

## Research Article

# Working Memory and Speech Recognition in Noise Under Ecologically Relevant Listening Conditions: Effects of Visual Cues and Noise Type Among Adults With Hearing Loss

Christi W. Miller,<sup>a</sup> Erin K. Stewart,<sup>a</sup> Yu-Hsiang Wu,<sup>b</sup> Christopher Bishop,<sup>a</sup>  
Ruth A. Bentler,<sup>b</sup> and Kelly Tremblay<sup>a</sup>

**Purpose:** This study evaluated the relationship between working memory (WM) and speech recognition in noise with different noise types as well as in the presence of visual cues.

**Method:** Seventy-six adults with bilateral, mild to moderately severe sensorineural hearing loss (mean age: 69 years) participated. Using a cross-sectional design, 2 measures of WM were taken: a reading span measure, and Word Auditory Recognition and Recall Measure (Smith, Pichora-Fuller, & Alexander, 2016). Speech recognition was measured with the Multi-Modal Lexical Sentence Test for Adults (Kirk et al., 2012) in steady-state noise and 4-talker

babble, with and without visual cues. Testing was under unaided conditions.

**Results:** A linear mixed model revealed visual cues and pure-tone average as the only significant predictors of Multi-Modal Lexical Sentence Test outcomes. Neither WM measure nor noise type showed a significant effect.

**Conclusion:** The contribution of WM in explaining unaided speech recognition in noise was negligible and not influenced by noise type or visual cues. We anticipate that with audibility partially restored by hearing aids, the effects of WM will increase. For clinical practice to be affected, more significant effect sizes are needed.

Compared with their peers with typical hearing, individuals with sensorineural hearing loss often experience difficulty in understanding speech in noise. Despite appropriately fit hearing aids (HAs) and increased audibility, many people with HAs continue to experience difficulty in these situations, with only half of listeners reporting satisfaction with their HA performance in noise (Kochkin, 2000). This poor performance in noise has the potential to lead to social isolation, frustration, and, ultimately, discontinued HA use, which can lead to an overall decrease in quality of life (Chia et al., 2007). For these reasons, it is essential to better understand the

elements associated with poor speech understanding in noise.

Speech understanding in noise is affected by many different factors, both auditory and nonauditory (George et al., 2007). Although sensorineural hearing loss is primarily characterized by sensory impairments such as increased thresholds and distortion of the auditory signal, there are also many cognitive factors that influence an individual's ability to understand speech, especially in the presence of background noise, reverberation, or other difficult listening situations (e.g., Working Group on Speech Understanding and Aging, Committee on Hearing, Bioacoustics, and Biomechanics, 1988; Gatehouse, Naylor, & Elberling, 2003; Humes, 2002). From a review of 20 studies, Akeroyd (2008) concluded that, in general, there is a link between cognitive abilities and speech recognition, but this is usually secondary to the predictive effects of hearing loss. Among a wide range of cognitive tests considered (e.g., nonverbal and verbal IQ, Scholastic Aptitude Test scores, attention, and memory), many were correlated with speech recognition, but results were mixed, and not all tests had high predictive power. Working memory (WM) abilities were shown

<sup>a</sup>Department of Speech and Hearing Sciences, University of Washington, Seattle

<sup>b</sup>Department of Communication Sciences and Disorders, University of Iowa, Iowa City

Correspondence to Christi W. Miller: christim@u.washington.edu

Editor: Frederick Gallun

Associate Editor: Mitchell Sommers

Received July 12, 2016

Revision received September 23, 2016

Accepted February 4, 2017

[https://doi.org/10.1044/2017\\_JSLHR-H-16-0284](https://doi.org/10.1044/2017_JSLHR-H-16-0284)

**Disclosure:** The authors have declared that no competing interests existed at the time of publication.

to be the most consistent predictive factors for speech understanding in noise, especially when measured using a reading span measure (RS; Daneman & Carpenter, 1980).

WM refers to the temporary storage system of maintaining information during ongoing processing (Baddeley & Hitch, 1974). WM is thought to be important for speech understanding because listeners are required to process the incoming speech signal and integrate it with stored knowledge while simultaneously anticipating upcoming signals (Daneman & Carpenter, 1980, 1983). People with hearing loss may need to allocate some of their WM resources to understanding a degraded incoming signal, therefore reducing their available capacity to process new information (e.g., Lunner, 2003; Rönnberg et al., 2013; van Rooij & Plomp, 1990). In fact, a robust relationship between WM and speech recognition in noise has been previously established in listeners with hearing loss (e.g., Lunner, 2003; Neher et al., 2009; Ng, Rudner, Lunner, Pedersen, & Rönnberg, 2013; Rudner, Foo, Rönnberg, & Lunner, 2009; Rudner, Rönnberg, & Lunner, 2011; Souza & Arehart, 2015). In general, people with higher WM capacity tend to have better speech understanding in noise than those with lower WM capacity, even when controlling for age and degree of hearing loss (e.g., Gatehouse et al., 2003; Lunner & Sundewall-Thorén, 2007; Souza & Arehart, 2015).

The purpose of this study was to determine if the relationship between WM and speech recognition in noise is maintained under test conditions that more closely resemble real-world listening conditions than those used in prior studies. Although previous research demonstrated significant results, the effect sizes were small (approximately 2%–17%; Ng et al., 2014; Souza & Arehart, 2015; see reviews in Akeroyd, 2008; Füllgrabe & Rosen, 2016), and it is unknown whether or not these effects will hold in more realistic environments. It is common for only a single target talker to be used (e.g., a single female voice in the Quick Speech-in-Noise test [QuickSIN]; Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004), which can lead to the learning of indexical cues throughout the test among other traits (e.g., Broadbent, 1952; Mullennix, Pisoni, & Martin, 1989). Such learning could have influenced speech understanding scores previously reported in the literature. Indexical cues refer to voice-specific features that are influenced by age, gender, regional dialect, native language, and anatomical differences between people, such as variations in the length and resonance of the vocal tract. The learning of indexical cues leads to better speech-recognition results and faster response times than when indexical cues are not available (Mullennix et al., 1989; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994). This is an important aspect to consider, because listeners encounter multiple talkers throughout the day or even within a single conversation. In fact, performance on understanding multiple target talkers (i.e., a different person speaking each sentence) correlates with self-reported hearing abilities, but not with single-target-talker conditions (Kirk, Pisoni, & Miyamoto, 1997). We therefore sought to add to the existing literature by examining the predictive ability of WM capacity to speech understanding under

multiple-target-talker conditions (e.g., the target talker varies from trial to trial). If similar or stronger WM effect sizes can be demonstrated in more typical situations, substantial evidence will be provided to support measuring WM in the clinic.

