Association Between Auditory Working Memory and Speech Recognition

To predict speech recognition performance in given listening conditions, auditory working memory capacity as measured using four tasks (DFS, DBS, WS, LS), age, and PTA were used in a stepwise multiple regression model. The unstandardized and standardized coefficients of the predictors and zero-order/partial correlations with dependent variables (SRSs) are shown in Table 3 for older listeners. Only WS ($\beta = 0.425$, $t = 4.043$, $p < 0.001$) significantly predicted the speech recognition performance on natural sentences ($R^2 = 0.181$), while both LS ($\beta = 0.359$, $t = 3.28$, $p = 0.002$) and DBS ($\beta = 0.265$, $t = 2.42$, $p = 0.018$) significantly predicted the speech recognition performance on sentences with a 30% TC ($R^2 = 0.28$). When the speech recognition performance on sentences with multi-talker

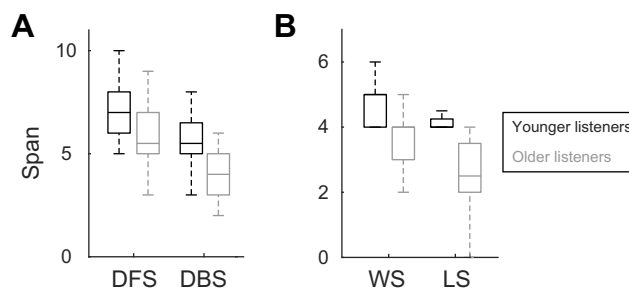


Figure 2 Auditory working memory capacity. **A.** Digit span (digit forward span: DFS, digit backward span: DBS) task. **(B).** Speech span (word span: WS, listening span: LS) task for younger and older listeners.

Table 2 Mean Sentence Recognition Scores (SRSs) and Standard Deviation in Four Listening Conditions for Younger and Older Listeners

	SRS1	SRS2	SRS3	SRS4
Younger listeners	100 (0)	100 (0)	94.88 (3.53)	84.56 (6.56)
Older listeners	99.38 (1.27)	98.52 (2.19)	84.05 (11.20)	69.16 (11.47)

Notes: SRS1: Natural Sentences, SRS2: Sentences with a 30% Time-Compression, SRS3: Sentences with Multi-Talker Noise, SRS4: Sentences with Multi-Talker Noise and a 30% Time-Compression

noise was predicted, PTA ($\beta = -0.319$, $t = -3.322$, $p = 0.001$), age ($\beta = -0.308$, $t = -3.337$, $p = 0.001$), and LS ($\beta = 0.312$, $t = 3.43$, $p = 0.001$) were the only significant predictors in the model ($R^2 = 0.489$). When predicting the speech recognition performance on sentences with multi-talker noise and a 30% TC, PTA ($\beta = -0.311$, $t = -3.385$, $p = 0.001$), age ($\beta = -0.366$, $t = -4.129$, $p < 0.001$), and LS ($\beta = 0.299$, $t = 3.428$, $p = 0.001$) were the only significant predictors in the model ($R^2 = 0.53$). Figure 4 displays the main results of the regression analysis suggesting that auditory working memory measured by LS, as well as PTA and age, primarily predict speech recognition performance in most adverse listening conditions. In addition, we examined if the relationship between a predictor and speech recognition performance is affected by the other predictors. For instance, the relationship between LS and speech recognition performance may exist whilst controlling for age and PTA. Partial correlations were obtained to determine the unique relationship between each predictor and speech recognition performance. Significant partial correlations were found between individual predictors (LS, age, PTA) and speech

