Research Article

Memory Span for Spoken Digits in Adults With Cochlear Implants or Typical Hearing: Effects of Age and Identification Ability

Miranda Cleary,^a Tracy Wilkinson,^a Lauren Wilson,^a and Matthew J. Goupell^a

Purpose: Short-term and working memory vary across individuals and life span. Studies of how cochlear implant (CI) users remember spoken words often do not fully disentangle perceptual influences from memory assessment because stimulus identification is rarely checked; instead, correct perception is assumed by using simple or practiced stimuli. Here, we examine whether the accuracy and speed of auditory digit identification are associated with short-term and working memory in adults with CIs as a function of age. Do perceptual errors and longer processing times depress memory span scores for such listeners? An intensive digit span assessment using manual responses is compared with traditional digit span, which uses only a few speech-based trials. We also assess how vocoder-based degradation impacts identification and digit span in older versus younger listeners with normal hearing (NH).

Method: We measured forward and backward recall for spoken digits in 25 adults with CIs ($M = 57$ years; range = 20–82 years), 10 older adults with NH (> 55 years; $M = 64$ years), and 11 younger adults with NH (< 22 years). Listeners identified recorded digits one at a time by clicking numerals on a computer screen and then were cued to recall the list. In a comparison condition, listeners only identified each item. Listeners with NH also completed the tasks under four conditions of signal degradation implemented using a channel vocoder: four or eight channels crossed with 0- or 6-mm simulated tonotopic shift, administered in randomized blocks. A mixed design was used, with the effects of

degradation examined within subject, across different listener groups, and as a function of age.

Results: CI users were near perfect at identifying auditory digits ($> 99\%$ correct, on average), both in isolation and while experiencing a memory load. Perceptual accuracy, thus, did not affect digit span in these listeners. Identification times, however, were significantly slower under memory load. Age was only weakly related to span in CI listeners. Instead, individual differences in span in the CI group were better predicted by simple no-load identification times. Among listeners with NH, spans for vocoded digits that were not frequency shifted did not differ from spans for unprocessed stimuli. Frequency-shifted vocoded digits, however, yielded lower scores that could mostly, but not entirely, be accounted for by crediting for perceptual errors. Frequency-shifted vocoded digits were particularly challenging for older adult listeners with NH to identify. Traditional digit span correlated more strongly with our primary span measure for listeners with CIs than for listeners with NH.

Conclusions: Digit span can be studied independently of perception in many adult CI users. For both CI listeners and listeners with NH, stimulus degradation does not appear to significantly affect memory span, unless an effect on simple identification is also present. Auditory degradation that may slow, but which does not ultimately prevent identification, appears to have negligible impact on short-term and working memory spans.

The ability to temporarily store and recall novel sets of information varies across individuals, circumstance, and life span. Entering a new phone number or repeating back a set of previously unfamiliar names,

for example, are common situations relying on our ability to store and access material in the temporary workspace of short-term and working memory. This study of adult cochlear implant users and listeners with normal hearing (NH) addresses how degradations and distortions in the input stimulus affect initial identification (ID), storage, and recall of target information.

Auditory stimuli are more difficult to classify when the target signals are degraded or obscured by other signals. Listeners might be less accurate at identifying a stimulus as the intended target, slower to assign a category to

^aDepartment of Hearing & Speech Sciences, University of Maryland, College Park

Correspondence to Miranda Cleary: miranda.cleary@gmail.com

Editor-in-Chief: Frederick (Erick) Gallun

Editor: Lori J. Leibold

Received June 22, 2017

Revision received January 30, 2018

Accepted March 27, 2018

https://doi.org/10.1044/2018_JSLHR-H-17-0245

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

the stimulus, or less confident in their classification. Hearing loss poses various challenges for auditory processing. Even when hearing-assistive devices, such as cochlear implants (CIs) or hearing aids, are used to restore some of the auditory input, the sensory encodings are atypical because of the design of the electronic interface and the state of the compromised auditory processing system (Abbas & Brown, 2015; Firszt, Chambers, Kraus, & Reeder, 2002).

Speech communication research has long studied the impact of noise and signal distortion on perception, with additional efforts made to measure their influence on retention. In general, the adult listener with NH functions quite well despite degraded speech information, particularly when a scaffolding of learned “top-down” information exists to draw upon (Miller & Licklider, 1950; Moore, 1997; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995). For new information presented in an arbitrary order, however, such scaffolding is less helpful. Thus, use of relatively context-neutral test materials, such as spoken strings of digits, provides insight into the basic workings of memory that is mostly irrespective of language structure and world knowledge.

Overview of Short-Term Memory, Working Memory, and Processing Speed

Short-term memory, a temporary storage system that operates within the timescale of several seconds to several minutes, appears markedly limited in capacity and is reset or cleared with the passage of time, unless effort (such as rehearsal) or other circumstance precipitates recoding into longer, lasting storage (see Baddeley, Eysenck, & Anderson, 2015, for review). Digit span is a traditional measure of short-term memory function; successful immediate recall of longer strings of digits corresponds to a higher “span.” Working memory, in contrast, subsumes short-term memory and incorporates additional attentional mechanisms for maintaining and manipulating information in conscious awareness. Backward digit span, which requires participants to recall list items in reverse order, is usually argued to be a simple assessment of working memory.¹

Processing speed generally refers to how quickly and efficiently individuals can act upon postperceptual representations of information. How processing speed is related to “capacity” (i.e., the number of units that can be retained in short-term memory) is a fundamental issue, as is the question of whether processing speed is domain general (not specific to the modality of the original input) or sensory domain specific, for a given individual. In many discussions of intelligence, development, and aging, processing speed has been offered as the primary factor behind individual differences and age-related changes in cognition (see Park & Festini, 2017; Salthouse, 1996; Verhaeghen, 2014, for review).