Cognitive resources are taxed as variability in the speech signal increases. Tamati, Gilbert, and Pisoni (2013) explored factors underlying differences on a highly variable speech test similar to that used in the current study, and they found that high performers on their speech-recognition test were better not only on measures of indexical processing (e.g., gender discrimination and regional dialect categorization), but also on measures of WM capacity, vocabulary size, and executive functioning. Evidence also exists to support a higher demand on WM resources for multiple-target-talker situations than single-target-talker situations (Martin, Mullennix, Pisoni, & Summers, 1989). However, these effects were exhibited in young listeners with typical hearing, and hearing loss requires greater WM resources due to the degraded peripheral input (Schneider, Daneman, & Pichora-Fuller, 2002; Wingfield & Tun, 2001). These results are in line with the Ease of Language Understanding model, which proposes that a coherent stream of information is automatically and rapidly formed at the cognitive level and then implicitly matched to stored phonological representations in long-term memory (Rönnberg, 2003; Rönnberg, Rudner, Foo, & Lunner, 2008). A failure to match representations in the long-term memory may occur in suboptimal conditions, such as in noisy environments or with distortion from hearing loss or HAs, and force explicit processing to occur (Rönnberg, 2003; Rönnberg et al., 2008). In a multiple-target-talker situation, explicit processing is likely to be more engaged than implicit processing due to a greater mismatch in phonological representations of the stimuli to long-term memory.

When exploring the effects of WM on speech understanding in noise with more ecologically relevant stimuli, another feature worth considering is the type of background noise. Researchers have demonstrated an interaction between noise type and WM on speech understanding in noise (for a review, see Rönnberg, Rudner, Lunner, & Zekveld, 2010). In particular, higher WM capacity leads to better speech understanding performance when the background contains amplitude modulations, and WM plays a smaller role in performance with steady-state backgrounds (e.g., Gatehouse, Naylor, & Elberling, 2006; Gatehouse et al., 2003; Lunner & Sundewall-Thorén, 2007; Rudner et al., 2011). However, not all findings have supported such an interaction between noise type and WM (Cox & Xu, 2010; Foo, Rudner, Rönnberg, & Lunner, 2007), and the effects of noise on WM may be task dependent (i.e., open or closed set; Marrone, Alt, DeDe, Olson, & Shehorn, 2015). Therefore, we included steady-state noises (SSNs) and modulated noises in the current study, and we expected WM capacity to predict modulated noise performance to a greater extent than in SSN.

Perhaps the most important feature to consider with ecologically relevant stimuli is the effects of visual cues on

speech understanding in noise, which are often available in real-world situations and lead to better performance (e.g., Wu, Stangl, Zhang, & Bentler, 2015). Visual cues are exceptionally helpful when the background noise is also speech (Helfer & Freyman, 2005) and when audibility is reduced or the listener cannot take advantage of the audibility provided (e.g., Bernstein & Grant, 2009). For example, less HA bandwidth is required for speech recognition with the addition of visual cues (Silberer, Bentler, & Wu, 2015). In addition, Wu and Bentler (2010) found that benefit from directional microphones and preference for directional microphones on HAs were both reduced in the presence of visual cues, which implies that audibility has less of an effect on performance when visual cues are present. Regarding WM, some evidence exists to suggest that the availability of visual cues increases later recall of speech, particularly in the beginning and ending positions of trials (as opposed to the middle position), and presumably due to better encoding of the signal in short-term memory (Mishra, Lunner, Stenfelt, Rönnberg, & Rudner, 2013). Electrophysiological evidence has also been presented to support faster and more efficient brain processing in auditory-visual (AV) conditions compared with auditory-only (AO) conditions, and that WM facilitated this effect, as illustrated by earlier peak (P3) latencies and higher peak amplitudes (Frtusova and Phillips, 2016). Despite the evidence found by Frtusova and Phillips (2016), suggesting that visual cues relieved some of the processing demands on WM, listening effort may be higher in AV tasks under challenging conditions when speech performance is equal (Fraser, Gagné, Alepins, & Dubois, 2010). In fact, visual cues may actually impede speech recognition when they are not needed to perform the task (i.e., auditory cues are sufficient to understand the speech), and they may even be considered a distractor in young listeners with typical hearing (Mishra et al., 2013). Therefore, if the task demands a high level of cognitive resources (e.g., speech understanding with multiple target talkers), the additional resources required for integration of audio-visual information may reduce resources that could be allocated to processing the AO information. Because the listeners in the current study have hearing loss (i.e., the auditory signal will be degraded), the prediction was that although visual cues will enhance speech recognition, greater reliance on WM to reconstruct the bimodal peripheral input would be required. In other words, the WM-speech-in-noise relationship will be stronger with AV speech than AO speech.

The aim of this experiment was to determine if the relationship between WM and speech recognition in noise is maintained under more ecologically relevant listening conditions than those used in prior studies, using two types of background noise as well as the absence or presence of visual cues. We measured verbal WM capacity presented in the auditory domain and in the visual domain in adults with hearing loss. In the visual domain, the RS (Baddeley, Logie, Nimmo-Smith, & Brereton, 1985; Daneman & Carpenter, 1980; Rönnberg, Arlinger, Lyxell, & Kinnefors, 1989) has been most commonly used in our field, and it has been the best predictor of speech recognition in noise

as compared with other WM measures (e.g., for a review, see Akeroyd, 2008). However, some researchers have suggested that WM capacity may be a domain-specific function, and that when measured in the same domain as the outcome of interest, a stronger relationship will be observed (e.g., Smith & Pichora-Fuller, 2015; Smith, Pichora-Fuller, & Alexander, 2016). Smith et al. (2016) developed a WM measure delivered in the auditory domain, called the Word Auditory Recognition and Recall Measure (WARRM). Smith and Pichora-Fuller (2015) were not able to provide evidence to support the hypothesis that the WARRM would be more predictive of speech outcomes than the RS, but they suggested it may have been due to their study being underpowered. For this reason, we measured WM capacity using both the WARRM and RS tests.