recognition performance whilst controlling for the other two predictors (Table 3). Figure 5 shows the scatterplots of the partial correlation between the residuals of these predictors and the residuals of speech recognition performance. As a post hoc analysis, we divided older listeners into normal-hearing (ONH, mean 13.96 dB HL, 41 participants) whose PTA was at or better than 20 dB HL and hearing-impaired (OHI, mean 27.96 dB HL, 35 participants) group whose PTA was worse than 20 dB HL to see whether speech recognition performance was predicted by auditory working memory in both groups. ONH and OHI listeners had no difference in their LS ($t = 1.965$, $p = 0.0532$) but significantly differ in age ($t = -2.418$, $p = 0.0181$). Speech recognition performance on sentences with multi-talker noise ($t = 3.737$, $p < 0.001$) and on sentences with multi-talker noise and a 30% TC ($t = 3.965$, $p < 0.001$) were significantly different across the two older listener groups. However, only OHI group showed the dependence on LS. To be more specific, for ONH group, their PTA ($\beta = -0.435$, $t = -3.017$, $p = 0.004$) significantly predicted the speech recognition performance on sentences with multi-talker noise in the model ($R^2 = 0.189$) while age ($\beta = -0.499$, $t = -3.594$, $p = 0.001$) was the only significant predictors in the model ($R^2 = 0.249$) when predicting the speech recognition performance on sentences with multi-talker noise and a 30% TC. For OHI group, their LS ($\beta = 0.392$, $t = 2.943$, $p = 0.006$), as well as, age ($\beta = -0.469$, $t = -3.517$, $p = 0.001$) significantly predicted the speech recognition performance on sentences with multi-talker noise in the model ($R^2 = 0.468$) while LS ($\beta = 0.35$, $t = 2.548$, $p = 0.016$) and age ($\beta = -0.304$, $t = -2.171$, $p = 0.038$) were significant predictors in the model ($R^2 = 0.518$) in predicting the speech recognition performance on sentences with multi-talker noise and a 30% TC. For younger listeners, none of the predictors were significant except PTA ($\beta = -0.485$, $t = -2.355$, $p = 0.030$) for speech recognition performance on sentences with multi-talker noise ($R^2 = 0.236$).

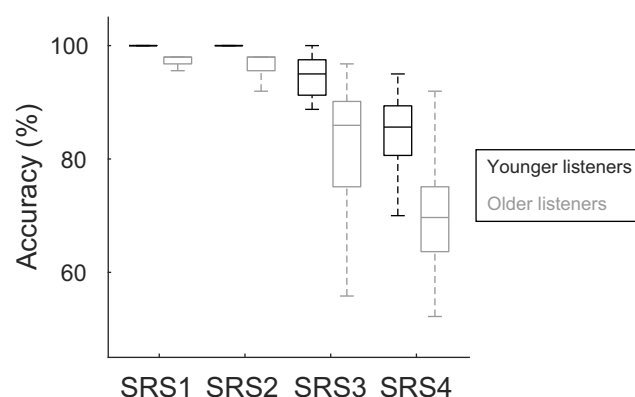


Figure 3 Sentence recognition scores (SRSs) in four listening conditions (SRS1: natural sentences, SRS2: sentences with a 30% time-compression, SRS3: sentences with multi-talker noise, SRS4: sentences with multi-talker noise and a 30% time-compression) for younger and older listeners.

Discussion

The main goal of this study was to describe the association between auditory working memory and speech recognition in unfavorable listening conditions for older listeners, in a systematic way, with multiple working memory tasks and various listening conditions for speech recognition tests. We found that auditory working memory, measured by LS, can predict speech recognition performance in adverse listening conditions driven by time-compression and multi-talker noise for older listeners even after controlling for the impacts of age and hearing sensitivity, but we did not find this association for younger listeners.

Table 3 Stepwise Regression Results for Older Listeners. The Dependent Variables are Speech Recognition Performance on Natural Sentences (SRS1), Sentences with a 30% Time-Compression (SRS2), Sentences with Multi-Talker Noise (SRS3), and Sentences with Multi-Talker Noise and a 30% Time-Compression (SRS4), Respectively

Model	B	SE B	β	t	p	Zero-Order Correlations	Partial Correlations
Dependent: SRS1, $R^2 = 0.181$							
Constant	96.417	0.744		129.675	<0.001		
WS	0.792	0.196	0.425	4.043	<0.001	0.425	0.425
Dependent: SRS2, $R^2 = 0.28$							
Constant	93.948	0.928		101.246	<0.001		
LS	0.857	0.261	0.359	3.28	0.002	0.471	0.358
DBS	0.578	0.239	0.265	2.42	0.018	0.417	0.273
Dependent: SRS3, $R^2 = 0.489$							
Constant	125.010	12.972		9.637	<0.001		
PTA	-0.437	0.132	-0.319	-3.322	0.001	-0.552	-0.365
Age	-0.614	0.184	-0.308	-3.337	0.001	-0.51	-0.366
LS	3.808	1.11	0.312	3.43	0.001	0.501	0.375
Dependent: SRS4, $R^2 = 0.53$							
Constant	119.440	12.747		9.370	<0.001		
PTA	-0.437	0.129	-0.311	-3.385	0.001	-0.563	-0.371
Age	-0.746	0.181	-0.366	-4.129	<0.001	-0.561	-0.438
LS	3.741	1.091	0.299	3.428	0.001	0.5	0.375

Note: PTA: The Pure-Tone Average Across 0.5, 1, 2, 4 kHz at the Better Ear.

