Research Article

Speech Recognition in Adults With Cochlear Implants: The Effects of Working Memory, Phonological Sensitivity, and Aging

Aaron C. Moberly,[a] Michael S. Harris,[a] Lauren Boyce,[a] and Susan Nittrouer[b]

Purpose: Models of speech recognition suggest that "top-down" linguistic and cognitive functions, such as use of phonotactic constraints and working memory, facilitate recognition under conditions of degradation, such as in noise. The question addressed in this study was what happens to these functions when a listener who has experienced years of hearing loss obtains a cochlear implant.
Method: Thirty adults with cochlear implants and 30 age-matched controls with age-normal hearing underwent testing of verbal working memory using digit span and serial recall of words. Phonological capacities were assessed using a lexical decision task and nonword repetition. Recognition of words in sentences in speech-shaped noise was measured.

Results: Implant users had only slightly poorer working memory accuracy than did controls and only on serial recall of words; however, phonological sensitivity was highly impaired. Working memory did not facilitate speech recognition in noise for either group. Phonological sensitivity predicted sentence recognition for implant users but not for listeners with normal hearing.
Conclusion: Clinical speech recognition outcomes for adult implant users relate to the ability of these users to process phonological information. Results suggest that phonological capacities may serve as potential clinical targets through rehabilitative training. Such novel interventions may be particularly helpful for older adult implant users.

Although most adults with postlingual deafness who receive cochlear implants (CIs) derive some benefit regarding speech recognition, enormous variability and individual differences in outcomes exist. Approximately half of this variability can be explained by factors that relate to differences in individuals' sensitivity to spectral and temporal cues after implantation, such as the condition of the peripheral auditory system (e.g., duration of deafness or amount of preoperative hearing) and the positioning of the electrode array (e.g., proximity to the modiolus or angle of insertion; Holden et al., 2013; Holden, Reeder, Firszt, & Finley, 2011; Kenway et al., 2015). However, there is increasing evidence from work with adults with various degrees of hearing loss that cognitive functions may explain additional variability in speech recognition

outcomes (Akeroyd, 2008; Arehart, Souza, Baca, & Kates, 2013; Lunner & Sundewall-Thorén, 2007; Pichora-Fuller & Souza, 2003; Rönnberg et al., 2013; Rudner, Foo, Sundewall-Thorén, Lunner, & Rönnberg, 2008).

For robust speech recognition, the listener must use previously developed language knowledge and cognitive abilities to make sense of the incoming speech signal, relating it to phonological and lexical representations in long-term memory (Pisoni & Cleary, 2003). These "top-down" processes are especially important when the "bottom-up" sensory input is degraded (e.g., in noise, when using a hearing aid, or when listening to the spectrally degraded signals transmitted by a CI); degraded input leads to greater ambiguity in how the information within that input should be organized perceptually. Several models of continuous speech recognition suggest that listeners use their knowledge of language-specific semantic, syntactic, and phonological structures to constrain variability in potential linguistic choices, thus reducing the amount of sensory information required to make accurate lexical selections (Ahissar, 2007; Boothroyd, 2010). For CI users, who have available to them only speech signals that are spectrally degraded relative to normal speech, these cognitive functions may play even greater roles. With the growing elderly population in

[a]Department of Otolaryngology–Head and Neck Surgery, Wexner Medical Center, The Ohio State University, Columbus
[b]Department of Speech, Language, and Hearing Sciences, University of Florida, Gainesville
Correspondence to Aaron C. Moberly: aaron.moberly@osumc.edu
Editor: Nancy Tye-Murray
Associate Editor: Richard Dowell

the United States, progressively larger numbers of older adults are receiving CIs (Dillon et al., 2013). This fact presents a clinical challenge to optimizing outcomes for the expanding population of CI users because cognitive functions decline with advancing age. That fact could further compromise speech recognition in this population of patients with hearing loss (Lin et al., 2011, 2013).

Defining Working Memory

Age-related declines in general cognition have been well recognized (Hällgren, Larsby, Lyxell, & Arlinger, 2001; Humes, 2005; Van Rooij & Plomp, 1990; Wilson, Leurgans, Boyle, Schneider, & Bennett, 2010). One cognitive function that is particularly relevant to success in speech recognition for individuals with hearing loss, and has been used to document evidence of age-related declines, is verbal working memory (WM; e.g., Bopp & Verhaeghen, 2005; Kumar & Priyadarshi, 2014; Nittrouer, Lowenstein, Wucinich, & Moberly, 2016; Rönnberg et al., 2013). WM is commonly defined as a limited capacity, temporary storage mechanism for holding information (Baddeley, 1992; Daneman & Carpenter, 1980; Daneman & Hannon, 2007). This mechanism serves a vital role in temporarily maintaining information for further processing, such as during the process of recognizing and comprehending spoken language. Models differ regarding whether WM is domain specific (i.e., independent processes are dedicated to phonological, visual, or spatial tasks) or domain general (i.e., the same processes are implemented, regardless of input modality), but most models share the property of dual mechanisms: a short-term storage component and a processing component (Andersson & Lyxell, 2007; Daneman & Hannon, 2007; Repovš & Baddeley, 2006).

WM can be measured using tasks that examine the storage and processing components somewhat separately. For example, the task may consist of having the participant recall a series of familiar items (e.g., digits or words) in correct serial order (Nittrouer, Caldwell-Tarr, & Lowenstein, 2013); accuracy scores may serve as indicators of success in storage, and response times may reflect processing speed. The demands of the task may place greater or lesser demands on storage or processing components of WM (Alloway, Gathercole, & Pickering, 2006; Pickering, 2001). For example, forward digit span requires a listener to repeat a sequence of digits in the correct order, whereas backward digit span requires a listener to repeat a digit sequence in reverse order (Tao et al., 2014). The former primarily involves storage, and the latter involves a larger processing component.

WM, Hearing Loss, and Speech Recognition

For adults with hearing loss, verbal WM assessment has been generally effective for predicting speech recognition in noise (Akeroyd, 2008). Foo, Rudner, Rönnberg, and Lunner (2007) examined verbal WM (using a reading span task) and aided speech recognition performance (using Hagerman vs. Swedish-HINT sentences) in 32 older adults

with mild to moderate hearing loss. Results revealed significant correlations of reading span and speech recognition in both static and modulated masking noise. Gatehouse, Naylor, and Elberling (2003, 2006) examined word recognition in static or modulated noise, both unaided and in several aided conditions, in 50 elderly listeners with mild to moderate hearing loss. In that study, WM was assessed using letter- and digit-monitoring tasks, and these data were combined and used to classify listeners into low-, mid-, or high-performing groups. For unaided listening, a significant effect on word recognition was found for WM performance, with a 9% difference between low- and high-performing WM groups; differences were larger at more challenging signal-to-noise ratios (SNRs) and in modulated noise, as compared with testing in static noise. For certain aided conditions, such as use of fast compression, WM performance level correlated with speech recognition benefit. Lunner (2003) investigated the relationships between speech reception thresholds for sentences in modulated noise and verbal WM on reading span in 72 elderly patients with mild to moderate hearing loss. Significant correlations ($r = 0.4$–0.5) were found. Arehart et al. (2013) identified WM as a significant factor in listeners' recognition of sentences in babble processed with frequency compression, accounting for 29.3% of the variance in recognition scores. These findings are consistent with numerous other reports of correlations of verbal WM and speech perception for patients, tested in unaided and/or aided conditions, when the speech has been degraded by peripheral hearing loss (Cervera, Soler, Dasi, & Ruiz, 2009), noise (Lunner & Sundewall-Thorén, 2007; Rudner et al., 2008), spectral degradation (Schvartz, Chatterjee, & Gordon-Salant, 2008), or dynamic range compression (Piquado, Benichov, Brownell, & Wingfield, 2012; Rudner, Rönnberg, & Lunner, 2011). Thus, there is generally support for the idea that verbal WM abilities are related to speech recognition skills in patients with hearing loss listening under noisy conditions and in participants with normal hearing (NH) listening to degraded speech. However, it is generally difficult to gauge from these studies whether the effects are due to problems in storage or in processing. That question was addressed in this study.

The studies above involved listeners with sufficient residual hearing to warrant hearing aids but not CIs. Far less is known about WM and speech recognition in listeners with CIs. In one of the few studies of WM in adults with severe to profound hearing loss who received early generation CIs, Lyxell et al. (1998) found a small but significant correlation between reading span scores before implantation and speech recognition abilities after 12 months of implant experience. Tao et al. (2014) found significant correlations of speech recognition (Mandarin disyllable recognition) and digit span scores (forward and backward) in adult CI users. And although verbal WM has received little attention in adult CI users, a number of researchers have examined this cognitive ability in pediatric CI users. Several studies have suggested that verbal WM plays an important role in speech and language outcomes for children with CIs (Dawson, Busby, McKay, & Clark, 2002; Pisoni, 2014). Pediatric

CI users consistently have poorer verbal WM than do age-matched NH peers, as determined through auditory tasks (Cleary, Pisoni, & Geers, 2001; Nittrouer et al., 2013). However, Dawson et al. failed to find significant differences between children with NH and children with CIs in tasks of visuospatial WM, where the test items were unlikely to be recoded into verbal form. That finding suggests that the WM problems of this group may have more to do with storage than with processing. In a study of children with CIs and NH peers, Nittrouer et al. found relative deficits in CI users on an auditory task of serial recall of monosyllabic words. CI users were less accurate, but processing speeds were equivalent. These findings corroborated the suggestion that WM deficits for children with CIs were a result of problems with the storage component of WM, while processing abilities were intact. In another highly relevant study, Geers, Pisoni, and Brenner (2013) used auditory digit span and visual reading span tasks to examine WM in adolescents with prelingual deafness who had used CIs since childhood. CI users had poorer auditory digit span scores but equivalent reading span scores.