## Method

### Participants

Fifty-three women and 23 men participated, with a mean age of 69 years ( $SD = 7.47$  years). The numbers of participants across decades were 38, 31, 6, 0, and 1 for the age ranges 70–79, 60–69, 50–59, 40–49, and 30–39 years, respectively. These participants were part of a larger study on HA outcomes. Adults with HAs were recruited from participation pool databases across two sites: the University of Washington and the University of Iowa. Approximately 366 participants were identified across all databases who were between the ages of 21 and 79 and who had a history of bilateral, symmetrical moderately severe sensorineural hearing loss. Twenty more people were recruited through word of mouth or advertising. Participants who responded to our initial contact attempt underwent further screening prior to being enrolled in the study. The inclusion criteria were as follows: fluent English speakers (self-reported); bilateral HA user (at least 8 hr/week over the last 6 months; self-reported); high-frequency average gain of 5 dB or greater according to American National Standard Institute specification S3.22 (ANSI, 2003); a total score exceeding 21 out of 30 on the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005) to rule out dementia; no self-reported comorbid health conditions (e.g., conductive hearing loss, active otologic disorders, or complex medical history involving the head, neck, ears, or eyes) that would interfere with the study procedures; and bilateral, symmetrical mild to moderately severe sensorineural hearing loss. Bilateral, symmetrical mild to moderately severe sensorineural hearing loss was defined as thresholds no poorer than 70-dB HL from 250 through 4000 Hz, air-bone gaps no greater than 10 dB with a one-frequency exception, and all interaural thresholds within 15 dB at each frequency with one-frequency exception. Even though the results reported here only reflect unaided performance, each participant's amplification status was known for the purpose of future research. To determine the amount of gain the HAs were providing, each HA was placed in an Audioscan Verifit (Audioscan; Dorchester,

Ontario, Canada) test chamber, and the automatic gain control test battery (ANSI, 2003) was performed at user settings. The high-frequency average was determined to be the amount of gain reported for a 50-dB input averaged at 1000, 1600, and 2500 Hz. Approximately 140 people were screened between sites. If individuals were disqualified, it was most often due to low gain in their HAs, followed by not meeting the asymmetry criteria. Two people did not pass the MoCA screening, and four people did not use their HAs often enough to meet criteria.

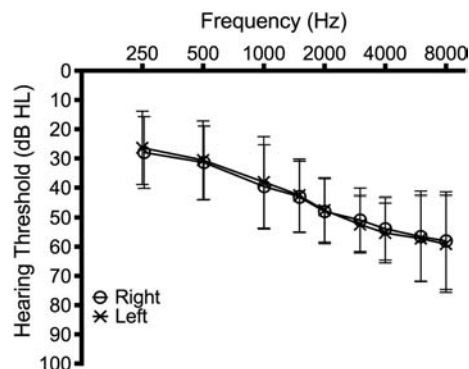
Seventy-six adults with HAs (53 women and 23 men) passed all screening criteria and participated in this experiment. Figure 1 illustrates the average audiogram across all participants for the right and left ears. The average gain measured was 14.7 dB ( $SD = 8.5$ ); however, all data reported in this article are for unaided conditions only. All participants provided informed consent to participate in the study approved by the Human Subjects Review Committee at the University of Washington and University of Iowa.

## Stimuli

### Speech Recognition

Unaided speech recognition in noise was assessed using the Multi-Modal Lexical Sentence Test for Adults (MLST-A; Kirk et al., 2012). The MLST-A is a clinically available measure of speech recognition (GN Otometrics, LIPread, available at <http://www.otometrics.com/>) that uses multiple target talkers, multiple presentation formats (audio-only, audio-video, visual-only), and optional background noise. The test is composed of 320 sentences, which are presented in 30 lists of 12 sentences each. Some sentences are repeated. Each sentence is seven to nine words in length and has three key words per sentence used for scoring purposes. During the development of the test (Kirk et al., 2012; Kirk, personal communication, November 5, 2015), 10 adults were selected to record the stimuli, and they represented diverse racial and ethnic backgrounds. Five women and five men served as the talkers. Individual lists were equated for intelligibility, regardless of the presentation

**Figure 1.** Mean hearing threshold levels across all participants for the right (o) and left (x) ears. Error bars represent the standard deviation at each frequency; HL = hearing level.



format. Two lists were presented and averaged in each condition. Sentences were presented through a loudspeaker (Tannoy Di5t; Coatbridge, North Lanarkshire, Scotland United Kingdom) at 65-dB SPL at 0° azimuth with the participant seated 1 m from the speaker. Only the AO and AV MLST-A presentation formats were used. The visual-only condition was not tested. During AV conditions, a 15-inch (13-cm) Acer (New Taipei City, Taiwan; University of Washington) or Dell (Round Rock, Texas; University of Iowa) monitor mounted directly below the 0° azimuth speaker presented the face of the person speaking each sentence.

Two types of background noise were used during speech-recognition testing. SSN was created from white noise and shaped to the long-term average spectrum of the speech stimuli in one-third octave bands. A 4-talker babble was made from the International Speech Test Signal (ISTS; Holube, Fredelake, Vlaming, & Kollmeier, 2010). Van Engen and Bradlow (2007) showed that babble in a language that is understandable to the listener (e.g., English) can lead to greater degrees of masking effects than with a babble in a foreign language, nonintelligible to the listener. Our primary interest was in the effects from multiple target talkers (i.e., the MLST-A speech); therefore, we wanted to minimize linguistic effects also occurring from the masker. To create the ISTS noise condition, the ISTS stimulus was temporally offset and wrapped to create four uncorrelated samples. In both noise conditions, noise was presented via four loudspeakers positioned at 0°, 90°, 180°, and 270° azimuth and 1 m from the participant at output levels to create an overall +8-dB signal-to-noise ratio (i.e., 57-dB SPL) at the location of the participant's head. The decision to present speech at 65-dB SPL and noise at a +8-dB signal-to-noise ratio was based on two reasons. First, previous data showed these speech and noise levels to be representative of those experienced in the real world (Pearsons, Bennett, & Fidell, 1977; Smeds, Wolters, & Rung, 2015). Second, in real-world listening environments, a person is unable to adjust the level of speech and noise signals. Thus, the levels used here permitted us to assess each person's capacity under similar testing conditions that resemble realistic listening situations.

### Working Memory

WM was quantified using the RS (Baddeley et al., 1985; Daneman & Carpenter, 1980; Rönnberg et al., 1989) and the WARRM (Smith & Pichora-Fuller, 2015; Smith et al., 2016). The RS is a verbal WM test administered in the visual domain using sentences presented on a computer monitor. In total, 54 sentences were presented in four set sizes (3, 4, 5, and 6), with three sequences in each set size. The participant sees one word of a sentence at a time and is instructed to read the words aloud as they come up. After each sentence, the participant makes a judgment about whether or not the sentence makes sense (i.e., is semantically valid). At the end of each sequence within a set size, the participant is asked to recall the first or last word in each sentence. The participant does not know if they will have to repeat the first or last words until the end of the set.

The percent of total words correctly recalled, regardless of order, was used in the analysis.

The WARRM is also a measure of verbal WM. The protocol of the WARRM is similar to the RS, but the stimuli are presented auditorily. The WARRM is composed of 100 target words presented in four set sizes (2, 3, 4, and 5) with five repetitions in each set size. The target words are preceded by the carrier phrase, “You will cite\_\_\_\_\_.” The carrier phrase and target words are presented through the loudspeaker at 0° azimuth at 80-dB SPL for participants with pure-tone averages (PTAs) less than 40-dB HL, and at 90-dB SPL for PTAs greater than 40-dB HL, as recommended by the test developers. The participant is instructed to both repeat the target word and judge whether the first letter of the word is from the first half (A–M) or second half (N–Z) of the alphabet immediately after sentence offset. At the end of each sequence, the participant is asked to recall all of the target words in the sequence in order, if possible. The words that each participant repeats are recorded. Regardless of whether or not the participant repeats the word correctly, if they recall the word later in the test procedure, the answer is recorded as correct. The percent of total target words correctly recalled was used for the analysis.