Abbreviations: DBS, Digit Backward Span, WS, Word Span, LS, Listening Span.

Auditory Working Memory Predicts Speech Recognition

Predicting speech recognition performance in unfavorable listening conditions may depend on the modality of working

memory tasks. Auditory working memory, measured by LS, showed a significant correlation with speech recognition performance in noise as well as with fast speech. In linear regression models used in the present study, when hearing sensitivity

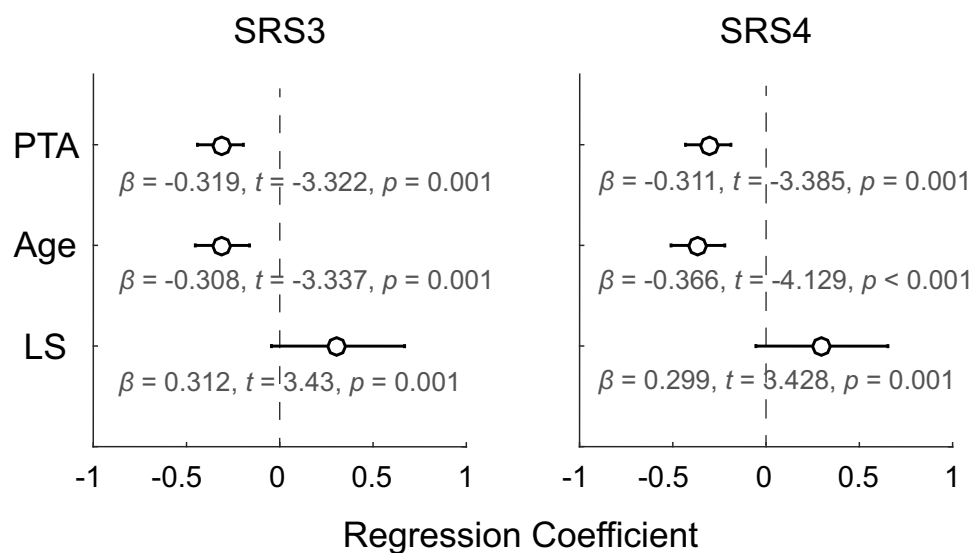


Figure 4 Graph of regression coefficients showing a significant relationship between three variables (PTA, age, LS) and speech recognition performance (SRS3: sentences with multi-talker noise, SRS4: sentences with multi-talker noise and a 30% time-compression) in older listeners. Error bars reflect ± 1 standard error. PTA: the pure-tone average across 0.5, 1, 2, 4 kHz at the better ear.

Abbreviation: LS, listening span.