¹The large “umbrella” construct of “executive function” incorporates working memory, along with mechanisms for planning and regulating behavior across tasks (e.g., AuBuchon, Pisoni, & Kronenberger, 2015b; Baddeley, Eysenck, & Anderson, 2015).

Errors of Perception Versus Errors of Memory

Studies of cognition in listeners with hearing loss have not always clearly distinguished between errors of perception and errors of memory. If an item is not perceived as intended, errors during “recall” might be due to initial misclassification of the input. Even when it is quite likely that ID was successful, the possibility of misperception remains, and if this influence is not explicitly quantified, its impact on purported “memory” scores is unclear (Shub, 2013). This study, therefore, incorporates procedures that permit a check of perception both prior to and during a memory task. Although this check may involve a longer delay before recall, when used in conjunction with other short-term/working memory measures that do not involve this additional requirement, a better estimate of perceptual influences on digit span may be reached for cases where it cannot be assumed that initial ID is near perfect (i.e., in people with relatively poor speech understanding).

Many findings regarding cognition in CI users come from the pediatric literature. This research suggests that a history of hearing loss, even with early use of a CI, generally results in poorer immediate recall of verbal (and verbally encodable) lists. Notably, even when children with CIs who demonstrate perfect accuracy on ID pretests for simple word stimulus sets are only considered, mean spans for lists of those same stimuli tend to be lower relative to those of age-matched peers with NH (Cleary, Pisoni, & Geers, 2001; Dawson, Busby, McKay, & Clark, 2002; Nitttrouer, Caldwell-Tarr, & Lowenstein, 2013). Although most of these studies have used well-familiarized or pretested stimuli, they typically have not measured correct ID accuracy or speed during actual presentation of the stimuli to be remembered.

Additional issues hinder interpretation of other data showing reduced spans in CI users. Traditional standardized digit span procedures, such as those included in the Wechsler Intelligence Scale for Children–Third Edition (Wechsler, 1991), typically use spoken responses. Such tests have sometimes been used with CI users (e.g., Pisoni & Cleary, 2003). Numerous studies have shown, however, that speaking rates are generally slower in speakers with hearing loss (e.g., Gold 1980; Nickerson, 1975; Uchanski & Geers, 2003). The time needed to initiate a spoken motor response also tends to be greater. Recently, AuBuchon, Pisoni, and Kronenberger (2015b) have reported that children and teens with CIs initiate naming of items more slowly and speak at a slower rate when repeating target sentences, on average, than typically hearing children. Whether these differences are related to motor output constraints or to aspects of long-term cortical representations of language, these facts complicate interpretation of spoken responses during a memory span task because memory decay becomes more likely as time elapses (Ebbinghaus, 1913).

Thus, inclusion of a condition requiring only a manual response method, as in this study, provides a useful

comparison point. Indeed, other research (e.g., AuBuchon, Pisoni, & Kronenberger, 2015a; Moberly, Harris, Boyce, & Nitttrouer, 2017; Nitttrouer et al., 2013) has used computerized versions of digit or simple word span tasks utilizing a manual response format. However, none of these studies have included a shadowing requirement for comparison, which is necessary to verify ID at time of list encoding. Shadowing refers to identifying a single auditory word as soon as it is recognized, either through verbal repetition or, as is implemented here, via a manual response choice among several labeled buttons.

Memory Span for Vocoded Words

Memory for vocoded speech has been studied in listeners with NH to better understand auditory memory in CI users. Vocoded speech is created by reducing the spectral resolution of the signal while preserving many of its temporal aspects and is the basis for the digital speech-processing strategies used with CIs (Loizou, 2006). Signals to be vocoded are first analyzed into a fixed number of frequency bands. Narrow bands of noise or sine tones at the center frequencies of each bandpass filter are then used as carrier signals to convey the relatively slow (e.g., < 400 Hz) amplitude modulations present within each frequency bandwidth “channel” (Loizou, 2006; Shannon et al., 1995). The summed output of the channels can be presented as is, or the individual channel information may be manipulated further to simulate the frequency-to-tonotopic-place mismatch (from here simply called *frequency shift*) that typically exists because CI electrode array insertions are relatively shallow (Landsberger, Svrakic, Roland, & Svirsky, 2015; Rosen, Faulkner, & Wilkinson, 1999).

An early study by Eisenberg, Shannon, Martinez, Wygonski, and Boothroyd (2000) obtained memory spans for eight-channel unshifted noise-vocoded digits from listeners with NH. They reported a small but statistically significant decrease in span for vocoded digits compared with unprocessed digits. ID scores for items in isolation were, however, not reported. Burkholder, Pisoni, and Svirsky (2005) subsequently collected spans from young adult listeners with NH using eight-channel noise-vocoded digits with a 6.5-mm upward frequency shift. After approximately 30 min of training with other vocoded speech samples and item-specific ID pretest and training trials, approximately half of their participants were less than perfect at identifying the vocoded digits 1 through 9. Spans for vocoded digits were found to be lower than those for unprocessed stimuli, but the authors also found that the reduction in span for individual subjects could be completely accounted for in terms of their errors during the ID pretest. Although no data from CI listeners were examined for comparison, these results suggest that if CI listeners make errors, even relatively infrequent ID errors, their span scores will be reduced due to perception, not memory limitations. Such a result could be further exacerbated by using tests involving relatively few trials, as is commonly the case in traditional digit span measures.