Phonological Sensitivity in CI Users

The WM deficits identified in CI users suggest that the locus of impairment resides in the storage component. According to one prominent model of WM, sensory input must be analyzed in one early operating component of the system and in the output of that component used to store items (Baddeley, 1992; Repovš & Baddeley, 2006). For verbal materials, that early operating component is the phonological loop, which recovers phonological structure and stores items with that structure. Previous work has demonstrated the importance of phonological structure to verbal WM by revealing an advantage in the serial recall of nonrhyming over rhyming words, with a greater advantage demonstrated by participants with better phonological sensitivity. For example, adults show a greater difference in recall of nonrhyming versus rhyming words than do children (Nittrouer & Miller, 1999), and children with typical reading skills show a greater difference than do children with phonologically based reading disorders (Mann & Liberman, 1984; Spring & Perry, 1983). Thus, children with reading disabilities are less able to store verbal material in short-term memory buffers than are typically reading children and adults.

On the basis of the important contribution of the phonological loop to the early operating component of WM, phonological sensitivity must be considered when studying verbal WM. The term *phonological sensitivity* refers to the listeners' abilities to recognize detailed phonological (including phonemic) structure in the speech stream. Previous studies have provided support for the concept that prolonged hearing loss can lead to degeneration of long-term phonological representations. Signal degradation also could impede the ability of listeners to recover phonological representations, even when those representations remain intact internally. Lyxell et al. (1998) found that adults with severe to profound hearing loss performed poorly on visually presented rhyme-judgment tasks, leading to the conclusion that phonological representations had deteriorated. Similar work by others has supported that conclusion (Andersson, 2001; Classon, Rudner, Johansson, & Rönnberg, 2013; Classon, Rudner, & Rönnberg, 2013; Lyxell et al., 1998; Lyxell, Andersson, Borg, & Ohlsson, 2003), but corroboration with varying paradigms is lacking. We recently found that a group of 30 adult CI users spanning a wide age range had less phonemic sensitivity than did age-matched NH peers on audiovisual tasks requiring selection of words with the same initial or final phoneme as a target word and that better performance on these tasks predicted better word recognition in quiet (Moberly, Lowenstein, & Nittrouer, 2016). We also found that older age was correlated with less phonemic sensitivity for the adults with CIs but not for their NH peers. Consequently, problems with storage might be predicted for older CI users.

Regarding the effects of aging on verbal WM, Nittrouer et al. (2016) sought to disentangle the effects of storage and processing. In that study, digit span and serial recall of words were assessed. Older listeners had significantly poorer accuracy and slower responses during the serial recall WM tasks. Phonological capacities were equivalent across age groups; nonetheless, within the older group, phonological sensitivity predicted serial recall accuracy. Findings from that study suggested that aging has a deleterious effect on verbal WM as a result of both cognitive slowing (processing) and diminished sensory inputs (storage) even when auditory thresholds are relatively good. However, the integrity of phonological capacities can help ameliorate these effects by facilitating better storage.

Goals of the Current Study

The overarching goal of this study was to examine what factors other than those associated with the damaged peripheral auditory system might explain variability in outcomes for adults with postlingual deafness who get CIs. The first objective was to examine whether verbal WM abilities were poorer for adult CI users with postlingual deafness than for age-matched NH peers. Two measures of verbal WM were included: digit span and serial recall of lists of established words. Digit span is used widely in clinical and research assessments of WM, and serial recall was included because its design could help identify whether any group differences in WM were due to differences in participants' access to phonological structure (storage) or to differences in WM response times (processing). During the serial recall task, listeners completed a picture-pointing task in which they heard a string of words and were asked to point to pictures on the screen in the order recalled as quickly as possible. Three kinds of words were incorporated: (a) nonrhyming nouns that are phonologically distinct, (b) rhyming nouns that are phonologically confusable, and (c) nonrhyming adjectives that are phonologically distinct but not as obviously related to pictures. The prediction was that within each group (CI or NH) recall accuracy

would be poorer for the rhyming nouns than for the non-rhyming words because rhyming words are phonologically confusable. However, the difference between nonrhyming and rhyming was predicted to be smaller for CI users than for NH peers because of anticipated problems in phonological sensitivity that likely are due to periods of degraded sensory input. This prediction was based on the premise that CI users would experience difficulty encoding phonological information because of the degraded quality of the sensory input but possibly also because of degradation of their phonological representations. Support for this idea comes from several early studies by Lyxell and colleagues (Andersson, 2001; Lyxell et al., 1998, 2003) and more recent studies by Classon and colleagues (Classon, Rudner, Johansson, & Rönnberg, 2013; Classon, Rudner, & Rönnberg, 2013): adults with severe to profound hearing loss have repeatedly demonstrated poorer performance on visually presented rhyme-judgment tasks. The authors of those articles concluded that cognitive deficits for adults with hearing loss, relative to those with NH, were primarily related to deficits in phonological processing.

During our serial recall task of verbal WM, response times were also measured, and the prediction was that these times would be slower for the adjectives than for the nouns because more processing is involved in matching adjectives to pictures. However, no Condition × Group interaction was predicted because degraded sensory input was not anticipated to have affected processing speed for these CI users.

The second objective of the current study was to examine the phonological contributions to WM for these listeners with CIs, specifically as a way to further explore potential problems in storage. To address this objective, two measures of phonological processing were collected: (a) a nonauditory lexical decision task of letter strings differing in phonological regularity and (b) an audiovisual task requiring repetition of short, nonword stimuli. The first of these tasks was designed to assess internal phonological representations, and the second was designed to assess how well participants could recover those representations through sensory inputs. The hypothesis was that CI listeners demonstrate deficits relative to NH peers on both of those measures, and these deficits predict deficits in WM accuracy, especially in the condition that benefits most from phonological sensitivity—that is, serial recall of nonrhyming words.

The third objective of this study was to examine the contribution of aging to verbal WM performance of adult CI users. Aging is associated with declines in both the storage and the processing components of WM in listeners with NH, but better phonological abilities help to ameliorate the detrimental effects of aging for older listeners, specifically through effects on storage (Nittrouer et al., 2016). Whether better phonological capacities in adult CI users can likewise mitigate the effects of advancing age could have implications for rehabilitative approaches for patients with hearing loss.

The fourth objective of this study was to test the hypothesis that verbal WM abilities or phonological sensitivity, as an independent effect, predicts speech recognition scores, especially for CI users. If supported, that hypothesis suggests that either verbal WM or phonological skills can serve as a reasonable target for aural rehabilitation. Some early, although not entirely consistent, evidence suggests that WM training may improve speech recognition for CI listeners (Ingvalson & Wong, 2013; Kronenberger, Pisoni, Henning, Colson, & Hazzard, 2011; but see Oba, Galvin, & Fu, 2013). That training focused on processing abilities. No study has yet tested the potential value of providing training focused specifically on phonological sensitivity.

In summary, a group of experienced adult CI users with postlingual deafness and a group of age- and gender-matched peers with age-normal hearing were tested using measures of verbal WM (forward and backward digit span and serial recall of nonrhyming nouns, rhyming nouns, and nonrhyming adjectives). Participants were also assessed using measures of phonological capacities (using a lexical decision task and repetition of short nonwords) and a measure of sentence recognition in noise.

Method

Participants

Sixty adults participated in this experiment. Thirty were experienced CI users between the ages of 50 and 82 years who were recruited from a pool of departmental patients. CI users had various etiologies of hearing losses and ages at implantation, but all experienced a progressive decline in their hearing during adulthood. All CI participants received their implants at or after the age of 35 years and were found to be candidates for implantation on the basis of clinical sentence recognition criteria in place at that time. Participants demonstrated CI-aided thresholds that were better than 35 dB HL at 0.25, 0.5, 1, and 2 kHz measured by clinical audiologists within the 12 months before enrollment. All had greater than 9 months of CI experience. All participants used Cochlear devices (Sydney, Australia), with an Advanced Combined Encoder speech processing strategy. Thirteen CI participants used a right CI, nine used a left device, and eight used bilateral devices. Thirteen participants used a contralateral hearing aid. During testing, participants wore their devices in everyday mode, including use of any contralateral aids, and were instructed to keep the same settings throughout the experiment. Unaided audiometric assessment was performed immediately prior to testing to assess residual hearing in each ear. Regarding production abilities, none of the CI users showed any evidence of diminished intelligibility.

Thirty NH participants were tested as a control group. They were matched closely in gender and age to the CI users, meaning ages between matched individuals were within 5% of the younger individual's age. Control participants were evaluated for NH, assessed immediately before testing, defined as four-tone (0.5, 1, 2, and 4 kHz) pure-tone average thresholds better than 25 dB HL in the better-hearing ear. Recognizing that this threshold might be difficult to meet for some older participants, it was relaxed to

30 dB HL for those over 60 years of age, of which only three had a value worse than 25 dB HL. Control participants were identified among patients with nonotologic complaints in the Department of Otolaryngology and by using ResearchMatch, a national research recruitment database.

Participants underwent screening of cognitive function to ensure that they had no frank evidence of cognitive impairment or dementia. For this purpose, the Mini-Mental State Examination (MMSE) was used, which is a validated screening assessment tool for memory, attention, and the ability to follow instructions (Folstein, Folstein, & McHugh, 1975). Raw scores were converted to T scores, which are based on age and education. A T score of less than 29 is concerning for cognitive impairment, and a score of 50 represents the mean and 1 SD above or below the mean of T scores of 60 or 40, respectively. Participants with T scores less than 29 would have been excluded from data analyses, but no participants had T scores less than 29. In addition to the MMSE, all participants were assessed for basic word-reading ability with the Word Reading subtest of the Wide Range Achievement Test, 4th edition (Wilkinson & Robertson, 2006) as a basic metric of language proficiency. All participants had standard scores ≥ 85, so no participant had a score poorer than 1 SD below the mean. Although significant differences were observed between the NH listeners and the CI users for the MMSE and the Word Reading subtest, scores for neither of these independent variables were correlated with any of the dependent variables that were measured in this study. Because some of our tasks required looking at a computer monitor, a screening test of near vision also was used. All participants had corrected near vision of better than or equal to 20/30, which is the criterion used for passing vision tests in educational settings (American Academy of Ophthalmology, Pediatric Ophthalmology Strabismus Panel, 2012).