### Procedures

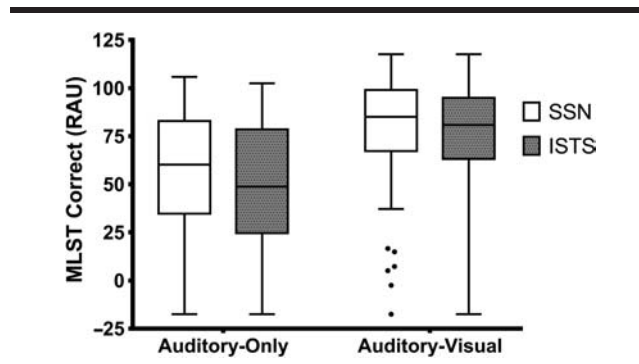
Testing took place in a sound-attenuated booth. All stimuli were presented using a combination of Windows Media Player (Microsoft 2011; Redmond, WA), and MATLAB and PsychToolBox (MathWorks 2012; Natick, MA), and a custom platform for presenting stimuli. In total, four MLST-A conditions were tested (AO-SSN, AO-ISTS, AV-SSN, and AV-ISTS) in a randomized order across participants. Noise was presented throughout each test block, and the participant was tasked to repeat any part of the sentence they could understand immediately following the sentence. For a response to be correct, the repeated word had to be identical to the target key word (e.g., plurals were not counted as correct). The experimenter recorded the responses before presenting the next sentence. The average of 24 sentences was used in the MLST-A analysis for each condition. The RS and WARRM were conducted after MLST-A testing.

## Results

### Speech Recognition

Mean (*SD*) percent correct on the MLST-A was 55.30 (31.67), 48.14 (32.03), 78.56 (27.33), and 75.27 (27.99) in the AO-SSN, AO-ISTS, AV-SSN, and AV-ISTS conditions, respectively. Percent correct absolute scores were transformed to rationalized arcsine units (RAU; Studebaker, 1985) to stabilize the variance. Mean (*SD*) performance on the MLST-A (RAU) was 55.52 (31.82), 48.33 (32.20), 78.62 (27.51), and 75.44 (28.14) in the AO-SSN, AO-ISTS, AV-SSN, and AV-ISTS conditions, respectively (Figure 2). As observed in Figure 2, the range in performance was from

**Figure 2.** Box and whisker plot showing performance on the Multi-Modal Lexical Sentence Test for adults (MLST-A) transformed to rationalized arcsine units (RAU) in each condition. Lines in the boxes represent the median, boxes represent the 25–75 percentiles, and the whiskers are the Tukey range, with filled circles representing potential outliers. SSN = steady-state noise; ISTS = International Speech Test Signal.

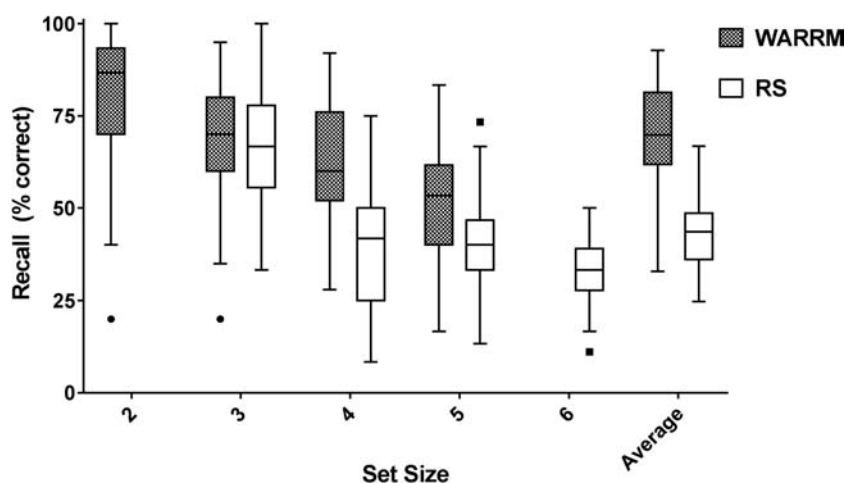


–17 to 117 RAU across noise types and format of presentation, and mean performance increased with the addition of visual cues and was slightly better in SSN than ISTS. A repeated measures analysis of variance (RM-ANOVA) was performed with noise type (SSN, ISTS) and presentation format (AO, AV) as within-participant factors. Although all MLST-A conditions were significantly and positively correlated with one another, the main and interaction effects in the ANOVA were all statistically significant, suggesting performance in each condition was different from one another. The main effect of noise,  $F(1, 75) = 37.769$ ;  $p < .0001$ ;  $\eta_p^2 = .335$ , indicated that performance was better in SSN than ISTS. The main effect of presentation format,  $F(1, 75) = 270.847$ ;  $p < .0001$ ;  $\eta_p^2 = .783$ , suggests that performance was better in AV conditions than in AO conditions. The interaction term between noise type and presentation format was also significant,  $F(1, 75) = 8.078$ ;  $p = .006$ ;  $\eta_p^2 = .097$ , and paired samples *t* tests showed a significant difference in performance between noise types in the AO condition (mean difference: 7.2 RAU;  $t = 6.919$ ;  $p < .0001$ ) and in the AV condition (mean difference: 3.2 RAU;  $t = 2.752$ ;  $p = .007$ ).

### Working Memory

Figure 3 shows percent correct recall on the RS and WARRM across set size. The mean (*SD*) percent correct recall across set sizes was 69.8% (12.94) on the WARRM and 42.51% (9.09) on the RS. Although recall performance values on the WARRM and RS were significantly correlated in this study ( $r = .58$ ;  $p < .0001$ ), previous work suggests that the strength of correlation between these two measures may depend on age, with older listeners showing a higher correlation between the two measures than younger listeners (Smith & Pichora-Fuller, 2015). Also, previous research has shown mixed results on whether WM measures presented auditorily (i.e., WARRM) or visually (i.e., RS) better predict speech recognition in noise (for a review, see

**Figure 3.** Box and whisker plot showing recall (percent correct) on both working memory tests, Word Auditory Recognition and Recall Measure (WARRM) and reading span measure (RS) across set sizes. Lines in the boxes represent the median, boxes represent the 25–75 percentiles, and the whiskers are the Tukey range, with filled circles representing potential outliers.



Smith & Pichora-Fuller, 2015). Therefore, WARRM and RS results remained separate in the subsequent analysis.