More recently, Nitttrouer and Lowenstein (2014) tested order recall for a fixed set of common monosyllabic vocoded words in 8-year-olds and adults with NH. In this case, the authors conducted both a pretest and posttest for simple ID of the eight-channel unshifted noise-vocoded stimuli and reported that no errors were made by any subject for unprocessed or vocoded stimuli during pretesting or posttesting. They, however, still found that recall performance for vocoded stimuli was poorer than for unprocessed speech. This suggests that degradation that does not measurably affect ID accuracy (as measured using small closed set of response alternatives) can nevertheless negatively influence order recall. In this study, as in the other vocoder studies reviewed above, perfectly accurate ID at time of list presentation was assumed, but not shown, given the task design.

Effects of Advancing Age on Memory Span and Reaction Time

Memory span and reaction time (RT) are both generally believed to be affected by advancing age. Aging after young adulthood is associated with a gradual increase in RT, although the root cause of the slowing is debated (Park & Festini, 2017; Salthouse, 1996; Verhaeghen, 2014). The data for memory span tasks are more ambiguous; although some research reports small to minimal decremental effects (Siegel, 1994; Wingfield, Stine, Lahar, & Aberdeen, 1988), several large-scale studies demonstrate clear aging effects on simple span (Gregoire & Van der Linden, 1997; Hester, Kinsella, & Ong, 2004; Myerson, Emery, White, & Hale, 2003).

In addition to debate regarding the precise effect size, there is controversy over whether the effect is equivalent for both forward and backward span. More complex working memory measures are often associated with larger age effects, perhaps because aging affects a variety of subcomponent processes, thereby magnifying the apparent effect of age. An often-stated, but possibly incorrect, related supposition is that backward digit span (which requires manipulation of elements) should decrease more steeply with advancing age than forward digit span. Data supporting steeper decreases for backward than for forward digit span (Bopp & Verhaeghen, 2005), as well as data showing equivalent slopes (Hester et al., 2004; Myerson et al., 2003), have been reported. It would be useful therefore, to examine not only whether a relationship exists between age and digit span in adult CI users but also whether this relationship changes when backward recall is required.

Current Study

This study has several goals. First, to measure the degree to which perception may influence span measures in CI users, we compare a shadowing version of a list recall task with results using a traditional span measure. To address the verbal production issue for this clinical population, we adopt a manual response method. We have also

chosen to test significantly more trials at each list length. If misperceptions are influencing the digit span scores of CI users as hypothesized, we should observe (a) ID errors during shadowing of list items and (b) higher spans when listeners are credited for “correctly” recalling their own misperceptions. If we observe no ID errors during shadowing, then we can assume that recall scores reflect memory function (or at least some postperceptual processes).

Second, few studies exist of RT in CI users. By collecting RTs both for successful simple ID and for successful ID under load, we will determine if memory load slows ID times. We will also assess if faster RTs predict longer spans within our listener groups. Additionally, we will characterize how RT changes for CI users as a function of advancing age, examined cross-sectionally, and determine if such a relationship must be accounted for when conducting other analyses of cognition in individuals with CIs.

Finally, our manipulation of signal degradation (unprocessed vs. eight or four channels of vocoding crossed with either 0 or 6 mm of simulated tonotopic shift) allows us to quantify signal quality and determine its relationship to both ID and recall in each NH listener group. We hypothesize that some mild levels of degradation will still allow perfect ID but will nevertheless hurt recall or slow ID, relative to no degradation. Alternatively, if recall is hurt only once ID becomes impaired, we may be able to determine if perceptual scores are fully predictive of recall scores.

Method

Participants

Twenty-five adult CI users, ranging in age from 20 to 82 years ($M = 57$ years; $SD = 19$ years), participated in the study. Participant characteristics for the CI group are shown in Table 1. All participants were regular users of their devices at time of testing and had used at least one of their CIs for at least a year.

Ten older adults with self-reported normal hearing, all over the age of 55 years (range = 58 to 71 years; $M = 64$ years; $SD = 4.5$ years), and 11 younger adults with self-reported normal hearing (range = 20 to 21 years) comprised two comparison groups. Hearing ability was further verified using a hearing threshold screening test (American Speech-Language-Hearing Association, 2005). Detection of pure tones was tested for each ear via headset (Maico, DD45) at 250, 500, 1000, 2000, and 4000 Hz within a sound-shielded testing booth (IAC Acoustics) using a portable pure-tone audiometer (Maico, MA41). The older normal-hearing (ONH) group was tested additionally at 8000 Hz. Responses ≤ 25 dB HL were accepted as normal thresholds, except for 8000 Hz, for which responses up to 35 dB HL were accepted if the opposite ear threshold was ≤ 25 dB HL. Thresholds for individual participants in the NH groups are shown in Table 2. Group mean audiometric thresholds are shown

in Figure 1. The older listeners with NH displayed relatively good hearing for their respective ages.