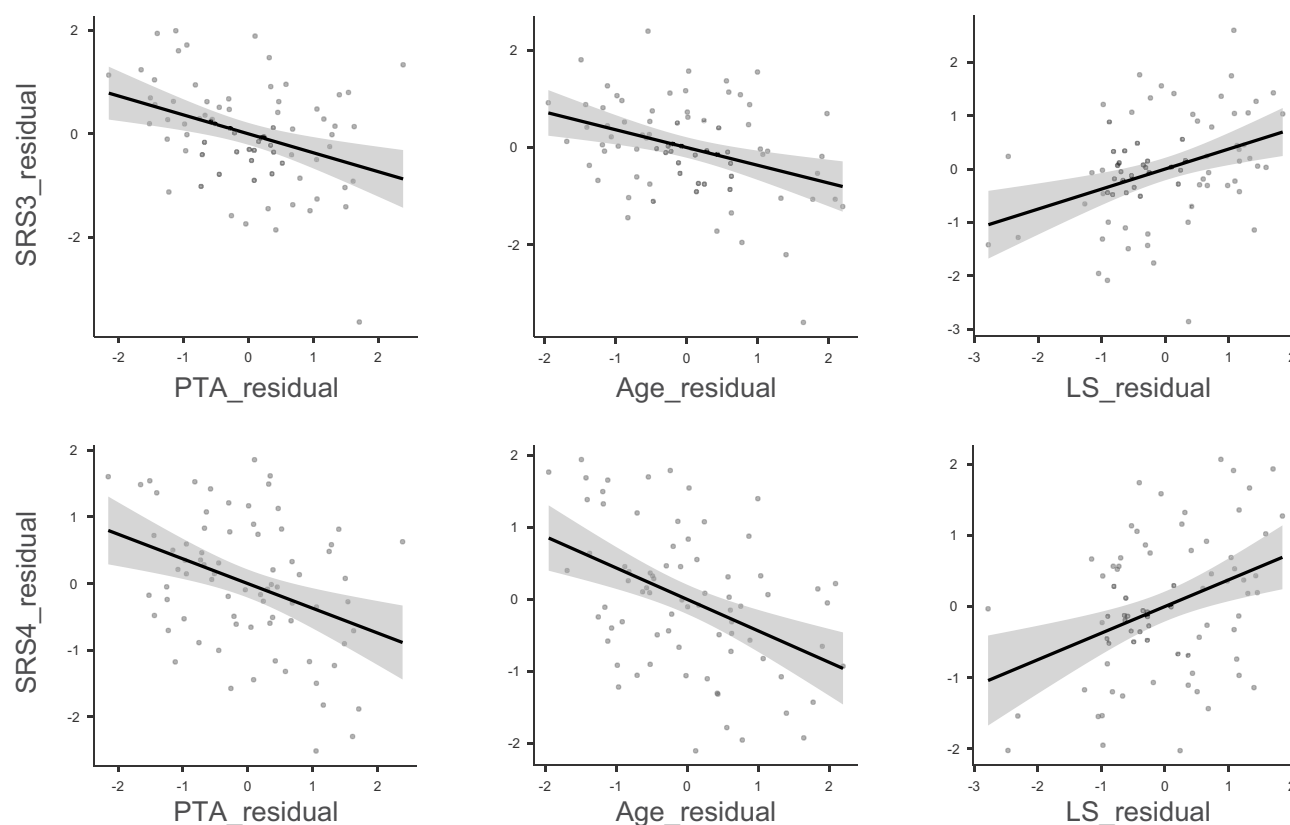


Figure 5 Partial correlation scatter plots where the dependent variables are the residual of the speech recognition performance on sentences with multi-talker noise (SRS3) (top row) and sentences with multi-talker noise and a 30% time-compression (SRS4) (bottom row) for older listeners. Each predictor (PTA, age, LS) has a significant partial correlation to speech recognition performance whilst controlling for the effect of the other two predictors. PTA: the pure-tone average across 0.5, 1, 2, 4 kHz at the better ear. **Abbreviation:** LS, listening span.

(PTA) and age were controlled, auditory working memory still accounted for individual differences in speech recognition performance in adverse listening conditions. As the working memory task was presented in the auditory modality in which speech recognition was also measured, significant correlations were found between two tasks in the present study. Exploring the association between working memory and speech recognition in the same (auditory) modality in the present study is in accordance with a recent study that develops the Word Auditory Recognition and Recall Measure (WARRM).¹⁸ Smith, et al¹⁸ demonstrated that the WARRM that incorporates auditory working memory task and speech recognition test is more feasible, reliable, and ecologically relevant. In addition, the WARRM measures the intraindividual difference in working memory for speech recognition across various listening conditions in a given subject, which may not be able to be measured when the working memory task is presented in the visual modality. Since the present study did not make direct

comparisons between auditory and visual working memory tasks, the relative usefulness of the auditory modality cannot be concluded. Prior studies showed that stronger correlations were found across tasks with the same sensory input than similar tasks that tested different modalities.³⁶ Behavioral studies reveal a discrepancy in LS performance between younger and older adults with normal or near-normal hearing, but not in RS performance.^{16,19} This supports the idea that auditory working memory tasks may be more sensitive to predicting the difficulty in speech recognition in older listeners. fMRI studies also support the modality-specific difference by revealing that different brain regions are involved in different modality tasks; auditory n-back tasks engaged the left hemisphere dorsolateral prefrontal cortex, while the left hemisphere posterior parietal cortex was activated during visual n-back tasks.^{20,21} Crottaz-Herbette and others also revealed bilateral cross-modal inhibition (auditory/visual cortex activity decreases during visual/auditory working memory), supporting the utility of

auditory working memory for predicting (auditory) speech recognition performance.