Participants in both groups were adults whose first language was American English. All participants had graduated from high school, except for one CI user who earned his General Educational Development certificate. A measure of socioeconomic status (SES) was included because SES may predict language abilities. SES was quantified with a metric defined by Nittrouer and Burton (2005), indexing the occupational and educational levels with two scales between 1 and 8 (a score of 8 indicated the highest occupation or educational level achievable). The two scores were multiplied, resulting in SES scores between 1 and 64. Demographic and audiologic data for individual CI users are shown in Table 1. Comparisons of demographic and SES scores between the 30 CI users and 30 NH listeners whose data were included in the analyses are shown in Table 2. No significant differences were found for age or SES, but CI participants did score significantly more poorly on the reading and cognitive screening tasks.

Equipment

All tasks were performed in a soundproof booth or sound-treated testing room. Audiometry was performed using a Welch Allyn TN262 audiometer with TDH-39 headphones. Auditory stimuli were presented through a Creative Labs Soundblaster sound card using a 44.1-kHz sampling rate and 16-bit digitization. Stimuli were presented via computer over a speaker placed 1 m from the participant at zero-degree azimuth. Scoring for the WM tasks was done automatically at the time of testing by the participant directly entering responses into the computer via a touchscreen. For the MMSE task and the sentence recognition task, participant responses were video- and audio-recorded for later scoring. Participants wore vests holding FM transmitters that sent signals to receivers, which provided input directly into the video camera. Responses for these tasks were scored later; two staff members could then independently score responses to check reliability. All participants were tested while using their usual devices (one CI, two CIs, or CI plus contralateral hearing aid) or no devices (for NH controls), and devices were checked at the beginning of testing by having the tester confirm sound detection by the participant through each device.

Custom software was used to present all stimuli. For the WM tasks, responses were collected using a 21-in. widescreen touchscreen monitor (HP Compaq L2105TM). For stimulus generation on sentence recognition measures, speech samples were collected from a male talker directly onto the computer hard drive via an AKG C535 EB microphone, a Shure M268 amplifier, and a Creative Laboratories Soundblaster sound card.

Stimuli and Stimuli-Specific Procedure

Five measures were used: two of WM (digit span and serial recall of words), two of phonological ability (lexical decisions and nonword repetition [NWR]), and one of recognition of words in sentences. The two WM measures were selected to examine performance on tasks with differing demands for the participant to access phonological structure to encode the stimuli. For digit span, the prediction was that listeners might be able to perform this task without detailed phonological encoding because digits are so readily recognized. For serial recall tasks, encoding strategies can apparently be more or less phonologically based, depending on the phonological sensitivity of the listener. Comparisons of accuracy scores on serial recall of nonrhyming versus rhyming nouns could provide information regarding the ability of listeners to encode and store words in verbal WM using phonological codes; comparison of response times for nonrhyming nouns versus adjectives should provide information regarding verbal WM processing abilities (Nittrouer & Miller, 1999; Nittrouer et al., 2013, 2016).

Digit Span

The first WM task was based on the test of the same name from the Wechsler Intelligence Scale for Children, 3rd edition (Wechsler, 1991). The task was presented using a computerized platform so the rate and vocal inflection of digit presentation could be consistent across participants

Table 1. Cochlear implant participant demographics.

Participant	Gender	Age (years)	Implantation age (years)	SES	Side of implant	Hearing aid	Etiology of hearing loss	Better ear PTA (dB HL)	Sentence recognition (% correct words)
1	Female	64	54	24	Both	No	Genetic	120.0	96.0
2	Female	66	62	35	Right	Yes	Genetic, progressive as adult	78.8	59.2
3	Male	66	61	18	Left	No	Noise, Meniere's disease	82.5	66.4
4	Female	66	58	12	Right	Yes	Genetic, progressive as adult	98.8	92.0
6	Male	69	65	24	Right	No	Genetic, progressive as adult	88.8	76.0
7	Male	58	52	36	Both	No	Rubella, progressive	115.0	25.6
8	Female	56	48	25	Right	Yes	Genetic, progressive	82.5	84.0
9	Male	79	67	49	Left	No	Genetic	120.0	0.0
10	Male	79	76	36	Right	Yes	Progressive as adult, noise, sudden	70.0	73.6
12	Female	68	56	12	Both	No	Otosclerosis, progressive as adult	112.5	25.6
13	Male	54	50	24	Both	No	Progressive as adult	120.0	84.8
16	Female	62	59	35	Right	No	Progressive as adult	115.0	17.6
19	Female	75	67	36	Left	No	Progressive as adult, autoimmune	120.0	1.6
20	Male	78	74	15	Left	No	Ear infections	108.8	0.0
21	Male	82	58	42	Left	Yes	Meniere's disease	71.3	55.2
23	Female	80	73	30	Right	No	Progressive as adult	87.5	35.2
25	Male	58	57	24	Right	Yes	Autoimmune, sudden	120.0	3.2
28	Male	77	72	12	Both	No	Progressive as adult	120.0	0.8
31	Female	67	62	25	Left	Yes	Progressive as child	102.5	16.8
34	Male	60	54	42	Left	Yes	Noise, Meniere's disease, sudden	98.8	1.6
35	Male	68	62	42	Both	No	Genetic, progressive as adult, noise	120.0	68.8
37	Female	50	35	35	Both	No	Progressive as child	120.0	97.6
38	Male	75	74	35	Left	Yes	Ototoxicity	96.3	3.2
39	Female	63	61	30	Right	No	Progressive as adult	107.5	16.0
40	Female	66	59	15	Both	No	Genetic, Meniere's disease	120.0	73.6
41	Female	59	56	15	Right	Yes	Sudden	87.5	60.8
42	Male	82	76	42	Right	Yes	Progressive as adult, noise	68.8	61.6
44	Female	72	66	25	Right	No	Progressive as adult	98.8	7.2
46	Male	75	74	42	Left	Yes	Progressive as adult	87.5	0.0
48	Female	78	48	15	Right	Yes	Progressive as adult	110.0	12.0

Note. SES = socioeconomic status; PTA = pure-tone average.

and response time could be measured precisely. Digits were recorded by a single college-age male talker and presented sound field at a rate of one digit per second. Participant responses consisted of tapping on digits on a computer

Table 2. Demographics for participants with normal hearing (NH) and cochlear implants (CIs).

Characteristic	NH (*n* = 30)		CI (*n* = 30)			
	M	*SD*	*M*	*SD*	*t*	*p*
Age (years)	68.3	9.4	68.4	8.9	0.03	.98
Reading (standard score)	107.0	12.5	100.5	11.1	2.13	.04
MMSE (*T* score)	55.8	10.7	49.8	9.4	2.29	.03
SES	34.0	13.9	28.9	10.9	1.55	.13

Note. MMSE = Mini-Mental State Examination; SES = socioeconomic status.

screen, so verbal responses were not required. To administer the task, the participant sat in front of the monitor with the hand used to respond on the table directly in front of the screen. First, participants completed a pretest in which all digits were shown on the monitor, and the participant heard each one presented. The participant tapped each digit as it was heard. This pretest ensured that all participants could match the digits heard to their numerical representations; no participant had difficulty completing this test. During testing, digits were not shown on the screen. The participant heard the sequence, and then digits appeared at the top of the monitor. The participant was asked to tap the digits in the order heard, as quickly as possible. As they did, each digit moved to the vertical middle of the screen, ordered left to right in the order tapped. Each sequence length was presented twice with different digit orders, and the length just preceding the sequence length at which both sequences were incorrect was taken as that participant's

digit span. Both forward and reverse digit spans were computed as were response time per digit.

Serial Recall of Words

For the second WM task, stimuli that have been used previously were used here (e.g., Nittrouer et al., 2013). These stimuli consisted of three sets of six words each: nonrhyming nouns, rhyming nouns, and nonrhyming adjectives. The purpose of including these three conditions was to differentially examine the storage and processing components of WM. Participants with NH would be expected to show poorer accuracy on rhyming versus nonrhyming words because rhyming words can have greater phonemic similarity interference effects (Goldstein, 1975). In contrast, response times for the adjective condition was expected to be longer than that for the nonrhyming noun condition because this task should impose a greater processing demand on the listener (Nittrouer et al., 2016). The nonrhyming nouns used were *ball, coat, dog, ham, pack,* and *rake.* Rhyming nouns used were *bat, cat, hat, mat, Pat* (represented by a picture of a woman), and *rat.* The adjectives used were *big* (represented by a picture of a big dog next to a small dog), *deep* (a deep swimming pool), *full* (a full glass of water), *hot* (a steaming cup of coffee), *sad* (a crying child), and *wet* (a wet cat). All words were spoken and recorded by a male talker. Words could not be equated across lists based on frequency of occurrence because of the restrictions on list construction. However, participants were familiarized with the words to be used before testing, so they knew what words were in each set. Thus, word frequency within the lists should not have affected lexical retrieval and memory during this closed-set task: once words in the lexicon have been activated, recognition probabilities are equalized across the set (Miller, Heise, & Lichten, 1951). Nonetheless, the mean frequency of occurrence per one million words was obtained for each word using the counts of Brysbaert and New (2009): 51 for the nonrhyming nouns, 26 for the rhyming nouns, and 156 for the adjectives.