All participants were native speakers of American English recruited from the local community. Most had (or were pursuing) a college-level education. All were free of self-reported cognitive, neurological, or developmental impairments.

Participants completed informed consent procedures prior to participating and received monetary compensation for their time. All procedures were approved by the University of Maryland–College Park Institutional Review Board.

Stimuli

The digits 1 through 10 (excluding the bisyllabic number 7) spoken by a male talker were used in their unprocessed form and also used to create the vocoded speech stimuli. The original sound files were adapted from the Boston University matrix test (Kidd, Best, & Mason, 2008, Talker 07). Files used a 44.1-kHz sampling rate and 32-bit amplitude quantization. All files had the same root-mean-square energy and were zero padded at the end of the file to create 702-ms duration files.

Participants listened to the stimuli via a matched pair of loudspeakers (Tannoy Reveal) located 45° to the left and right of the seated participant, at a distance of approximately 18 in. All recorded stimuli were presented at approximately 65 dB-A within a double-walled, sound-attenuating booth (IAC Acoustics).

The sine-vocoded stimuli were created using MATLAB routines. The original stimuli were bandpass filtered (third-order Butterworth, using the “filtfilt” forward-and-reverse MATLAB filtering function) to minimize temporal distortion (e.g., Goupell et al., 2017) into four or eight contiguous channels. The corner frequencies of these filters spanned 200 to 5000 Hz and were logarithmically spaced (see Figure 1 for center frequencies). The Hilbert amplitude envelope of each channel was extracted and the resulting signals were then low-pass forward-and-reverse filtered using a second-order Butterworth filter with a cutoff frequency of 32 Hz. The low cutoff frequency was employed to keep the sidebands from the temporal envelope modulations contained within a single auditory filter (Souza & Rosen, 2009).

For synthesis of the channels, the sine tone carriers coincided with the center frequency of each bandpass filter (0-mm shift) or were converted to higher values that would be consistent with a shallow CI insertion depth (6-mm basalward shift; Rosen et al., 1999) using the conversion function detailed in Greenwood (1990). A correction for loudness after shift as per Faulkner, Rosen, and Stanton (2003) was applied. The channels were then summed to form the acoustic waveform. Root-mean-square energy levels were next adjusted to match the unprocessed stimulus set.

A custom-made MATLAB program was used to create, at time of testing, lists of 1 to 9 digits in length. Digit selection within a list was pseudorandomized with the

Table 1. Cochlear implant (CI) group: participant characteristics.

Listener	Code	Age (years)	CI processor right ear	CI processor left ear	Onset severe-to-profound hearing loss
1	CCE	20	Nucleus 6 w/N22	Freedom	Congenital
2	CBQ	22	AB-Harmony	—	Infancy
3	CAR	25	Freedom	—	Preschool
4	CAT	27	Freedom	Freedom	School age
5	CBP	36	—	Freedom	School age
6	CBU	41	—	AB-Naida	Toddlerhood
7	CAW	53	CP810 Freedom	CP810 Freedom	Progressive
8	CBA	54	Nucleus 5	Nucleus 5	Progressive
9	CAQ	56	Freedom	Freedom	Young adulthood
10	CBK	57	—	Nucleus 6	Teen
11	CBF	58	CP910-N6	Freedom	School age
12	CAY	59	Nucleus 5	Nucleus 5	Middle age
13	CBG	62	Nucleus 5	(HA)	Progressive
14	CCC	62	AB-Naida	(HA)	Middle age
15	CAJ ^a	65	Nucleus 5	—	Congenital
16	CCI	65	CP910	CP910	Young adulthood
17	CAK	69	Freedom	CP920	Middle age
18	CBY	69	AB-Naida	—	Middle age
19	CAM	71	Freedom	(HA)	Middle age
20	CAO	71	Nucleus 5	(HA)	School age
21	CAN	75	Freedom	—	Middle age
22	CCA	76	CP910	CP910	Middle age
23	CBC	78	CP810 Freedom	CP810 Freedom	Middle age
24	CBD	80	Freedom	—	Older adulthood
25	CBB	82	CP810 Freedom	—	Older adulthood

Note. Symbol “—” means no device in this ear. Code = arbitrarily assigned identifier—not CI user’s initials; (HA) = hearing aid typically used but replaced with earplug for study duration.

^aVision issue. Cursor enlarged and made higher contrast for participant.

caveats that no digit repeat within a list and immediately consecutive numbers were disallowed. In order to collect 10 responses at each list length in each condition, two blocks each containing five trials per list length were obtained for each condition. Older participants with NH were tested only on list lengths of two, four, six, and eight. Ordering of the list lengths was randomized; each trial, therefore, contained a list of unpredictable length.

CI participants and younger participants with NH completed 45 lists (five trials $\times$ nine list lengths) per block of trials for each task crossed by stimulus type. Older participants with NH completed 20 lists (five trials $\times$ four list lengths) per block. Two blocks of trials were collected under each condition, for a total of 90 (or 40) lists per task crossed by stimulus type.

Procedure

Preparation

Prior to testing, CI users were asked to report any issues with their devices. Unilateral CI users who reported residual hearing or hearing aid use in the opposite ear were fitted with an earplug in that ear. Participants with acoustic hearing had their hearing thresholds measured.