Several studies show that hearing loss and age play primary roles in predicting unaided speech recognition performance for older listeners.^{3,4} A review from Akeroyd suggests that working memory, mostly measured by visual tasks, has only a secondary effect on speech recognition. Also, recent studies using visual working memory tasks show that the correlation between working memory and speech recognition performance in adverse listening conditions becomes insignificant after controlling for age.^{4,22,37} These results may imply that a decline in visual working memory merely reflects the general cognitive decline in older listeners. However, auditory working memory has the unique ability to predict speech recognition performance in adverse listening conditions in the present study. Our linear regression model indicates that auditory working memory can still explain the variance of speech recognition performance in given listening conditions even after controlling for the impacts of age and hearing sensitivity. Although our results do not show an increase in the predictive effect of auditory working memory as the listening condition becomes harder, auditory working memory has consistent, significant effects across the listening conditions that involve multi-talker noise and TC. These findings may imply that auditory working memory tasks are useful tools to predict older listeners' speech recognition performance in unfavorable listening conditions. Recent studies found auditory working memory tests presented with fully audible words useful, showing that hearing aid signal processing can provide more cognitive spare capacity that is crucial in learning and auditory rehabilitation.^{38,39}

Systematic Approach to the Association Between Auditory Working Memory and Speech Recognition

Working memory capacity declined significantly in older listeners in both digit and speech span tasks. However, only LS showed a correlation with speech recognition performance in older listeners in more adverse conditions. The linguistic features of the auditory stimuli used in the working memory tasks may contribute to the association between working memory and speech recognition. LS uses multisyllabic words (two or more than two syllables) in the last (target) word position, while the digit span and word span tasks use monosyllabic and bisyllabic words,

respectively. In addition, since LS contains sentence-level linguistic information, it may better reflect lexical complexity that listeners need to utilize to recognize sentences in adverse listening conditions.^{40,41} Heinrich, Henshaw, Ferguson⁴² showed that the association between cognition (working memory) and speech perception could be affected by the linguistic complexity of speech material (digit vs sentence). In addition, the complexity of the working memory task can determine the association between working memory and speech recognition performance. In the present study, LS is included as the most complex span task that represents processing as well as storage of auditory information. The result is consistent with the finding that adult listeners' working memory measured by complex span tasks better predicts speech recognition in adverse listening conditions.¹⁴ However, we found that older listeners did not necessarily show a stronger association between LS and speech recognition in harder conditions (sentences with multi-talker noise vs sentences with multi-talker noise and TC). These results are not consistent with the findings from the ELU-model⁴³ that predicted higher involvement of working memory under adverse listening conditions for speech recognition, but consistent with recent studies^{24,44} that included participants with a narrow age range or the control for age.

Different Associations Between Auditory Working Memory and Speech Recognition in Younger and Older Listeners

Speech recognition performance significantly dropped when the signal was degraded by multi-talker noise or TC in both younger and older listeners. However, it was only the older listener group that showed associations between auditory working memory and speech recognition performance in these listening conditions. This is consistent with studies that show little association between visual working memory and speech recognition for younger listeners, but a strong association for older listeners (Füllgrabe, Rosen²⁴). The present study demonstrates that working memory has different contributions to predicting speech recognition in adverse listening conditions for younger and older listeners even when the target speech signal and working memory tasks share the same modality (auditory). It is also interesting that OHI group in the present study showed more dependence on auditory working memory compared to ONH group, who had better hearing sensitivity and was relatively younger, although