Prior to testing, the participant saw a series of six blue squares and was required to tap the squares in order from left to right as quickly as possible. Five trials were completed, and average time across those trials (the calibration time) was used to normalize response times to test items. Testing then began. The order of presentation of the three types of lists was randomized across participants. For each list type, the participant was trained to associate pictures with words by seeing the pictures at the top of the monitor and hearing each word presented by itself. The participant needed to tap the picture representing that word to indicate that the association was made. This procedure was done before and after testing as a way of verifying that the participant recognized the words. During testing, words were presented at a rate of one per second without the pictures being shown; following presentation of the six words, all the pictures appeared at once (randomly positioned). The participant was instructed to tap the pictures in the order heard, again as quickly as possible. Ten trials of each condition were included. Both response accuracy and response time

per item, normalized for general response time and computed as the average time per trial minus calibration time, were used as dependent measures.

Lexical Decision

This task examined participants' propensities to recover phonological structures and has been used in a study of younger and older listeners with NH (Nittrouer et al., 2016). In this task, letter strings were presented on a computer monitor, and the participant had to decide whether the string was a real word. Items were divided into five categories, each with 32 items: (a) high-frequency, phonemically regular real words (e.g., *dog, father,* and *song*); (b) low-frequency, phonemically less regular words (e.g., *aisle, ewe,* and *ostrich*); (c) homophones of real words with a wide range of frequencies (e.g., *fraun, oshin,* and *toste*); (d) nonwords that are somewhat phonemically regular (e.g., *drint, kalife,* and *snald*); and (e) nonwords that are best described as letter strings (e.g., *cifkr, pljuf,* and *zcbnm*). Mean frequency of occurrence for the real words, according to Brysbaert and New (2009), was 182 (range, 42 to 774) for the Category 1 words, 2 (range, < 1 to 7) for the Category 2 words, and 217 (range, < 1 to 5,721) for the words corresponding to the homophones of Category 3.

During testing, items appeared on the computer screen in large letters, one item at a time. The participant's task was to decide as quickly as possible whether the item was a real word. When the decision was that the item was a real word, the participant hit one key, marked in green. When the decision was that the item was not a real word, the participant hit a different key, marked in red. Response time served as the measure of primary interest, but percentage of correct responses was also examined to ensure that any observed group differences were not related to spelling or lexical abilities. The prediction was that participants with intact phonological skills would respond more slowly to nonword homophones as a result of a propensity to automatically recode orthographic text into phonological codes. Because the recoded phonological sequences produced real words, interference in decision making occurs, slowing response times.

Nonword Repetition

This task was used to assess participants' phonological processing, with recognition of items optimized by combined audiovisual presentation of stimuli. Sixteen nonwords between one and four syllables in length, developed by Dollaghan and Campbell (1998), were video- and audio-recorded by a female talker who is a trained phonetician. Equal stress was placed on all syllables for all stimuli, and fundamental frequency was kept consistent and flat. Stimulus amplitude was constant. Instructions for the task were placed at the start of the video recording of the stimuli. During the task, participants saw and heard the talker saying each nonword and were asked to repeat each nonword immediately. Four nonwords were presented at each syllable length. Participant responses were recorded and scored later. For this

task, phonemes were scored as wrong when they were omitted or when substitutions were used. Distortions were not scored as wrong. Total percentage of correct phonemes across all syllable string lengths was used in analyses.

Recognition of Words in Sentences

The sentences used to examine speech recognition were five-word, highly meaningful sentences. Twenty-seven of the 72 five-word sentences (two for practice and 25 for testing) used by Nittrouer and Lowenstein (2010) were used here. These sentences are semantically predictable and syntactically correct and follow a subject-predicate structure (e.g., *the sun melted the snow*). They originally came from the Hearing in Noise Test (Nilsson, Soli, & Sullivan, 1994). To avoid ceiling and floor effects, participants were tested in different amounts of speech-shaped noise, with the presentation of signal and noise at 68 dB SPL. For CI participants, the SNR was +3 dB; NH listeners were tested at −3 dB SNR. Percentage of correct words was the measure of interest.

General Procedure

All procedures were approved by The Ohio State University Institutional Review Board. Participants were tested in one session of approximately 2 hr. Hearing thresholds and screening measures were obtained first. The order of presentation of tasks was randomized across participants, except that serial recall was always followed by digit span.

Results

All data were screened for normal distributions and homogeneity of variances. An α of .05 was set, but significance was considered to be $p < .10$. At $p > .10$, outcomes were considered not significant.

Reliability

An estimate of interscorer reliability was obtained for the tests that involved audiovisual recording and later scoring of responses. Responses were scored by one trained scorer and scored a second time by the third author for 25% of all participants (14 participants). Mean agreement in scores for the two scorers across these participants ranged from 92% to 100% for the measures of word reading, sentence recognition, and NWR. These outcomes were considered to indicate good reliability, and the scores from the staff member who initially scored all the samples were used in further analyses.

WM: CI versus NH

The first question of interest was whether CI users had poorer WM than did age-matched NH peers. The hypothesis was that CI users have poorer WM, primarily as a result of deficits in phonological storage rather than processing. The prediction was that accuracy of responses on the tasks of WM would be poorer for CI users, but

response times for those same measures would be similar. Two types of verbal WM tasks were used: digit span (a commonly used measure of WM) and serial recall of monosyllabic words, which in general would be expected to require more phonological processing. The serial recall task should reveal differences between CI and NH groups in use of phonological coding on the basis of accuracy for nonrhyming nouns and rhyming nouns and differences in processing times for nonrhyming nouns and adjectives.

Starting with forward and backward digit span, all 30 participants from each group successfully completed the tasks, so their data were included in analyses. Figure 1A shows group mean forward and backward digit span scores; Figure 1B shows mean response times for the same tasks. No obvious group differences in accuracy or response times are apparent. For analyses, first a two-way repeated-measures analysis of variance (ANOVA) was performed with digit span as the dependent measure, condition (forward or backward) as the within-participant variable, and group (CI vs. NH) as the between-participants variable. Results showed a significant effect of condition, $F(1,58) = 81.39$, $p < .001$, $\eta^2 = .584$, but no Condition × Group interaction. No significant effect of group was found. Thus, participants scored more poorly on backward span than on forward span, but CI and NH listeners had similar spans. A similar analysis was then performed for response times. Again, a significant effect of condition was found, $F(1,58) = 81.39$, $p < .001$, $\eta^2 = .592$, but there was no Condition × Group interaction. No significant effect of group was identified. Thus, participants had longer response times for backward span than for forward span, but CI and NH listeners had similar times.

For serial recall, data were excluded from analyses for two CI participants and one NH listener because response times were longer than 10 s. This time was so much longer than that of other participants and so much longer than times needed by these participants on the initial calibration trials that the researchers concluded that these participants were not responding "as quickly as possible." Data were excluded also for four CI users for rhyming words and for one CI user for adjectives because those participants could not pass the task of associating the words with the pictures, either before or after testing.

The first hypothesis regarding the serial recall task was that CI users have poorer verbal WM accuracy and show a relatively smaller difference in accuracy scores between the nonrhyming and rhyming words. NH listeners should be able to take greater advantage of the phonological distinctiveness of the nonrhyming nouns relative to the rhyming words and so show a larger discrepancy between these accuracy scores. Figure 2A shows accuracy scores for serial recall of nonrhyming nouns, rhyming nouns, and nonrhyming adjectives across list positions for each group separately. Typical primacy and recency effects are evident for both groups, with better recall of first and final words, respectively. CI users appear to have slightly poorer recall accuracy across conditions, and a smaller difference in accuracy is apparent between nonrhyming and rhyming nouns for CI users compared with NH controls. To test
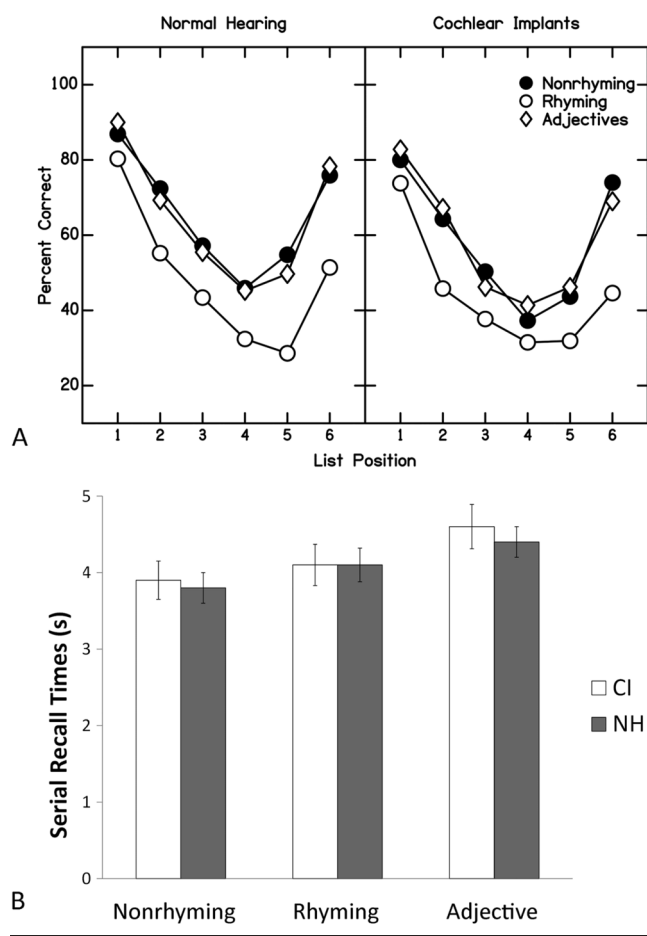
Figure 1. Digit span accuracy (A) and response times (B). Error bars represent standard errors of the means. CI = cochlear implants; NH = normal hearing.