Traditional Digit Spans

All participants completed the Wechsler Adult Intelligence Scale–Third Edition (WAIS-III) Digit Span subtest

(Wechsler, 1997). We collected traditional spans to provide continuity with previous studies and to compare results obtained using the two methods. As already reviewed, WAIS Digit Span is a brief, standardized measure of short-term and working memory, similar to the Wechsler Intelligence Scale for Children Digit Span subtest (Wechsler, 1991) used in some past research. On each trial, participants were read a list of digits by an experimenter, at a rate of approximately 1 digit per second and, then, immediately repeated the list back aloud. Two lists at each length were tested, starting at list length 2, until both lists at a given length were incorrectly recalled. As per the standard scoring procedure, one point was awarded for each list correctly recalled in its entirety. Forward recall was tested first and, then, backward recall, requiring participants to reverse the list order when recalling the numbers. The maximum possible score for forward recall was 16 points, and 14 points for backward. Participants sat facing the examiner during presentation and recall.

Training With Vocoded Speech for NH Groups

To provide some familiarity with vocoded speech, listeners with NH received training with a subset of the vocoded stimuli. Participants were trained to identify the most severely distorted stimulus set (four channels with 6-mm shift) by listening to each stimulus and clicking on an on-screen button to show the matched digit. After each attempted ID, the correct button was illuminated, and the correctly matched unprocessed version of the stimulus

Table 2. Hearing thresholds in decibel hearing level.

ONH group														
Listener	Code	Age (years)	Right ear thresholds						Left ear thresholds					
			250	500	1000	2000	4000	8000	250	500	1000	2000	4000	8000
1	SQB	58	5	5	0	5	15	20	15	5	0	15	15	20
2	SQI	60	15	15	0	10	15	10	15	10	5	20	15	5
3	SQQ	60	15	15	10	10	20	20	10	10	5	5	10	25
4	SQC	62	10	10	15	25	20	25	5	10	15	25	20	35
5	SKV	63	25	20	20	20	15	15	15	20	20	20	5	15
6	SKP	63	5	10	15	15	5	15	5	10	10	15	15	20
7	SLD	65	15	15	5	5	-5	20	10	10	10	5	0	30
8	SRB	67	0	5	0	0	5	5	5	0	5	5	0	15
9	SJJ	71	10	10	10	5	-5	15	15	15	10	5	-5	15
10	SSA	71	5	0	15	15	15	15	5	10	10	15	10	15

YNH group														
Listener	Code	Age (years)	Right ear thresholds					Left ear thresholds						
			250	500	1000	2000	4000	250	500	1000	2000	4000		
*1	SPN	20	20	20	20	20	20	20	20	20	20	20		
*2	SPS	20	20	20	20	20	20	20	20	20	20	20		
3	SPF	20	10	5	5	10	0	0	0	5	10	-10		
4	SQU	20	5	10	10	5	10	5	15	10	0	0		
5	SRE	20	5	5	0	0	5	15	0	5	0	0		
6	SRL	20	0	-5	-5	-5	0	-5	-5	-5	-5	-5		
7	SRP	20	10	15	15	5	0	15	25	5	5	5		
8	SYE	20	10	5	0	0	10	5	5	0	5	0		
9	SPR	21	10	0	0	0	10	15	-5	5	5	10		
10	SRF	21	5	5	5	0	5	5	5	0	-5	5		
11	SRR	21	5	0	5	5	5	5	0	10	5	5		

Note. Participants marked with * did not receive threshold testing but, rather, passed a hearing screening conducted at 20 dB HL. These thresholds are conservatively estimated as 20 dB HL and were not included in averaged threshold data shown in Figure 1. Code = arbitrarily assigned identifier—not listener initials; ONH = older normal-hearing; YNH = younger normal-hearing.

Figure 1. Averaged audiometric hearing thresholds for ONH (N = 10, mean age = 64 years) and YNH (N = 9, mean age = 20 years; see text) listener groups. Bandpass filter center frequencies for vocoded digits overlaid for comparison. ONH = older normal-hearing; YNH = younger normal-hearing.

was played, immediately followed by another playing of the processed token. Each of the nine processed stimuli was trained 25 times, for a total of 225 trials. This training took approximately 25 min.

Primary Tasks

The three main experimental tasks were Shadowing Alone, Forward Digit Span, and Backward Digit Span. The main tasks were administered in the listed order and then, again, in reverse order. In the Shadowing Alone task, listeners saw the numbers 1 through 12 arranged clockwise in a circle on a computer screen with the numbers 7, 11, and 12 deactivated. On each trial, a single auditory digit was presented, and participants reported what they heard, via mouse click.

CI listeners completed the Shadowing Alone task using only the unprocessed digits. Listeners with NH, after completing the Shadowing Alone task with the unprocessed digits, completed the same task using the vocoded speech tokens. Trials using each of the four vocoded speech conditions were blocked, and the ordering of these blocks was randomized across listeners with NH.

Listeners next completed the Forward Digit Span task. For this task, listeners again identified the digits one at a time by clicking on-screen and then were visually cued to recall the entire list at the end of the trial. Responses were self-paced, and the manually entered recalled list could be erased and reentered if desired. Participants were encouraged to enter in as much as they could recall even if they could not remember the whole list. No cues to list length were provided. No audio was played during the listener response.

CI listeners completed the Forward Digit Span task using only the unprocessed speech tokens. Listeners with NH completed the Forward Digit Span task first in the unprocessed speech condition and then in the vocoded speech conditions using the same randomized order as in the Shadowing Alone task.