both groups had no significant difference in their working memory ability measured by LS. The increase in the contribution of working memory with age and hearing-impairment may result from the loss of sensitivity to temporal cues.^{45,46} Unfortunately, otoacoustic emissions tests were not conducted in the present study due to the time constraint and the lack of equipment. Older listeners in general may have the decline of the medial olivocochlear system with age, that precedes outer hair cell degeneration and occurs before the change in hearing sensitivity,⁴⁷ or the loss of outer hair cell function.⁴⁸ In addition, it is also possible that older listeners' deficits in supra-threshold auditory processing strongly engage auditory working memory.²⁴ Older listeners may have poor acoustic representation, despite normal hearing thresholds, due to age-related loss of neural coding fidelity.⁴⁹ Due to incomplete supra-threshold auditory processing in older listeners, they may employ different cognitive strategies during speech recognition tasks in adverse listening conditions. In other words, older listeners may need more involvement of auditory working memory to compensate for auditory processing deficits. Alternatively, for younger listeners, the same tasks might require less cognitive engagement. These younger listeners might need more challenging conditions to be more dependent on auditory working memory capacity.⁴ Gordon-Salant & Cole²⁶ showed a strong contribution of working memory to speech recognition in noise in both younger and older listeners when working memory capacity and hearing sensitivity between the two groups were matched. Therefore, different associations between working memory and speech recognition in younger and older listeners in the present study may stem from significantly different working memory capacities and hearing sensitivity between the two groups.

Conclusions

Older listeners' auditory working memory capacity predicts speech recognition in unfavorable listening conditions after controlling for the impact of age and hearing sensitivity. The association between auditory working memory and speech recognition performance depends on the type of working memory tasks, listening conditions, and participant population. Our findings suggest that understanding modality-specific characteristics of working memory may provide better insight into the difficulty of speech recognition in older listeners and successful hearing intervention.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and National Research Foundation of Korea (NRF2016S1A5A8020353).

Disclosure

The authors report no conflicts of interest in this work.

References

- Pichora-Fuller MK, Souza PE. Effects of aging on auditory processing of speech. *Int J Audiol*. 2003;42(sup2):11–16. doi:10.3109/14992020309074638
- Gordon-Salant S, Fitzgibbons PJ. Temporal factors and speech recognition performance in young and elderly listeners. *J Speech Hear Res*. 1993;36(6):1276–1285. doi:10.1044/jshr.3606.1276
- Akeroyd MA. Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int J Audiol*. 2008;47(sup2):53–71. doi:10.1080/14992020802301142
- Besser J, Koelewijn T, Zekveld AA, Kramer SE, Festen JM. How linguistic closure and verbal working memory relate to speech recognition in noise: a review. *Trends Amplification*. 2013;17(2):75–93. doi:10.1177/1084713813495459
- Rabbitt PM. Channel-capacity, intelligibility and immediate memory. *Q J Exp Psychol*. 1968;20(3):241–248. doi:10.1080/14640746808400158
- Rabbitt PM. Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngol Suppl*. 1991;476(sup476):167–176. doi:10.3109/00016489109127274
- Lunner T. Cognitive function in relation to hearing aid use. *Int J Audiol*. 2003;42:49–58. doi:10.3109/14992020309074624
- Rudner M, Fransson P, Ingvar M, Nyberg L, Rönnerberg J. Neural representation of binding lexical signs and words in the episodic buffer of working memory. *Neuropsychologia*. 2007;45(10):2258–2276. doi:10.1016/j.neuropsychologia.2007.02.017
- Pichora-Fuller MK, Kramer SE, Eckert MA, et al. Hearing impairment and cognitive energy: the Framework for Understanding Effortful Listening (FUEL). *Ear Hear*. 2016;37 Suppl 1:5s–27s. doi:10.1097/AUD.0000000000000312
- Baddeley A. The episodic buffer: a new component of working memory? *Trends Cognit Sci*. 2000;4(11):417–423. doi:10.1016/S1364-6613(00)01538-2
- Baddeley A. *Working Memory*. New York, NY: Clarendon Press/Oxford University Press; 1986.
- Daneman M, Carpenter PA. Individual differences in working memory and reading. *J Verbal Learn Verbal Behav*. 1980;19(4):450–466. doi:10.1016/S0022-5371(80)90312-6
- Rönnerberg J, Holmer E, Rudner M. Cognitive hearing science and ease of language understanding. *Int J Audiol*. 2019;58(5):247–261. doi:10.1080/14992027.2018.1551631
- Rönnerberg J, Lunner T, Zekveld A, et al. *The Ease of Language Understanding (ELU) Model Theoretical, Empirical, and Clinical Advances*. 2013.
- Arehart KH, Souza P, Baca R, Kates JM. Working memory, age, and hearing loss susceptibility to hearing aid distortion. *Ear Hear*. 2013;34(3):251–260. doi:10.1097/AUD.0b013e318271aa5e
- Pichora-Fuller MK, Schneider BA, Daneman M. How young and old adults listen to and remember speech in noise. *J Acoust Soc Am*. 1995;97(1):593–608. doi:10.1121/1.412282
- Smith SL, Pichora-Fuller MK. Associations between speech understanding and auditory and visual tests of verbal working memory: effects of linguistic complexity, task, age, and hearing loss. *Front Psychol*. 2015;6:1394. doi:10.3389/fpsyg.2015.01394