### Relationship Between Working Memory and Speech Performance

To describe relationships between the variables of interest, partial correlations among age, WM capacity, and MLST-A performance were calculated, controlling for better-ear PTA (thresholds at 500, 1000, and 2000 Hz). Two-tailed significant values are shown in Table 1. Age was not significantly correlated to MLST performance in any condition, nor to RS; however, age was significantly and negatively correlated to WARRM. One participant, who was 32 years old, was an outlier in terms of age, but his performance on the WM measures was within the middle 50% of the distribution (i.e., WARRM: 68%, RS: 48%). Further, the significance of the partial correlations did not change upon excluding his data. Therefore, his data were included in the remaining analyses. Figures 4–7 plot WM capacity (percent correct recall) on both measures, and the correlation with MLST-A, for each noise type and presentation format. The data in Table 1 show a significant partial correlation (controlling for PTA) between both WM measures and all four MLST-A conditions, in that as WM capacity increased, speech performance also increased.

A linear mixed model was used to examine the effect of WM on understanding of speech (MLST-A) and the dependency of this effect on different conditions (visual cue and noise type) for the data shown in Figures 4–7. Fixed effects considered in the model were WM (WARRM/RS), noise type (SSN/ISTS), visual cues (AO/AV), better-ear PTA, and age. A random intercept for participants was also included in the model. Noise type and visual cues were binary variables, whereas WM measures were continuous.

All interactions were also examined, yet none was significant. Better-ear PTA and visual cues were the only significant main effects on speech understanding,  $F(1, 75) = 32.22$ ,  $p < .0001$  for visual cue;  $F(1, 71) = 43.51$ ,  $p < .0001$  for PTA. Listeners with more severe hearing loss had poorer performance scores on the MLST-A. Also, MLST performance was better in the AV condition than in the AO condition. Neither of the WM measures had a significant effect on MLST-A performance. Table 2 shows the output of the model, including the only interaction that approached statistical significance.

### Discussion

The goal of the current study was to determine if the effects of WM on speech understanding in noise are maintained under more ecologically relevant test conditions than those used in previous investigations. Ecologically relevant stimuli were operationalized as multiple talkers in the target stimuli, steady-state and 4-talker babble noise, and in the absence and presence of visual cues. Our results from the linear mixed model suggest that WM had no effect on unaided speech understanding, even when controlling for hearing loss (i.e., PTA) and age. Minimal effects of WM on speech understanding have been found previously (e.g., Ng et al., 2014; Smith & Pichora-Fuller, 2015), whereas others have demonstrated a robust effect (e.g., Gatehouse et al., 2003; Lunner, 2003; Souza & Arehart, 2015). These conflicting results may be due to differences between studies in the characteristics of the speech task, the participants, or the statistical analyses.

### Effects of Minimal Indexical Cues

The MLST-A minimizes the learning of indexical cues by using multiple target talkers. In a recent survey

**Table 1.** The partial correlations ( $r$ ) among predictors and dependent variables, controlling for the better-ear pure-tone average.

Variable	1	2	3	4	5	6	7
1. Age							
$r$	–						
$p$	–						
2. RS							
$r$	–.18						
$p$	.128						
3. WARRM							
$r$	–.24*	.58**					
$p$	.040	< .001					
4. MLST-A AO SSN							
$r$	.05	.29**	.33**				
$p$	.692	.012	.004				
5. MLST-A AO ISTS							
$r$	.02	.26*	.26*	.94**			
$p$	.876	.022	.024	< .001			
6. MLST-A AV SSN							
$r$	.01	.34**	.27*	.84**	.79**		
$p$	.946	.003	.019	< .001	< .001		
7. MLST-A AV ISTS							
$r$	–.02	.25*	.26*	.85**	.83**	.91**	
$p$	.901	.030	.025	< .001	< .001	< .001	

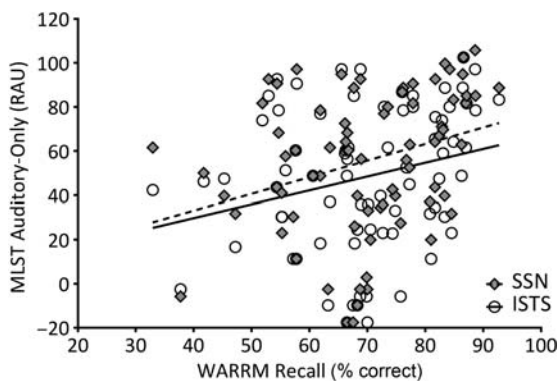
Note. RS = reading span measure; WARRM = Word Auditory Recognition and Recall Measure; MLST-A = Multi-Modal Lexical Sentence Test for adults; AO = auditory-only; SSN = steady-state noise; ISTS = International Speech Test Signal; AV = auditory–visual.

\*Correlation is significant at the .05 level (2-tailed). \*\*Correlation is significant at the .01 level (2-tailed).

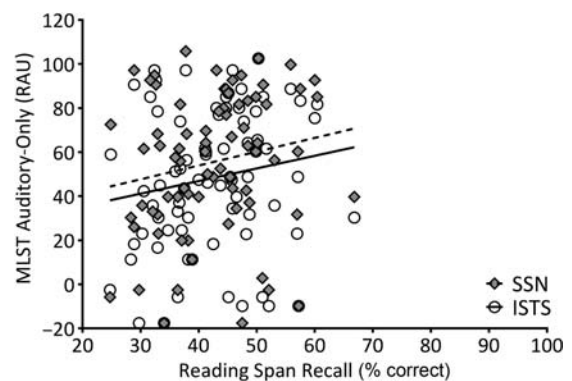
of data reporting relationships between WM and speech recognition in noise, Füllgrabe and Rosen (2016) reported inconsistent trends in correlations across the difficulty of the listening condition, with some studies showing more reliance on WM for harder tasks and some studies even showing the opposite effect (greater reliance on WM for easier tasks). Previous work using single-talker, AO stimuli

(e.g., QuickSIN) has shown RS to significantly explain 2% of variance over that explained by audibility (Souza & Arehart, 2015). Smith and Pichora-Fuller (2015) used several tests for speech understanding in noise to test whether linguistic complexity (words, sentences, discourse) resulted in differential effects of WM on performance. Using a similar group of older adults with hearing loss as those

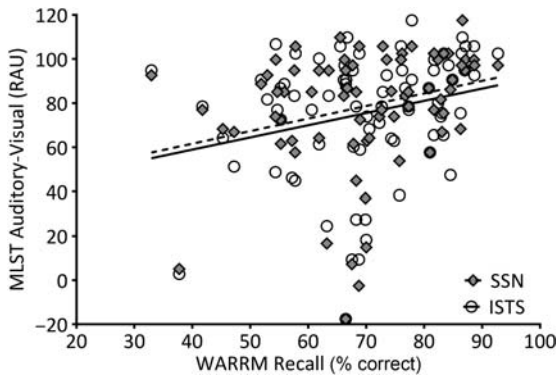
**Figure 4.** Working memory capacity on the Word Auditory Recognition and Recall Measure (WARRM; percent correct recall) plotted relative to Multi-Modal Lexical Sentence Test for adults (MLST-A) performance transformed to rationalized arcsine units (RAU) for auditory-only conditions with steady-state noise (SSN; filled diamonds) and International Speech Test Signal (ISTS; open circles) noise. A linear regression line was fit to SSN (dotted line) and ISTS (solid line) data sets. See Tables 1 and 2 for the statistical analyses.