All listeners next completed the Backward Digit Span task. Participants were instructed to identify the digits one at a time, as before, but to recall the list in reverse order when given the visual cue. Again, CI listeners completed the Backward Digit Span task using only the unprocessed stimuli. Listeners with NH completed the Backward Digit Span task first in the unprocessed speech condition and then in the processed speech conditions using the same randomized order as used for that participant in the other two tasks.

Following the completion of one block of Backward Digit Span trials, listeners took a break and then completed their second and final block of Backward Digit Span trials. The second blocks of the Forward Digit Span task and the Shadowing Alone task were then completed. The randomized order used for the vocoded speech was reversed for the second block of each task. The unprocessed condition was then tested last for each task.

One block of the Shadowing Alone task took approximately 7 min to complete and involved either 225 ID

responses for the CI and younger normal-hearing (YNH) groups or 100 ID responses for the ONH group. One block of either the Forward or Backward Digit Span task took approximately 13 min to complete and involved either 225 ID responses and 45 list recalls for the CI and YNH groups or 100 ID responses and 20 list recalls for the ONH group. CI listeners completed all three tasks under only the unprocessed condition for a total testing time with breaks of approximately 1.75 hr. Younger listeners with NH took approximately 8.5 hr to complete testing. Older listeners with NH, in pilot testing, tended to need more time for the tasks and longer breaks. Older listeners with NH were therefore tested on only a subtest of list lengths and took approximately 6 hr to complete all tasks.

Data Analysis

One CI participant (CAN) completed only a half-session of digit span trials because of time constraints and did not return to complete the study due to health issues. Another participant (CBD) completed only a half-session of digit span trials and not the traditional digit span task and, due to relocation, did not return to finish the study. One additional CI participant (CBB) completed only 87/90 test trials due to a computer issue. These three data-sets were retained in the analysis.

For analysis, the primary dependent measure was a digit span estimate reflecting the list length at which participants were able to correctly recall all items in the correct order 50% of the time. This measure was obtained by calculating the mean percent correct lists at each length and fitting the psychometric function with a sigmoid (psignifit; Wichmann & Hill, 2001) to find the 50% cross-over point.

We also calculated the percentage of digits correctly shadowed in the Shadowing Alone task and during the shadowing portions of the Digit Span Forward and Digit Span Backward tasks. The recall data for the digit spans were scored in multiple ways. Our primary focus was on comparing (a) recall scored relative to the target stimuli and (b) recall scored relative to what the listener reported hearing during the shadowing task. Crediting listeners for correctly recalling list items that they had misidentified during shadowing by using their own responses was expected to result in longer spans than would be obtained using traditional digit span procedures.

RTs were collected for all shadowing responses in order to assess the impact of having to retain items in memory while identifying stimuli. RT is reported here as the time elapsed from the onset of the speech file to each mouse-click response, in milliseconds. Responses could not be given during the stimulus presentation, ~0.7 s, thereby introducing a minimum or floor value for the RT measurement. Only RTs from correct responses were retained in the analysis. Mean RT as a function of age was a main measure of interest. Additionally, we planned to examine the relationship between mean RT and digit spans collected using the automated versus traditional methods.

Faster RTs were expected to be associated with longer spans.

Results

Training With Vocoded Speech

Figure 2 displays results from the training task administered to the NH groups using the highly distorted four-channel, 6-mm-shift digit-name stimuli. Percent correct for each group is shown as a function of training set, where a set consisted of five trials with each processed digit. There were listeners in both NH groups who achieved over 90% correct by the final training set; however, there were also listeners in both groups who achieved only about 30% correct. The overall score distribution for each group appeared unimodal in shape and relatively continuous, suggesting that identifying this form of distorted speech is not an “all or nothing” ability. Collapsing across training set, the ONH group mean was 42.3% correct ($SD = 19.5\%$; chance = 11.1% correct). The YNH group mean was markedly higher at 62.6% correct ($SD = 21.5\%$). The mean difference between groups was 20.3% correct (95% confidence interval [1.5%, 39.1%]). A mixed analysis of variance showed main effects of set, $F(4, 76) = 35.9, p < .001$, and age group, $YNH > ONH, F(1, 19) = 4.9, p = .038$. The Set $\times$ Age Group interaction, however, did not reach statistical reliability ($p = .14$). Rationalized-arcsine-unit transformed scores corrected for chance responding yielded equivalent results.²

Shadowing Alone Versus Shadowing Under Load

Accuracy

Digit ID means are shown in Table 3 for the Shadowing Alone and Shadowing During Memory Load tasks. All groups were essentially at ceiling on the Shadowing Alone task using unprocessed stimuli. The CI group averaged 99.8% correct. Only one listener with a CI scored < 99%, with a score of 96.4% (CBK). Zero errors were made by the ONH group, and only one error by the YNH group. Furthermore, although we hypothesized that shadowing might be less accurate if listeners were required to concurrently hold items in memory, unprocessed digits were just as accurately shadowed under memory load as without memory load (bottom two panels vs. top panel, Table 3).