18. Smith SL, Pichora-Fuller MK, Alexander G. Development of the word auditory recognition and recall measure a working memory test for use in rehabilitative audiology. *Ear Hear.* 2016;37(6):e360–e376. doi:10.1097/AUD.0000000000000329
19. Baldwin CL, Ash IK. Impact of sensory acuity on auditory working memory span in young and older adults. *Psychol Aging.* 2011;26(1):85–91. doi:10.1037/a0020360
20. Crottaz-Herbette S, Anagnoson RT, Menon V. Modality effects in verbal working memory: differential prefrontal and parietal responses to auditory and visual stimuli. *NeuroImage.* 2004;21(1):340–351. doi:10.1016/j.neuroimage.2003.09.019
21. Rodriguez-Jimenez R, Avila C, Garcia-Navarro C, et al. Differential dorsolateral prefrontal cortex activation during a verbal n-back task according to sensory modality. *Behav Brain Res.* 2009;205(1):299–302. doi:10.1016/j.bbr.2009.08.022
22. Besser J, Zekveld AA, Kramer SE, Rönnberg J, Festen JM. New measures of masked text recognition in relation to speech-in-noise perception and their associations with age and cognitive abilities. *J Speech Lang Hear Res.* 2012;55(1):194–209. doi:10.1044/1092-4388(2011/11-0008)
23. Zekveld AA, Rudner M, Johnsrude IS, Rönnberg J. The effects of working memory capacity and semantic cues on the intelligibility of speech in noise. *J Acoust Soc Am J Acoust Soc Am.* 2013;134(3):2225–2234. doi:10.1121/1.4817926
24. Füllgrabe C, Rosen S. On the (Un)importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Front Psychol.* 2016;7.
25. Dryden A, Allen HA, Henshaw H, Heinrich A. The association between cognitive performance and speech-in-noise perception for adult listeners: a systematic literature review and meta-analysis. *Trends Hear.* 2017;21.
26. Gordon-Salant S, Cole SS. Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing. *Ear Hear.* 2016;37(5):593–602. doi:10.1097/AUD.0000000000000316
27. Lee D-Y, Lee KU, Lee JH, Kim KW, Jhoo JH, Youn J. A normative study of the mini-mental state examination in the Korean elderly. *J Korean Neuropsychiatr Assoc.* 2002;41:508–525.
28. O'Bryant SE, Humphreys JD, Smith GE, et al. Detecting dementia with the mini-mental state examination in highly educated individuals. *Arch Neurol.* 2008;65(7):963–967. doi:10.1001/archneur.65.7.963
29. Kim J-S, Lim D, Hong H-N, et al. Development of Korean Standard Monosyllabic Word Lists for Adults (KS-MWL-A). *Audiol.* 2008;4(2):126–140.
30. Bahng J, Lee J. Hearing thresholds for a geriatric population composed of Korean males and females. *J audiol Otol.* 2015;19(2):91–96. doi:10.7874/jao.2015.19.2.91
31. Cho S-J, Lim D, Lee K-W, Han H-K, Lee J-H. Development of Korean standard bisyllabic word list for adults used in speech recognition threshold test. *Audiol.* 2008;4(1):28–36.
32. Lee B-T. Reliability of the reading span test. *Psychol Sci.* 2002;11(1):15–33.
33. Jang H, Lee J, Lim D, Lee K, Jeon A, Jung E. Development of Korean standard sentence lists for sentence recognition tests. *Audiol.* 2008;4(2):161–177.
34. Wingfield A, Tun PA, Koh CK, Rosen MJ. Regaining lost time: adult aging and the effect of time restoration on recall of time-compressed speech. *Psychol Aging.* 1999;14(3):380–389. doi:10.1037/0882-7974.14.3.380
35. Kawahara H, Masuda-Katsuse I, de Cheveigne A. Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: possible role of a repetitive structure in sounds. *Speech Commun.* 1999;27(3):187. doi:10.1016/S0167-6393(98)00085-5