Figure 2. Serial recall accuracy (A) and response times (B) for nonrhyming words, rhyming words, and adjectives. Error bars represent standard errors of the means. CI = cochlear implants; NH = normal hearing.
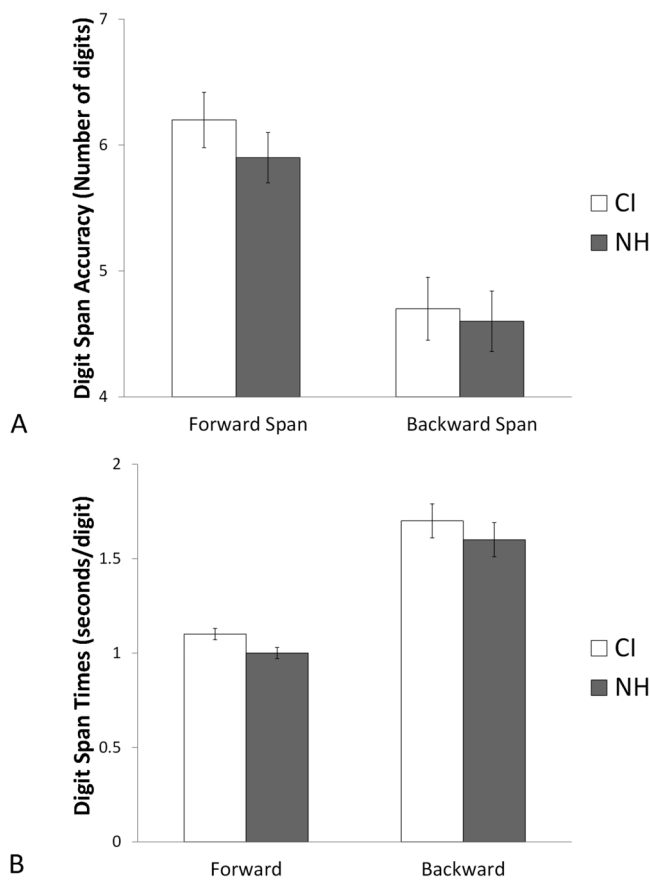
this observation, a two-way repeated-measures ANOVA was performed with serial recall accuracy scores as dependent measures, but only scores for the nonrhyming and rhyming nouns were used to maintain consistency in word class. Because responses across list positions were as expected for both groups, mean percent correct scores across all positions were used as the dependent measure. Results showed a significant effect of condition, $F(1,51) = 54.10$, $p < .001$, $\eta^2 = .515$, but no Condition × Group interaction was identified. These analyses revealed no significant evidence of a disproportionately greater benefit from phonological encoding for the NH listeners and no significant group effect, although the differences were close to significant, $F(1,51) = 3.06$, $p = .086$, $\eta^2 = .057$. This finding prompted further inspection, so individual mean scores across all three conditions were calculated. Mean accuracy was 55.3% ($SD = 14.4$) for CI users and 59.6% ($SD = 13.3$) for NH listeners. Although not significant, Cohen's d for this difference was 0.31.

The next question regarding serial recall was whether CI users would differ from NH peers in verbal WM processing speed. Figure 2B shows mean response times for each group separately. Listeners in both groups appear to have

similar response times, and both groups appear to have similar degrees of prolongation in response times when recalling adjectives. To examine this observation more closely, a two-way repeated-measures ANOVA was performed with serial recall response time as the dependent measure. Only response times for nonrhyming nouns and adjectives were used, again to maintain as much consistency across conditions to be compared as possible. Response times were computed as mean time across the 10 trials in each condition minus the calibration time for that participant. This approach was taken, even though a t test revealed that CI users and NH listeners did not differ significantly in calibration time, because it controlled for individual differences in speed of responding. Results revealed a significant within-participant effect of condition, $F(1,54) = 18.39$, $p < .001$, $\eta^2 = .254$; no Condition × Group interaction was identified nor was there a significant group effect. These findings suggest that CI users and NH listeners have similar processing speeds and similar degrees of slowed processing during serial recall for adjectives as compared with nonrhyming words.

The findings from the digit span and serial recall tasks provided modest support for our first hypothesis: CI users perform similarly to NH listeners on digit span tasks but are slightly less accurate on serial recall tasks. No group difference was apparent for the effect of phonological similarity. Processing speeds were indistinguishable.

Phonological Contributions to WM

Our first hypothesis—that WM storage is poorer for CI users than NH listeners but processing is similar—was supported to some extent. The second hypothesis was that declines in WM storage abilities are attributable specifically to diminishing sensitivity to the phonological structure by CI listeners relative to NH listeners. This hypothesis arose from previous findings that adults with acquired hearing loss of clinical significance have diminished phonological capacities (Andersson, 2001; Classon, Rudner, Johansson, & Rönnberg, 2013; Classon, Rudner, & Rönnberg, 2013; Lyxell et al., 2003). To investigate this possibility in adult CI users, two measures of phonological capacities were used: a lexical decision task and a task of NWR.

First, responses for the lexical decision task were examined. Figures 3A and 3B show response accuracy for the lexical decision task for participants in the two groups. Accuracy was defined by the percentage of responses in which the word was judged correctly either as a real word (high-frequency and low-frequency words) or as a nonword (homophones, phonologically regular nonwords, and phonologically irregular nonwords). Although some variability was noted in response accuracy, mean performance was better than 92% correct in all conditions for all listeners. A two-way repeated-measures ANOVA was performed with lexical decision accuracy scores as dependent measures, lexical condition as the within-group variable, and group as the between-participants variable. Results revealed a significant effect of condition, $F(4,232) = 14.94$, $p < .001$, $\eta^2 = .205$, but there was no significant Condition × Group effect nor was there a significant group effect. Thus, CI users and NH listeners were similar in their abilities to decide the lexical status of the test items when reading items presented visually.

For the response times for the lexical decision task, Figures 4A and 4B show group mean response times for each condition. Adults in both groups showed similar response time patterns: they were fastest for the items that were most clearly real words or nonwords (high-frequency real words and phonologically irregular nonword conditions) and slowest for the homophone nonwords. A two-way repeated-measures ANOVA performed on these response times revealed a significant effect of condition, $F(4,232) = 77.35$, $p < .001$, $\eta^2 = .571$, but again there was no significant Condition × Group interaction and no significant group effect. In summary, CI and NH listeners seem to have engaged in phonological recoding to a similar extent when reading letter strings, suggesting that this phonological skill did not decline as a result of auditory deprivation. Therefore, CI users were clearly capable of performing this phonological task when reading and had intact phonological representations.
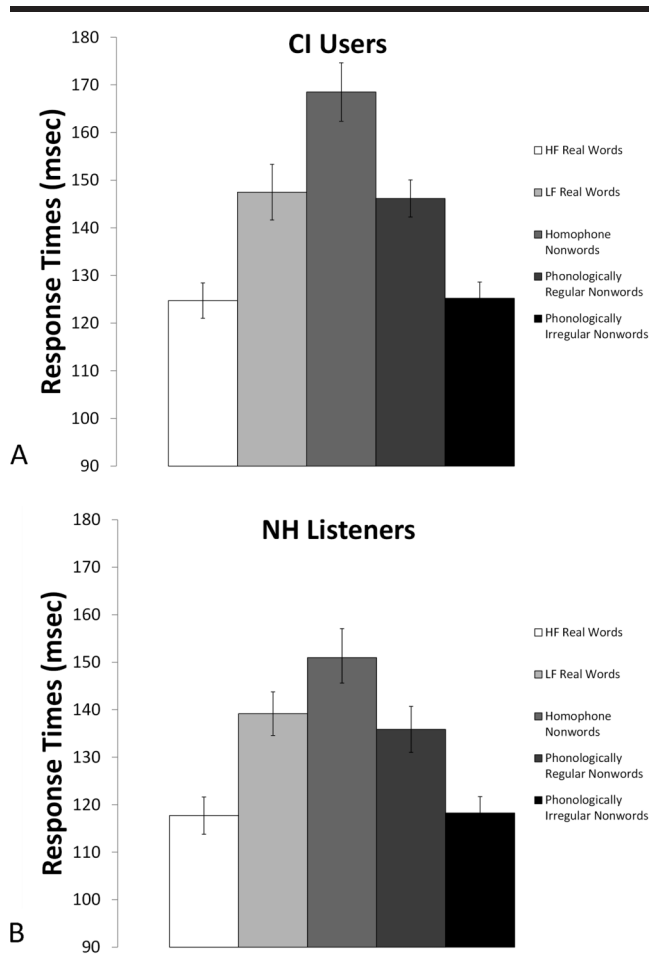
Figure 3. Lexical decision task accuracy for cochlear implant (CI) users (A) and listeners with normal hearing (NH) (B). Error bars represent standard errors of the means. HF = high frequency; LF = low frequency.
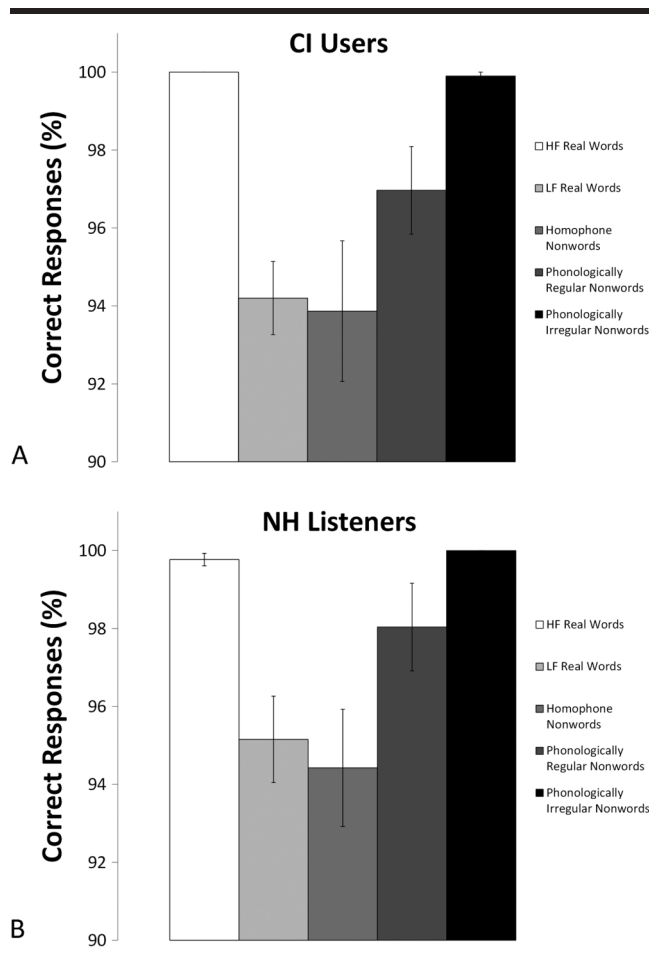

Responses during the NWR task also were examined. Figure 5 shows group mean accuracy scores for the nonword stimuli of various lengths (one to four syllables). The results shown in Figure 3 indicate that CI users performed more poorly when asked to repeat nonwords and that the performance gap between CI and NH participants was larger for longer syllable strings than for single-syllable strings, possibly because of a ceiling effect for the NH participants on one-syllable strings. Mean total phoneme percent correct score across all syllable lengths was 71.7% ($SD = 12.6$) for CI users and 89.6% ($SD = 5.0$) for NH listeners. A t test revealed a highly significant difference in NWR scores between CI users and NH peers, $t(58) = 7.19$, $p < .001$, with a Cohen's d of 1.87. These results confirm that phonological sensitivity was poorer in CI users.