**Figure 5.** Working memory capacity on the reading span measure (percent correct recall) plotted relative to Multi-Modal Lexical Sentence Test for adults (MLST-A) performance transformed to rationalized arcsine units (RAU) for auditory-only conditions with steady-state noise (SSN; filled diamonds) and International Speech Test Signal (ISTS; open circles) noise. A linear regression line was fit to SSN (dotted line) and ISTS (solid line) data sets. See Tables 1 and 2 for the statistical analyses.



**Figure 6.** Working memory capacity on the Word Auditory Recognition and Recall Measure (WARRM; percent correct recall) plotted relative to Multi-Modal Lexical Sentence Test for adults (MLST-A) performance transformed to rationalized arcsine units (RAU) for auditory–visual conditions with steady-state noise (SSN; filled diamonds) and International Speech Test Signal (ISTS; open circles) noise. A linear regression line was fit to SSN (dotted line) and ISTS (solid line) data sets. See Tables 1 and 2 for the statistical analyses.

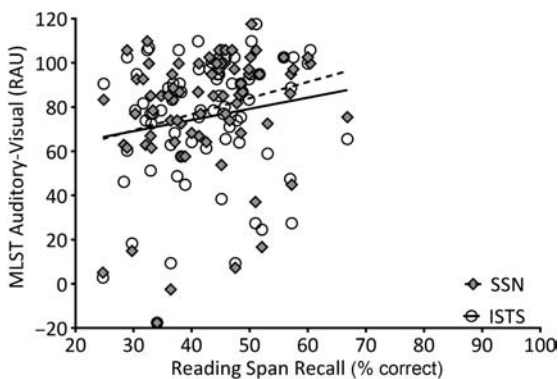


in the current study, they found no significant relationships between either measure of WM (RS and WARRM) and their speech-recognition tasks. Altogether, WM influences on verbal communication may be greater, or even stronger, with stimuli differences in talker characteristics (e.g., dialect, speech rate, gender) and not with linguistic differences (e.g., Smith & Pichora-Fuller, 2015).

### Noise Type and Visual Cues: Effects of WM Did Not Differ

Previous work has shown a slightly larger effect of WM on speech understanding in modulated noise (e.g., 10% in 2-talker babble; Gatehouse et al., 2003) compared

**Figure 7.** Working memory capacity on the reading span measure (percent correct recall) plotted relative to Multi-Modal Lexical Sentence Test for adults (MLST-A) performance transformed to rationalized arcsine units (RAU) for auditory–visual conditions with steady-state noise (SSN; filled diamonds) and International Speech Test Signal (ISTS; open circles) noise. A linear regression line was fit to SSN (dotted line) and ISTS (solid line) data sets. See Tables 1 and 2 for the statistical analyses.



with steady-state noise (e.g., 8% in steady-state noise; Gatehouse et al., 2003). In the current study, the difference in variability explained by WM between noise types was not significant in the linear mixed model, as indicated by the nonsignificant interactions. It is possible that the task difficulty was already higher with multiple target talkers compared with the single-talker speech used in previous research (Mullennix et al., 1989; Nygaard & Pisoni, 1998; Nygaard et al., 1994), which perhaps overrode the stronger WM effects that we expected with modulated noise. It is also possible that the use of the ISTS to create the babble led to less informational masking than expected, which limited the influence of WM under this noise condition.

We had hypothesized that adding visual cues would improve speech understanding, yet with a greater reliance on WM due to reconstructing of the auditory and visual peripheral input (presumably by the episodic buffer; Baddeley, 2000). However, reliance on WM was not greater under AV conditions compared with AO conditions in the current study, contrary to our prediction, suggesting that bimodal processing did not require additional WM resources. One explanation is that although a +8-dB signal-to-noise ratio is representative of real-world experiences, it may not have been challenging enough, especially when visual cues were present. As a result, the participants may have not needed to use top-down processing in this condition, resulting in the minimal WM effects and nonsignificant interactions. Although WM capacity did not differ with bimodal input, it is possible that reaction time may have been improved (e.g., Frtusova and Phillips, 2016), which was not measured in the current study.

### Effects of Hearing Loss and Age

Hearing loss (defined as better-ear PTA in the current study) was the primary predictor in our speech-recognition outcomes. These results were not surprising, given that hearing loss has been the most consistent predictor in speech recognition for years (e.g., Humes, 2002). However, age did not play a role in our speech-recognition outcomes once PTA was considered; this finding is in contrast to previous work (e.g., Dubno, Horwitz, & Ahlstrom, 2002; Gordon-Salant & Fitzgibbons, 1997), yet in agreement with others (e.g., Souza & Arehart, 2015). It should be noted that the age range of our participants (52–79 years, with one listener at 32 years) may not have been wide enough to show an age effect. Further, age correlated weakly with the WARRM, but not with the RS results, which is consistent with findings by Arehart, Souza, Baca, and Kates (2013). These age effects could be due to the distribution of WARRM recall scores being wider than those for RS or also from not recruiting enough listeners in the younger decades. Studies that have found more robust age effects on WM tasks have included individuals in their 20s and 30s (e.g., Bopp & Verhaeghen, 2005). Another possibility is that other resources are required to understand speech with multiple target talkers that were not measured in the current study, and these may have dominated any effects of



**Table 2.** Effect of factors in the linear mixed model (with compound symmetric covariance matrix).

Effect	Level of noise type	Level of visual cues	Estimate	Standard error	t	Wald p	Type 3 test	
							F	p
Intercept			76.79	33.44	2.23	.03		
Age			0.31	0.36	0.87	.39	0.76	.39
PTA			-1.55	0.26	-6.60	< .01	43.51	< .01
WARRM Recall			0.32	0.25	1.26	.21	2.16	.15
RS Recall			0.50	0.35	1.45	.15	2.11	.15
Noise Type	ISTS		-0.18	5.98	-0.03	.98	0.00	.98
	SSN		Reference					
Visual Cues		AO	-33.92	5.98	-5.68	< .01	32.22	< .01
		AV	Reference					
WARRM × Noise Type	ISTS	AO	0.06	0.12	0.46	.65	2.25	.09
× Visual Cues	ISTS	AV	-0.04	0.09	-0.49	.62		
	SSN	AO	0.16	0.09	1.82	.07		
	SSN	AV	Reference					

Note. PTA = pure-tone average; WARRM = Word Auditory Recognition and Recall Measure; RS = reading span measure; ISTS = International Speech Test Signal; SSN = steady-state noise; AO = auditory-only; AV = auditory-visual.

age that could have occurred. For example, under the Framework for Effortful Listening (Pichora-Fuller et al., 2016), other sources of input-related demands (e.g., suprathreshold distortion), attention-related processes, or fatigue may have influenced results.