Listeners with NH had similarly high performance when shadowing digits in the eight-channel, 0-mm-shift condition. The other vocoded conditions, however, yielded many ID errors. As shown in Table 3, the ONH group generally had more difficulty identifying the frequency-shifted digits, both with eight channels and four channels

Figure 2. Training progress over time (Set#) for ONH versus YNH groups. Training used the most highly distorted, four-channel, 6-mm-shift, digit-name stimuli. ONH = older normal-hearing; YNH = younger normal-hearing.

of information, than the YNH group. For example, in the four-channel, 6-mm-shift condition, mean scores were twice as high in the YNH group (~87% correct, across tasks) than in the ONH group (~44% correct, across tasks). Conditions in which the difference between groups was statistically significant are marked in Table 3. These results suggest that degradation greater than that imposed by the eight-channel, 0-mm shift might not be suitable for “equating” the perceptual accuracy of listeners with NH with that of CI users under these testing conditions.

RTs

Although the listener groups appeared very similar in their accuracy at identifying unprocessed digits, examination of RTs for correct responses showed marked differences, as shown in Table 4. Younger listeners with NH responded more quickly on average than CI listeners and older listeners with NH, in all tasks and conditions. Conditions that yielded statistically significant differences between groups are marked in Table 4. As the table also shows, RTs under memory load were, for all groups, longer by several hundred milliseconds, as compared to when no concurrent task was present (paired t -tests, two-tailed, CI: $t(24) = 6.7, p < .001$, ONH: $t(9) = 3.5, p < .01$, YNH: $t(10) = 4.9, p < .01$).

Examining the effect of signal degradation on the ONH listener group, we found that changes in number of channels had minimal effect on RTs but that the frequency shift elicited clearly longer RTs. For the YNH group, RTs lengthened moderately both as channels were reduced and as frequency shift was increased. Again, the difference between NH groups was clearest for the frequency-shifted stimuli.

To better understand the contribution of age to RTs, we examined the simple correlations between these variables in our groups. For the CI listeners, age correlated reliably with Shadowing Alone RTs for unprocessed digits ($r = +.70, p < .001$; see Figure 3, filled triangles). RTs slowed

²A mixed analysis of variance on rationalized-arcsine-unit transformed scores (Sherbecoe & Studebaker, 2004) corrected for chance showed main effects of set, $F(4, 76) = 32.1, p < .001$, and age group, $YNH > ONH, F(1, 19) = 5.2, p = .034$. The Set $\times$ Age Group interaction, however, did not reach statistical reliability ($p = .15$).

Table 3. Average digit identification scores (% correct).

Task type	Group			Mean difference
Scores for shadowing alone				
Listening condition	CI	ONH	YNH	
Unprocessed	99.76	100.00	99.98	
8-channel, 0-mm shift		99.65	99.82	
4-channel, 0-mm shift		94.62	94.90	
8-channel, 6-mm shift		43.55	90.03	$p < .001$
4-channel, 6-mm shift		46.45	84.24	$p = .001$
Scores for shadowing during memory load: forward recall				
Listening condition	CI	ONH	YNH	
Unprocessed	99.41	99.76	99.75	
8-channel, 0-mm shift		99.60	99.68	
4-channel, 0-mm shift		95.65	98.71	$p = .05$
8-channel, 6-mm shift		41.10	95.21	$p < .001$
4-channel, 6-mm shift		43.55	85.10	$p < .001$
Scores for shadowing during memory load: backward recall				
Listening condition	CI	ONH	YNH	
Unprocessed	99.59	99.95	99.82	
8-channel, 0-mm shift		99.95	99.84	
4-channel, 0-mm shift		96.15	99.72	$p = .03$
8-channel, 6-mm shift		40.60	96.64	$p < .001$
4-channel, 6-mm shift		43.35	90.67	$p < .001$

Note. p values reported for unpaired t -tests, two-tailed, $df = 19$. Use of rationalized-arcsine-unit transformed scores (Sherbecoe & Studebaker, 2004) yielded equivalent statistical outcomes. CI = cochlear implant; ONH = older normal-hearing; YNH = younger normal-hearing.

cross-sectionally by approximately 100 ms for every decade of life (see Figure 3, trendline). Among the listeners with NH (see Figure 3, open circles), age also correlated with RT in the unprocessed condition (simple bivariate Pearson $r = +.82$, $p < .001$; point biserial $r = +.78$ with age treated

categorically as younger vs. older; $r(8) = +.63$ for ONH group alone), with RTs increasing by approximately 74 ms for each decade of life. Similar-sized correlations with age were obtained for shadowing RTs in each of the degraded speech conditions.

Table 4. Average digit identification RTs (ms).

Task type	Group			Mean difference		
RTs for shadowing alone						
Listening condition	CI	ONH	YNH	CI vs. ONH	CI vs. YNH	ONH vs. YNH
Unprocessed	1519	1371	1063	<i>ns</i>	<i>p</i> < .001	<i>p</i> < .001
8-channel, 0-mm shift		1433	1075			<i>p</i> < .001
4-channel, 0-mm shift		1546	1114			<i>p</i> < .001
8-channel, 6-mm shift		2364	1165			<i>p</i> < .001
4-channel, 6-mm shift		2249	1258			<i>p</i> < .001
RTs for shadowing during memory load: forward recall						
Listening condition	CI	ONH	YNH	CI vs. ONH	CI vs. YNH	ONH vs. YNH
Unprocessed	1887	1869	1505	<i>ns</i>	<i>p</i> = .002	<i>p</i> = .09
8-channel, 0-mm shift		1921	1497			<i>ns</i>
4-channel, 0-mm shift		2059	1503			<i>p</i> = .03
8-channel, 6-mm shift		3040	1583			<i>p</i> < .001
4-channel, 6-mm shift		2745	1652			<i>p</i> < .001
RTs for shadowing during memory load: backward recall						
Listening condition	CI	ONH	YNH	CI vs. ONH	CI vs. YNH	ONH vs. YNH
Unprocessed	1891	2026	1464	<i>ns</i>	<i>p</i> < .001	<i>p</i> = .02
8-channel, 0-mm shift		2050	1488			<i>p</i> = .02
4-channel, 0-mm shift		2170	1429			<i>p</i> < .01
8-channel, 6-mm shift		2920	1451			<i>p</i> < .001
4-channel, 6-mm shift		2920	1532			<i>p</i> < .001