The next question was whether this deficit in phonological sensitivity could explain the poorer performance of CI users, most notably on the serial recall WM task of nonrhyming words, which should require the participant to encode and store words using phonological structure.

Moberly et al.: *Speech Recognition in Adults With Cochlear Implants* **1055**

Figure 4. Lexical decision task response times for cochlear implant (CI) users (A) and listeners with normal hearing (NH) (B). Error bars represent standard errors of the means. HF = high frequency; LF = low frequency.

Figure 5. Mean nonword repetition (NWR) scores across syllable lengths for cochlear implant (CI) users and listeners with normal hearing (NH). Error bars represent standard errors of the means.
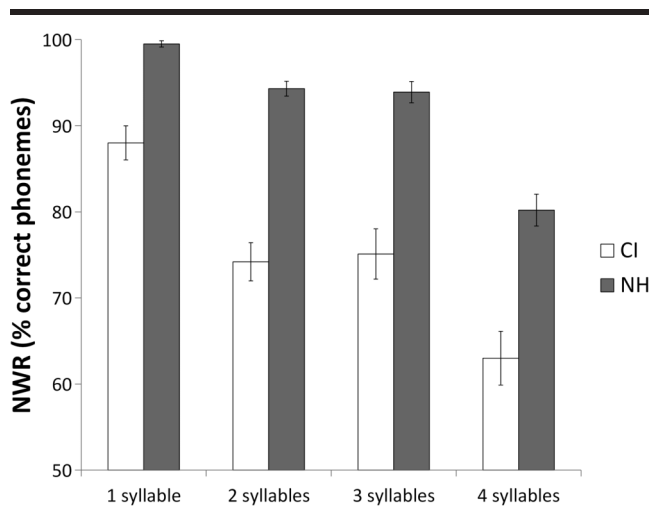
To answer this question, bivariate correlation analyses were performed for each group separately between NWR accuracy scores and the serial recall accuracy for nonrhyming words. For CI users, no correlation was found. However, for NH participants, NWR scores were correlated with accuracy scores for serial recall of nonrhyming words, $r(29) = .63$, $p < .001$. Thus, for these listeners with CIs, phonological sensitivity, as indicated by NWR accuracy scores, did not explain within-group variability in verbal WM accuracy on serial recall of nonrhyming words. However, this lack of a within-group correlation does not necessarily mean that the large group difference in phonological sensitivity did not account for the group difference in verbal WM accuracy that was observed for serial recall. Instead, it seems that CI users were not strongly using phonological codes at all to store words in WM; if they had been, a relationship between NWR and WM accuracy would have been observed, as was found for NH listeners.

Effects of Age on Verbal WM

The third main goal of this study was to examine age effects on verbal WM in adult CI users. This goal was formulated on the basis of results from the recent study by Nittrouer et al. (2016) in which older NH adults had poorer verbal WM accuracy and longer response times than did younger NH controls. Although younger and older adults had similar phonological capacities in that study, better phonological capacities ameliorated some of the detrimental effects of advancing age on WM for the older group. The goal here was to examine whether phonological capacities would mitigate the effects of advancing age on verbal WM in this group of CI users.

Bivariate correlation analyses of age with verbal WM scores were computed for each group separately (see Table 3). For both groups, advancing age was correlated with poorer and/or slower performance on tasks across several measures of verbal WM. However, advancing age was also negatively correlated with NWR scores of phonological sensitivity for the CI group, $r(28) = -.52$, $p = .003$; for the NH group, age and NWR were not correlated. Thus, for CI users, advancing age was associated with poorer verbal WM and poorer phonological sensitivity. These two situations appear to exist, even though CI listeners were not relying on phonological structure for encoding words into a short-term memory buffer. Nonetheless, declines in both abilities could impact speech recognition for these CI users independently.

An important consideration was whether advancing age might be serving as a surrogate for duration of deafness, which can negatively affect phonological capacities for adults with severe to profound hearing loss (Lyxell et al., 1998). A correlation analysis was performed for CI users between age and duration of deafness (computed as current age minus reported age at onset of hearing loss). Age and duration of deafness were not significantly correlated. Thus, the identified declines in verbal WM and phonological sensitivity in CI users appear attributable to advancing

Table 3. Bivariate correlation coefficients for normal hearing (NH) and cochlear implant (CI) groups among age and measures of verbal working memory.

Group	Accuracy					Response time				
	Span		Accuracy			Span		Time		
	Forward	Backward	Nonrhymes	Rhymes	Adjectives	Forward	Backward	Nonrhymes	Rhymes	Adjectives
NH age	−.23	−.49**	−.30	−.50**	−.51**	.16	.29	.44*	.32	.23
CI age	−.33	−.21	−.51**	−.69**	−.52**	.45*	.17	.60**	.18	.46*

*$p < .05$. **$p < .01$

age and not to longer duration of hearing loss. However, advancing age did not impact phonological sensitivity in the NH group.

Relationships of WM and Phonological Sensitivity With Sentence Recognition

The next question posed was whether verbal WM abilities or phonological sensitivity explained accuracy of recognition of words in sentences. Looking at WM first, the prediction was that better WM skills (either storage accuracy or processing speed) would facilitate better recognition of rapidly processing degraded speech signals. Mean recognition scores of words in sentences in noise were 40.5% correct ($SD = 35.0$) for CI users and 81.7% ($SD = 9.3$) for NH listeners. These group mean scores were not compared directly because CI listeners were tested at +3 dB SNR and NH listeners were tested at −3 dB SNR to prevent ceiling and floor effects within each group. Nonetheless, the poor performance of CI users in spite of the SNR used for presentation highlights just how difficult it is for CI users to recognize speech in degraded conditions. The large standard deviation highlights the great variability in performance within this clinical population.

Because some CI users scored close to the floor (see Table 1), arcsine transformations of sentence recognition scores were computed, and correlation analyses were performed using the results. Accuracy scores and response times for serial recall of nonrhyming words were used as the WM variables of interest because among the WM tasks included in this study, this condition is most apt to represent the verbal WM capacity of an individual listening to spoken language, consisting of relatively phonologically distinct and rapidly presented words in sequence. Bivariate correlation analyses were performed for each group separately, correlating scores of sentence recognition in noise (percentage of words correct) and verbal WM accuracy scores for serial recall of nonrhyming words and the response times for the same task. For neither group was a significant outcome found. Thus, verbal WM did not explain speech recognition for these listeners. However, this result may have been due to the fact that the sentences were short. Different outcomes may have been obtained for longer sentences.

For phonological sensitivity, the hypothesis was that better sensitivity to phonological structure facilitates better speech recognition because better predictions can be made of that structure in the speech signal. Overall, phonological sensitivity was poorer for these CI users when required to access phonological structure from an auditory (or audiovisual) signal, but the CI users did not seem to use phonological structure to encode words into a WM buffer, as the NH listeners did. Thus, outcomes were difficult to predict for this speech recognition task; accessing phonological structure through the auditory signal may be so difficult for these listeners that they cannot use this structure for any linguistic processes. However, when bivariate correlation analyses were performed for each group separately, correlating scores of sentence recognition and NWR scores, a significant outcome was obtained for the CI users, $r(28) = .50$, $p = .007$, but not for the NH listeners. Although impaired, better phonological skills seemed to aid these CI users in recognition of degraded signals.

Effects of Age on Speech Recognition

The final set of analyses was conducted to examine how age affected spoken word recognition for CI users and NH listeners separately. Bivariate correlational analyses were performed using age and word recognition for these sentences. Significant effects were observed for both the CI users, $r(28) = −.43$, $p = .023$, and the NH listeners, $r(29) = −.45$, $p = .014$. The question originally asked was whether phonological sensitivity, which can affect speech recognition for CI users, could help to ameliorate the consequences of advancing age for these listeners. However, results already obtained revealed that phonological sensitivity was adversely affected by advancing age in CI users. A related question, then, was whether the negative effect of age on recognition of words in sentences in CI listeners could be attributed to those age-related declines in phonological sensitivity. To answer that question, partial correlation coefficients were obtained for age and speech recognition, controlling for the effect of phonological sensitivity. This approach revealed in a lack of a significant relationship between age and speech recognition for CI users, indicating that the effect of advancing age on speech recognition was completely explained by declining phonological sensitivities. For NH listeners, that relationship between age and speech recognition remained significant, partial $r(26) = −.48$, $p = .011$, as would be expected due to the previously demonstrated lack of relationship between phonological sensitivity and speech recognition in these listeners.