## Conclusion

To summarize, under ecologically relevant listening conditions, WM influences on verbal communication appear to be negligible for unaided listening; when controlling for the effects of hearing loss and age, effect sizes were not sufficient to warrant clinical use of WM tests. However, future studies should evaluate the relevance of WM on aided performance under similar test conditions, as well as self-report measures. Previous work suggests that after enhancing audibility with amplification, WM plays a more important role in explaining speech-recognition variability (e.g., Akeroyd, 2008; Lunner, 2003; Ng et al., 2014). If effect sizes for WM hold for aided outcomes in more ecologically relevant situations, substantial evidence may exist to change clinical practice regarding counseling, selecting, and fitting of HAs on the basis of WM performance.

## Acknowledgments

We would like to thank the community practitioners for advertising our study, the participants for their time, and our funding sources for making this work possible: National Institute on Deafness and Other Communication Disorders Grants R01 DC012769-04 (awarded to Kelly Tremblay and Ruth A. Bentler) P30 DC004661 (awarded to the Core Center, NIDCD Research Core Center P30 [Rubel, PI]), and the Institute for Clinical and Translational Science at the University of Iowa is supported by the National Institutes of Health (NIH) Clinical and Translational Science Award (CTSA) program, grant U54TR001356. We also thank Ashley Moore, Kelley Trapp, Elizabeth Stangl, and Kelsey Dumanch for data collection and entry and Xuyang Zhang for statistical assistance.

## References

- Akeroyd, M. A. (2008). 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. *International Journal of Audiology, 47*, S53-S71.
- American National Standards Institute. (2003). *American National Standard Specification of Hearing Aid Characteristics (ANSI S3.22-2003)*. New York: American National Standards Institute.
- Arehart, K. H., Souza, P., Baca, R., & Kates, J. M. (2013). Working memory, age, and hearing loss: Susceptibility to hearing aid distortion. *Ear and Hearing, 34*(3), 251-260.
- Baddeley, A. D. (2000). The phonological loop and the irrelevant speech effect: Some comments on Neath (2000). *Psychonomic Bulletin & Review, 7*(3), 544-549.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 47-89). London, England: Academic Press.
- Baddeley, A. D., Logie, R., Nimmo-Smith, I., & Breerton, N. (1985). Components of fluid reading. *Journal of Memory and Language, 24*, 119-131.
- Bernstein, J. G., & Grant, K. W. (2009). Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 125*, 3358-3372.
- Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *Journal of Gerontology, 60B*(5), 223-233.
- Broadbent, D. E. (1952). Listening to one of two synchronous messages. *Journal of Experimental Psychology, 44*, 51-55.
- Chia, E. M., Wang, J. J., Roctchina, E., Cumming, R. R., Newall, P., & Mitchell, P. (2007). Hearing impairment and health-related quality of life: The Blue Mountains Hearing Study. *Ear and Hearing, 28*, 187-195. <https://doi.org/10.1097/AUD.0b013e31803126b6>
- Cox, R. M., & Xu, J. (2010). Short and long compression release times: Speech understanding, real-world preferences, and association with cognitive ability. *Journal of the American Academy of Audiology, 21*, 121-138.
- Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450-466.

- Daneman, M., & Carpenter, P. A.** (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(4), 561–584.
- Dubno, J. R., Horwitz, A. R., & Ahlstrom, J. B.** (2002). Benefit of modulated maskers for speech recognition by younger and older listeners with normal hearing. *Journal of the Acoustical Society of America*, 111, 2897–2907.
- Foo, C., Rudner, M., Rönnerberg, J., & Lunner, T.** (2007). Recognition of speech in noise with new hearing instrument compression release settings requires explicit cognitive storage and processing capacity. *Journal of the American Academy of Audiology*, 18, 618–631.
- Fraser, S., Gagné, J. P., Alepins, M., & Dubois, P.** (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: The effects of providing visual speech cues. *Journal of Speech, Language, and Hearing Research*, 53, 18–33.
- Frtusova, J. B., & Phillips, N. A.** (2016). The auditory–visual speech benefit on working memory in older adults with hearing impairment. *Frontiers in Psychology*, 7, 491. <https://doi.org/10.3389/fpsyg.2016.00490>
- Füllgrabe, C., & Rosen, S.** (2016). On the (un)importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Frontiers in Psychology*, 7, 1268. <https://doi.org/10.3389/fpsyg.2016.01268>
- Gatehouse, S., Naylor, G., & Elberling, C.** (2003). Benefits from hearing aids in relation to the interaction between the user and the environment. *International Journal of Audiology*, 42, S77–S85.
- Gatehouse, S., Naylor, G., & Elberling, C.** (2006). Linear and non-linear hearing aid fittings—2. Patterns of candidature. *International Journal of Audiology*, 45(3), 153–171.
- George, E. L., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M., & Houtgast, T.** (2007). Auditory and nonauditory factors affecting speech reception in noise by older listeners. *Journal of the Acoustical Society of America*, 121, 2362–2375.
- Gordon-Salant, S., & Fitzgibbons, P. J.** (1997). Selected cognitive factors and speech recognition performance among young and elderly listeners. *Journal of Speech, Language, and Hearing Research*, 40, 423–431.
- Helfer, K. S., & Freyman, R. L.** (2005). The role of visual speech cues in reducing energetic and informational masking. *Journal of the Acoustical Society of America*, 117, 842–849.
- Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B.** (2010). Development and analysis of an International Speech Test Signal (ISTS). *International Journal of Audiology*, 49(12), 891–903.
- Humes, L. E.** (2002). Factors underlying the speech-recognition performance of elderly hearing-aid wearers. *Journal of the Acoustical Society of America*, 112, 1112–1132.
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., & Banerjee, S.** (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 116, 2395–2405.
- Kirk, K. I., Pisoni, D. B., & Miyamoto, R. C.** (1997). Effects of stimulus variability on speech perception in listeners with hearing impairment. *Journal of Speech, Language, and Hearing Research*, 40, 1395–1405.
- Kirk, K. I., Prusick, L., French, B., Gotch, C., Eisenberg, L. S., & Young, N.** (2012). Assessing spoken word recognition in children who are deaf or hard of hearing: A translational approach. *Journal of the American Academy of Audiology*, 23(6), 464–475.
- Kochkin, S.** (2000). MarkeTrak V: “Why my hearing aids are in the drawer”: The consumers’ perspective. *The Hearing Journal*, 53(2), 34–41.
- Lunner, T.** (2003). Cognitive function in relation to hearing aid use. *International Journal of Audiology*, 42, S49–S58.
- Lunner, T., & Sundewall-Thorén, E.** (2007). Interactions between cognition, compression, and listening condition: Effects on speech-in-noise performance in a two-channel hearing aid. *Journal of the American Academy of Audiology*, 18(7), 604–617.
- Marrone, N., Alt, M., DeDe, G., Olson, S., & Shehorn, J.** (2015). Effects of steady-state noise on verbal working memory in young adults. *Journal of Speech, Language, and Hearing Research*, 58, 1793–1804.
- Martin, C. S., Mullennix, J. W., Pisoni, D. B., & Summers, W. V.** (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(4), 676–684.
- Mishra, S., Lunner, T., Stenfelt, S., Rönnerberg, J., & Rudner, M.** (2013). Visual information can hinder working memory processing of speech. *Journal of Speech, Language, and Hearing Research*, 56, 1120–1132.
- Mullennix, J. W., Pisoni, D. B., & Martin, C. S.** (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, 85, 365–378.
- Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., . . . Chertkow, H.** (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699.
- Neher, T., Behrens, T., Carlile, S., Jin, C., Kragelund, L., Petersen, A. S., & Schaik, A. V.** (2009). Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, and cognition. *International Journal of Audiology*, 48(11), 758–774.
- Ng, E. H. N., Classon, E., Larsby, B., Arlinger, S., Lunner, T., Rudner, M., & Rönnerberg, J.** (2014). Dynamic relation between working memory capacity and speech recognition in noise during the first 6 months of hearing aid use. *Trends in Hearing*, 18, 1–10.
- Ng, E. H. N., Rudner, M., Lunner, T., Pedersen, M. S., & Rönnerberg, J.** (2013). Effects of noise and working memory capacity on memory processing of speech for hearing-aid users. *International Journal of Audiology*, 52(7), 433–441.
- Nygaard, L. C., & Pisoni, D. B.** (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, 60, 355–376.
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B.** (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5(1), 42–46.
- Pearsons, K. S., Bennett, R. L., & Fidell, S.** (1977). *Speech levels in various noise environments*. Report No. 600/1-77-025. Washington, DC: U.S. Environmental Protection Agency.
- Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., . . . Wingfield, A. D.** (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, 37(S1), S5–S27.
- Rönnerberg, J.** (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Journal of Audiology*, 42(Suppl. 1), S68–S76. <https://doi.org/10.3109/14992020309074626>
- Rönnerberg, J., Arlinger, S., Lyxell, B., & Kinnefors, C.** (1989). Visual evoked potentials: Relations to adult speechreading and cognitive function. *Journal of Speech and Hearing Research*, 32, 725–735.
- Rönnerberg, J., Lunner, T., Zekveld, A., Sorqvist, P., Danielsson, H., Lyxell, B., . . . Rudner, M.** (2013). The Ease of Language

- Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, 7, 31.
- Rönnerberg, J., Rudner, M., Foo, C., & Lunner, T.** (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology*, 47(Suppl. 2), S99–S105.
- Rönnerberg, J., Rudner, M., Lunner, T., & Zekveld, A. A.** (2010). When cognition kicks in: Working memory and speech understanding in noise. *Noise & Health*, 12(49), 263–269.
- Rudner, M., Foo, C., Rönnerberg, J., & Lunner, T.** (2009). Cognition and aided speech recognition in noise: Specific role for cognitive factors following nine-week experience with adjusted compression settings in hearing aids. *Scandinavian Journal of Psychology*, 50(5), 405–418.
- Rudner, M., Rönnerberg, J., & Lunner, T.** (2011). Working memory supports listening in noise for persons with hearing impairment. *Journal of the American Academy of Audiology*, 22(3), 156–167.
- Schneider, B. A., Daneman, M., & Pichora-Fuller, M. K.** (2002). Listening in aging adults: From discourse comprehension to psychoacoustics. *Canadian Journal of Experimental Psychology*, 56(3), 139–52.
- Silberer, A. B., Bentler, R., & Wu, Y.** (2015). The importance of high-frequency audibility with and without visual cues on speech recognition for listeners with normal hearing. *International Journal of Audiology*, 54(11), 865–872.
- Smeds, K., Wolters, F., & Rung, M.** (2015). Estimation of signal-to-noise ratios in realistic sound scenarios. *Journal of the American Academy of Audiology*, 26(2), 183–196.
- Smith, S. L. & Pichora-Fuller, M. K.** (2015). Associations between speech understanding and auditory and visual tests of verbal working memory: Effects of linguistic complexity, task, age and hearing loss. *Frontiers in Psychology*, 6, 1394.
- Smith, S. L., Pichora-Fuller, M. K., & Alexander, G.** (2016). Development of the Word Auditory Recognition and Recall Measure: A working memory test for use in rehabilitative audiology. *Ear and Hearing*, 37(6), e360–e376. <https://doi.org/10.3389/fpsyg.2015.01394>
- Souza, P., & Arehart, K. H.** (2015). Robust relationship between reading span and speech recognition in noise. *International Journal of Audiology*, 54(10), 705–713.
- Studebaker, G. A.** (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 455–462.
- Tamati, T. N., Gilbert, J. L., & Pisoni, D. B.** (2013). Some factors underlying individual differences in speech recognition on PRESTO: A first report. *Journal of the American Academy of Audiology*, 24(7), 616–634.
- Van Engen, K. J., & Bradlow, A. R.** (2007). Sentence recognition in native- and foreign-language multi-talker background noise. *Journal of the Acoustical Society of America*, 121, 519–526.
- van Rooij, J. C., & Plomp, R.** (1990). Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses. *Journal of the Acoustical Society of America*, 88, 2311–2324.
- Wingfield, A., & Tun, P. A.** (2001). Spoken language comprehension in older adults: Interactions between sensory and cognitive change in normal aging. *Seminars in Hearing*, 22(3), 287–302.
- Working Group on Speech Understanding and Aging, Committee on Hearing, Bioacoustics, and Biomechanics, Commission on Behavioral and Social Sciences and Education, National Research Council.** (1988). Speech understanding and aging. *Journal of the Acoustical Society of America*, 83, 859–895.
- Wu, Y. H., & Bentler, R. A.** (2010). Impact of visual cues on directional benefit and preference: Part I—laboratory tests. *Ear and Hearing*, 31(1), 22–34.
- Wu, Y. H., Stangl, E., Zhang, X., & Bentler, R. A.** (2015). Construct validity of the Ecological Momentary Assessment in audiology research. *Journal of the American Academy of Audiology*, 26(10), 872–884.