36. Humes LE, Burk MH, Coughlin MP, Busey TA, Strauser LE. Auditory speech recognition and visual text recognition in younger and older adults: similarities and differences between modalities and the effects of presentation rate. *J Speech Lang Hear Res.* 2007;50(2):283–303. doi:10.1044/1092-4388(2007/021)
37. Koelewijn T, Zekveld AA, Festen JM, Rönnberg J, Kramer SE. Processing load induced by informational masking is related to linguistic abilities. *Int J Otolaryngol.* 2012;2012:1–11. doi:10.1155/2012/865731
38. Rudner M. Cognitive spare capacity as an index of listening effort. *Ear Hear.* 2016;37(Suppl 1):69s–76s. doi:10.1097/AUD.0000000000000302
39. Lunner T, Rudner M, Rosenbom T, Agren J, Ng EH. Using speech recall in hearing aid fitting and outcome evaluation under ecological test conditions. *Ear Hear.* 2016;37 Suppl 1:145s–154s. doi:10.1097/AUD.0000000000000294
40. McElree B, Foraker S, Dyer L. Memory structures that subserve sentence comprehension. *J Mem Lang.* 2003;48(1):67. doi:10.1016/S0749-596X(02)00515-6
41. Zekveld AA, Kramer SE, Festen JM. Cognitive load during speech perception in noise: the influence of age, hearing loss, and cognition on the pupil response. *Ear Hear.* 2011;32(4):498–510. doi:10.1097/AUD.0b013e31820512bb
42. Heinrich A, Henshaw H, Ferguson MA. The relationship of speech intelligibility with hearing sensitivity, cognition, and perceived hearing difficulties varies for different speech perception tests. *Front Psychol.* 2015;6.
43. Rönnberg J, Rudner M, Lunner T, Zekveld AA. When cognition kicks in: working memory and speech understanding in noise. *Noise Health.* 2010;12(49):263–269. doi:10.4103/1463-1741.70505
44. Nuesse T, Steenken R, Neher T, Holube I. Exploring the link between cognitive abilities and speech recognition in the elderly under different listening conditions. *Front Psychol.* 2018;9.
45. Fullgrabe C. Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. *Am J Audiol.* 2013;22(2):313–315. doi:10.1044/1059-0889(2013/12-0070)
46. Fullgrabe C, Meyer B, Lorenzi C. Effect of cochlear damage on the detection of complex temporal envelopes. *Hear Res.* 2003;178(1–2):35–43. doi:10.1016/S0378-5955(03)00027-3
47. Kim S, Frisina DR, Frisina RD. Effects of age on contralateral suppression of distortion product otoacoustic emissions in human listeners with normal hearing. *Audiol Neurotol.* 2002;7(6):348–357. doi:10.1159/000066159
48. Hoben R, Easow G, Pevzner S, Parker MA. Outer hair cell and auditory nerve function in speech recognition in quiet and in background noise. *Front Neurosci.* 2017;11:157. doi:10.3389/fnins.2017.00157
49. Bharadwaj HM, Masud S, Mehraei G, Verhulst S, Shinn-Cunningham BG. Individual differences reveal correlates of hidden hearing deficits. *J Neurosci.* 2015;35(5):2161–2172. doi:10.1523/JNEUROSCI.3915-14.2015

Clinical Interventions in Aging

Dovepress

Publish your work in this journal

Clinical Interventions in Aging is an international, peer-reviewed journal focusing on evidence-based reports on the value or lack thereof of treatments intended to prevent or delay the onset of maladaptive correlates of aging in human beings. This journal is indexed on PubMed Central, MedLine, CAS, Scopus and the Elsevier

Bibliographic databases. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/clinical-interventions-in-aging-journal>