Discussion

Speech recognition is usually impacted significantly by postlingual hearing loss and subsequent CI, although the magnitude of that impact varies widely. This study was conducted to examine the contributions to variability in outcomes of WM, phonological sensitivity, and aging. The study was motivated by findings from earlier work demonstrating that (a) phonological sensitivity underlies individual differences in verbal WM in older adults with NH and in children with CIs, (b) auditory deprivation in adults with acquired deafness results in degradation of phonological sensitivity, and (c) verbal WM skills are important for making sense of impoverished auditory input during spoken language recognition for individuals with hearing loss, including those listening through hearing aids. On the basis of these previous findings, the current study was conducted to test four hypotheses: (a) CI users have poorer verbal WM skills relative to age-matched NH peers, and group differences are most pronounced on tasks that rely heavily upon phonological sensitivity; (b) CI users have deficits in phonological skills, relative to their NH peers, as a result of auditory deprivation and their experience listening to impoverished input through their CIs; (c) advancing age is associated with declines on tasks of verbal WM; and (d) verbal WM performance or phonological sensitivity predicts recognition of words in sentences.

Several factors examined in this study contribute to speech recognition, especially under conditions of degradation such as while listening in noise or with a CI. During continuous speech recognition, as encountered in most daily living situations, the listener can use knowledge of topic and language structures, including phonological structure, to form hypotheses of what is being said. In conditions where the listener has stronger knowledge of topic and language structures, less sensory evidence is required to confirm those hypotheses. It also helps to be able to process sensory information quickly, before it degrades. This rapid processing would be especially useful when the signal is degraded to start with, such as in conditions of noise or when listening through a CI. The ability to store more sensory information— that is, having larger spans—also would be useful. Past studies of WM in listeners with hearing loss have supported these suggestions by showing that larger memory spans facilitate speech recognition for listeners with hearing loss (e.g., Foo et al., 2007; Lunner, 2003). The current study was conducted to explore further the influence of WM on speech recognition and the potential influences of sensitivity on phonological structure and aging.

The results of this study only partially support our hypotheses. Addressing the first and second hypotheses—that phonological sensitivity underlies verbal WM capacity and phonological sensitivity is eroded by deafness—participants with CIs performed on par with NH peers on measures of verbal WM that did not explicitly tax phonological skills: forward and backward digit span. However, on tasks of WM that placed greater demands on phonological capacities, serial recall of words, CI users were slightly less accurate,

suggesting that poor phonological sensitivity accounted for the difference in performance; response times were similar. Thus, any deficit on this serial recall task experienced by the CI users can be attributed to a problem in storage, not processing. But the novel finding here was that the listeners with CIs did not appear to use phonological structure for the most part to store words in a short-term memory buffer. Similar conclusions have been offered regarding children with dyslexia (Brady, Shankweiler, & Mann, 1983) and children with CIs (Nittrouer et al., 2013). This finding is new for adults who previously had normal hearing and presumably typical phonological sensitivities. Investigators working with children have explained this outcome as evidence that these listeners must rely on broader acoustic forms for storing language in a memory buffer than the acoustic structure that supports recovery of phonemic segments. This broader acoustic structure is best described as the relatively slow undulations in whole spectral shapes, a form that is well preserved in the auditory system of a CI user; more detailed structure is lacking because of limits on CI signal processing and the spread of excitation. Thus, although WM accuracy was not very different for NH and CI listeners in the current study, listeners in these two groups were apparently relying on different kinds of codes for storage; the NH listeners used the phonological codes that derive from the phonological loop, and the CI users needed a coarser kind of code.

The suggestion that CI users have access to only broad patterns of spectral structure is also supported by the finding of highly deficient phonological sensitivity, as measured by the most robust index, the NWR task. However, when presented with the exclusively visual stimuli of our lexical decision task, CI users performed on par with NH listeners, reflecting intact abilities to perform phonemic recoding during reading. Thus, CI users can access phonological structure when processing visually presented language but are hindered in their abilities to do so during listening tasks. They must have intact internal phonological representations that they cannot access through acoustically presented signals. These CI users did not appear to encode test items in the serial recall task using phonological structure.

Regarding our third hypothesis—that advancing age deleteriously affects verbal WM skills in both groups— results largely confirmed the predictions. Advancing age in CI users was associated with declining phonological sensitivity. This finding is in agreement with previous work that demonstrated a negative correlation of phonological skills and age in CI users (Moberly et al., 2016).

Our fourth hypothesis pertained to the relationships between verbal WM, phonological sensitivity, and recognition of words in sentences. Among CI users and NH listeners, recognition of words in sentences was not related to WM accuracy or speed of processing. Thus, WM capacity was unable to compensate for any phonological deficit in these CI users because capacity did not affect speech recognition for these test materials. This result may appear to contrast with those of other investigations in younger adults with hearing impairment where WM capacity seems

to have a protective effect. For example, Classon, Rudner, Johansson, and Rönnberg (2013) used a visually presented rhyme-judgment task to address this question. On this task, participants had to decide whether two words presented orthographically rhymed. Results showed that participants with hearing loss but good WM capacity performed similarly to NH participants on rhyme judgment; in contrast, performance of participants with hearing loss and poor WM capacity was significantly poorer than that of NH participants, even those with poor WM capacity. However, it is difficult to compare outcomes across studies because the dependent measure is so different: visual rhyme judgment versus auditory speech recognition. Classon et al. also found that the participants with hearing loss and good WM capacity performed exceptionally poorly when asked to recall words in the rhyme judgment task.

Although potentially in conflict with some previous results, the findings of the current study dovetail with earlier findings by our group and others with regard to the influence of phonological skills on verbal WM and speech recognition. Lyxell et al. (1998) in their study of adults with postlingual deafness who received CIs reported that performance was significantly poorer among deaf adults, relative to NH participants, on tasks in which use of phonological representations was a key task demand (i.e., rhyme judgment and lexical decision making). When phonological representations were less important, such as during a task of reading span, differences between deaf adults and NH controls were less prominent. Lyxell et al. also found that when they grouped participants based on functional communicative ability at 12 months postimplantation, those who were in the highest functioning group (i.e., on the basis of the ability to understand a speaker outside of direct view) had preoperative scores commensurate with those of NH controls on tasks with large phonological demands. This finding is consistent with recent work by our group (Moberly et al., 2016) demonstrating that auditory deprivation experienced by adults with postlingual deafness and CIs led to declines in phonemic sensitivity indexed by an audiovisually presented assessment of initial consonant choice and final consonant choice and that phonemic sensitivity predicted 25% to 40% of variability in word recognition in quiet. In contrast with the findings by Lyxell et al., in the current sample of CI users and in our earlier work, older age was associated with poorer phonological skills in CI users, but duration of deafness did not predict the extent of degradation of phonemic sensitivity. However, we may be limited in our ability to draw meaningful conclusions regarding duration of hearing loss because participants' subjective reports, which are subject to recall bias, were our source of this information, and we were unable to confirm when their hearing loss progressed to a severe degree.

Although our findings suggest that phonological skills play a pivotal role in speech recognition by adult CI users with postlingual deafness, it may be an oversimplification to focus on phonological structure alone. The lack of association between sentence recognition and phonological scores in NH listeners suggests less dependence on phonological structure by the NH participants. These NH listeners may

all have sufficient sensitivity to phonological structure to perform the speech recognition task used here, but they also may make use of other kinds of acoustic structure not available to CI users. Support for this idea comes from studies of NH listeners indicating that indexical structure (e.g., speaker age, gender, or dialect) can influence perceptual encoding and retention of spoken words (Mullennix, Pisoni, & Martin, 1989; Pisoni, 1997). For CI users, whose access to indexical structure may be restricted due to poor spectral resolution and poor transmission of pitch cues (Li & Fu, 2011; Luo, Fu, & Galvin, 2007), phonological skills may take on a more substantial role, resulting in stronger relationships of phonological capacities with speech recognition.

An apparent contradiction in relationships among phonological sensitivity, WM, and sentence recognition was found in this study. As predicted, better phonological sensitivity was correlated with better serial recall accuracy in NH listeners, suggesting that words were encoded and stored in WM using phonological codes for these participants. Sentence recognition scores in the NH group were not correlated with phonological sensitivity, possibly because their phonological skills were generally sufficient for performing a task of sentence recognition. In contrast, the opposite effects were seen for CI users: phonological sensitivity was not correlated with serial recall accuracy, suggesting that words were not being stored in WM using phonological codes. However, sentence recognition scores of CI users were related to phonological sensitivity, suggesting that access to phonological structure constrains their speech recognition abilities. This apparent contradiction deserves further exploration. Nonetheless, the finding remains that CI users experience deficits in their ability to access the phonological structure of spoken language when presented with a degraded signal through their CIs, and these deficits contribute to poor speech recognition skills.

Conclusion

The findings from this study support the idea that clinical outcomes in speech recognition for adult CI users are related to sensitivity of these users to phonological structure and only minimally to verbal WM skills, at least as gauged by the tasks used in this study. Phonological capacities may serve as potential targets for clinical interventions, such as by use of phonological training programs or through clinical rehabilitation approaches that focus on preserving or restoring phonological skills. These types of novel intervention strategies may be particularly helpful for older adult CI users.

Acknowledgments

References

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences, 11,* 458–465.

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.

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development, 77,* 1698–1716.

American Academy of Ophthalmology, Pediatric Ophthalmology Strabismus Panel. (2012). Preferred Practice Pattern® guidelines. Pediatric eye evaluations. San Francisco, CA: Author. Retrieved from http://www.aao.org/ppp

Andersson, U. (2001). *Cognitive deafness: The deterioration of phonological representations in adults with an acquired severe hearing loss and its implications for speech understanding* (Unpublished doctoral dissertation). Linköping University, Sweden.

Andersson, U., & Lyxell, B. (2007). Working memory deficit in children with mathematical difficulties: A general or specific deficit? *Journal of Experimental Child Psychology, 96,* 197–228.

Arehart, K. H., Souza, P., Baca, R., & Kates, J. (2013). Working memory, age and hearing loss: Susceptibility to hearing aid distortion. *Ear and Hearing, 34,* 251–260.

Baddeley, A. (1992). Working memory. *Science, 255,* 556–559.

Boothroyd, A. (2010). Adapting to changed hearing: The potential role of formal training. *Journal of the American Academy of Audiology, 21,* 601–611.

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 60,* P223–P233.

Brady, S., Shankweiler, D., & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology, 35,* 345–367.

Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41,* 977–990.

Cervera, T. C., Soler, M. J., Dasi, C., & Ruiz, J. C. (2009). Speech recognition and working memory capacity in young-elderly listeners: Effects of hearing sensitivity. *Canadian Journal of Experimental Psychology, 63,* 216–226.

Classon, E., Rudner, M., Johansson, M., & Rönnberg, J. (2013). Early ERP signature of hearing impairment in visual rhyme judgment. *Frontiers in Psychology, 4,* 241.

Classon, E., Rudner, M., & Rönnberg, J. (2013). Working memory compensates for hearing related phonological processing deficit. *Journal of Communication Disorders, 46,* 17–29.

Cleary, M., Pisoni, D. B., & Geers, A. E. (2001). Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Ear and Hearing, 22,* 395.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466.

Daneman, M., & Hannon, B. (2007). What do working memory span tasks like reading span really measure? In N. Osaka, R. H. Logie, & M. D'Esposito (Eds.), *The cognitive neuroscience of working memory* (pp. 21–42). Oxford Scholarship Online. https://doi.org/10.1093/acprof:oso/9780198570394.001.0001

Dawson, P. W., Busby, P. A., McKay, C. M., & Clark, G. M. (2002). Short-term auditory memory in children using cochlear implants and its relevance to receptive language. *Journal of Speech, Language, and Hearing Research, 45,* 789–801.

Dillon, M. T., Buss, E., Adunka, M. C., King, E. R., Pillsbury, H. C., III, Adunka, O. F., & Buchman, C. A. (2013). Long-term speech perception in elderly cochlear implant users. *Otolaryngology—Head & Neck Surgery, 139,* 279–283.

Dollaghan, C., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41,* 1136–1146.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Foo, C., Rudner, M., Rönnberg, 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.

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—1. Patterns of benefit. *International Journal of Audiology, 45,* 130–152.

Geers, A. E., Pisoni, D. B., & Brenner, C. (2013). Complex working memory span in cochlear implanted and normal hearing teenagers. *Otology & Neurotology, 34,* 396–401.

Goldstein, E. F. (1975). Selective phonemic and semantic coding in short-term recall. *Memory & Cognition, 3,* 619–626.

Hällgren, M., Larsby, B., Lyxell, B., & Arlinger, S. (2001). Cognitive effects in dichotic speech testing in elderly persons. *Ear and Hearing, 22,* 120–129.

Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., . . . Skinner, M. W. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear and Hearing, 34,* 342–360.

Holden, L. K., Reeder, R. M., Firszt, J. B., & Finley, C. C. (2011). Optimizing the perception of soft speech and speech in noise with the advanced bionics cochlear implant system. *International Journal of Audiology, 50,* 255–269.

Humes, L. E. (2005). Do 'auditory processing' tests measure auditory processing in the elderly? *Ear and Hearing, 26,* 109–119.

Ingvalson, E. M., & Wong, P. (2013). Training to improve language outcomes in cochlear implant recipients. *Frontiers in Psychology, 4,* 263.

Kenway, B., Tam, Y. C., Vanat, Z., Harris, F., Gray, R., Birchall, J., . . . Axon, P. (2015). Pitch discrimination: An independent factor in cochlear implant performance outcomes. *Otology & Neurotology, 36,* 1472–1479.

Kronenberger, W. G., Pisoni, D. B., Henning, S. C., Colson, B. G., & Hazzard, L. M. (2011). Working memory training for children with cochlear implants: A pilot study. *Journal of Speech, Language, and Hearing Research, 54,* 1182–1196.

Kumar, N., & Priyadarshi, B. (2014). Differential effect of aging on verbal and visuo-spatial working memory. *Aging and Disease, 4,* 170–178.

Li, T., & Fu, Q. J. (2011). Voice gender discrimination provides a measure of more than pitch-related perception in cochlear implant users. *International Journal of Audiology, 50,* 498–502.

Lin, F. R., Ferrucci, L., Metter, E. J., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology, 25,* 763–770.

Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., . . . Simonsick, E. M. (2013). Health ABC Study Group. Hearing loss and cognitive decline in older adults. *Journal of the American Medical Association, Internal Medicine, 173,* 293–299.

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 conditions: Effects on speech-in-noise performance in a two-channel hearing aid. *Journal of the American Academy of Audiology, 18,* 604–617.

Luo, X., Fu, Q. J., & Galvin, J. J. (2007). Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends in Amplification, 11,* 301–315.

Lyxell, B., Andersson, J., Andersson, U., Arlinger, S., Bredberg, G., & Harder, H. (1998). Phonological representation and speech understanding with cochlear implants in deafened adults. *Scandinavian Journal of Psychology, 39*(3), 175–179.

Lyxell, B., Andersson, U., Borg, E., & Ohlsson, I. S. (2003). Working-memory capacity and phonological processing in deafened adults and individuals with a severe hearing impairment. *International Journal of Audiology, 42,* S86–S89.

Mann, V. A., & Liberman, I. Y. (1984). Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities, 17,* 592–599.

Miller, G. A., Heise, G. A., & Lichten, W. (1951). The intelligibility of speech as a function of the context of the test materials. *Journal of Experimental Psychology, 41,* 329–335.

Moberly, A. C., Lowenstein, J. H., & Nittrouer, S. (2016). Word recognition variability with cochlear implants: The degradation of phonemic sensitivity. *Otology & Neurotology, 37,* 470–477.

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.

Nilsson, M., Soli, S. D., & Sullivan, J. A. (1994). Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. *Journal of the Acoustical Society of America, 95,* 1085–1099.

Nittrouer, S., & Burton, L. T. (2005). The role of early language experience in the development of speech perception and phonological processing abilities: Evidence from 5-year-olds with histories of otitis media with effusion and low socioeconomic status. *Journal of Communication Disorders, 38,* 29–63.

Nittrouer, S., Caldwell-Tarr, A., & Lowenstein, J. H. (2013). Working memory in children with cochlear implants: Problems are in storage, not processing. *International Journal of Pediatric Otorhinolaryngology, 77,* 1886–1898.

Nittrouer, S., & Lowenstein, J. H. (2010). Learning to perceptually organize speech signals in native fashion. *Journal of the Acoustical Society of America, 127,* 1624–1635.

Nittrouer, S., Lowenstein, J. H., Wucinich, T., & Moberly, A. C. (2016). Verbal working memory in older adults: The roles of phonological capacities and processing speed. *Journal of Speech, Language, and Hearing Research, 59,* 1520–1532.

Nittrouer, S., & Miller, M. E. (1999). The development of phonemic coding strategies for serial recall. *Applied Psycholinguistics, 20,* 563–588.

Oba, S. I., Galvin, J. J., III, & Fu, Q. J. (2013). Minimal effects of visual memory training on the auditory performance of adult cochlear implant users. *Journal of Rehabilitation Research and Development, 50,* 99–110.

Pichora-Fuller, M. K., & Souza, P. E. (2003). Effects of aging on auditory processing of speech. *International Journal of Audiology, 42,* 11–16.

Pickering, S. J. (2001). The development of visuo-spatial working memory. *Memory, 9,* 423–432.

Piquado, T., Benichov, J. I., Brownell, H., & Wingfield, A. (2012). The hidden effect of hearing acuity on speech recall, and compensatory effects of self-paced listening. *International Journal of Audiology, 51,* 576–583.

Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In K. Johnson & J. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9–32). San Diego, CA: Academic Press.

Pisoni, D. B. (2014). Rapid phonological coding and working memory dynamics in children with cochlear implants. In A. W. Farris-Trimble & J. A. Barlow (Eds.), *Language acquisition and language disorders: Vol. 56. Perspectives on phonological theory and development: In honor of Daniel A. Dinnsen* (pp. 91–112). Amsterdam, the Netherlands: John Benjamins.

Pisoni, D. B., & Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing, 24*(1, Suppl.), 106S.

Repovš, G., & Baddeley, A. (2006). The multi-component model of working memory: Explorations in experimental cognitive psychology. *Neuroscience, 139,* 5–21.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, 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), 1–17.

Rudner, M., Foo, C., Sundewall-Thorén, E., Lunner, T., & Rönnberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing-aid users. *International Journal of Audiology, 47,* 91–98.

Rudner, M., Rönnberg, J., & Lunner, T. (2011). Working memory supports listening in noise for persons with hearing impairment. *Journal of the American Academy of Audiology, 22,* 156–167.

Schvartz, K. C., Chatterjee, M., & Gordon-Salant, S. (2008). Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners. *Journal of the Acoustical Society of America, 124,* 3972–3988.

Spring, C., & Perry, L. (1983). Naming speed and serial recall in poor and adequate readers. *Contemporary Educational Psychology, 8,* 141–145.

Tao, D., Deng, R., Jiang, Y., Galvin, J. J., III, Fu, Q. J., & Chen, B. (2014). Contribution of auditory working memory to speech understanding in Mandarin-speaking cochlear implant users. *PLoS ONE, 9,* e99096.

Van Rooij, J. C. G. M., & Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. II. Multivariate analyses. *Journal of the Acoustical Society of America, 88,* 2611–2624.

Wechsler, D. (1991). *WISC-III: Wechsler intelligence scale for children* (Manual). San Antonio, TX: The Psychological Corporation.

Wilkinson, G. S., & Robertson, G. J. (2006). *WRAT 4: Wide range achievement test* (Professional manual). Lutz, FL: Psychological Assessment Resources.

Wilson, R. S., Leurgans, S. E., Boyle, P. A., Schneider, J. A., & Bennett, D. A. (2010). Neurodegenerative basis of age-related cognitive decline. *Neurology, 75,* 1070–1078.