Note. p values reported for unpaired t -tests, two-tailed. α adjustment, seven comparisons per task: $.05/7 = .007$. RT = reaction time; CI = cochlear implant; ONH = older normal-hearing; YNH = younger normal-hearing; *ns* = not significant.

Figure 3. Average RT for Shadowing Alone of unprocessed stimuli, plotted as a function of age. Filled black triangles indicate CI users, and unfilled circles indicate listeners with NH. RT = reaction time; CI = cochlear implant; NH = normal hearing.

Unlike RTs for Shadowing Alone, RTs for Shadowing Under Load did not associate clearly with age. For the CI group, no significant correlations with age were found ($p > .25$ for all comparisons). For the listeners with NH, RTs under load were not strongly correlated with age when the digits were unprocessed or mildly degraded ($r = +.40$ to $+.50$, p values ranging from .011 to .067, across conditions). However, when the frequency shift was involved, RTs under load showed strong correlations with age ($r = +.70$ to $.80$, $p < .01$ for all comparisons, two-tailed). This was found both for the entire NH group and for the ONH group considered on its own.

Digit Spans

Data Fitting

In general, the curve-fitting routine worked well describing the data. In only 11 of 512 attempted fits (23 participants $\times$ two conditions $\times$ two scoring criteria + 21 participants $\times$ 10 conditions $\times$ two scoring criteria) did the algorithm fail to fit the curve, all due to very low scores (2.1% of cases). All 11 of these cases occurred for the ONH group in the frequency-shifted conditions for the strict scoring method, which compared the participant's response to the presented target list. For these cases, the data were individually examined, and a score of 1.0 was assigned when no lists at any tested length (two, four, six, and eight) were correctly recalled (all cases). This value was chosen because the lowest well-fit score was 1.13 digits. Three additional fits could not be completed on the full dataset for that condition, but a fit was possible using half of the data (one block). In each case, this score fell below 2.0.

Impact of Scoring Method on Span

We first assessed the impact of crediting participants for "recalling" their shadowed response as compared with scoring strictly on the basis of whether the participant's recalled response matched the intended target list.

For CI group with their highly accurate shadowing, there was no effect of the two different scoring methods. Only one listener with a CI (CBK) made enough shadowing errors to have this negatively impact a span score, decreasing an estimated span by 0.8 digits. For four CI users, scoring versus the target increased scores by ~ 0.1 digit; these were cases in which participants appeared to recognize that he or she had misidentified the item and then self-corrected his or her response when prompted to recall.

Scoring method, likewise, had no effect on span in the YNH and ONH groups when unprocessed speech was used in the span task. However, for vocoded speech, crediting for "correctly recalled" mis-shadows had an observable effect, particularly for the ONH group. These results are shown by contrasting the top and bottom panels of Figure 4.

For the ONH group, crediting for correctly recalled misidentifications increased mean forward and backward spans by two to three digits for the frequency-shifted stimuli, regardless of number of channels, but had little effect in the unshifted conditions. For the YNH group, forward and backward spans were increased by about 1 digit in the most degraded four-channel, 6-mm-shift condition and by less than half a digit in the eight-channel, 6-mm-shift condition. There was no effect of scoring method for the YNH group in any of the unshifted conditions.

As can be seen in Figure 4, crediting for recall of shadowed responses did not entirely equate the spans obtained in different conditions. The ONH group's mean span, for example, was about 0.8 digits lower, on average, in the degraded listening conditions, compared with the unprocessed condition, even when the more lenient scoring method was used. Other types of errors, beyond simple item misclassifications, occurred in the degraded conditions.

Statistical Analysis of Primary Variables

For our statistical analyses, we focused on spans scored relative to shadowed responses, thereby minimizing perceptual contributions to performance. Examining spans for unprocessed digits, we found statistically significant main effects of Group, CI vs. ONH vs. YNH; $F(2, 42) = 9.5$, $p < .001$, and Recall Direction, Forward vs. Backward; $F(1, 42) = 12.5$, $p = .001$. Backward spans tended to be shorter than forward spans, and post hoc Bonferroni comparisons indicated that younger listeners with NH displayed higher spans using unprocessed digits than CI (adjusted $p < .001$) listeners and older listeners with NH (adjusted $p = .03$). The difference between ONH and CI groups did not reach significance. No significant interactions were observed.³

Examination of vocoded digit spans in the NH groups revealed main effects of Age Group, $F(1, 19) = 14.9$, $p = .001$, and Frequency Shift, $F(1, 19) = 37.2$, $p < .001$, but no main effects for Channels or Recall Direction (p values $> .05$). No reliable multiway interactions were observed, although the Channels $\times$ Age Group interaction and the Shift $\times$ Channels $\times$ Age Group interaction were marginal at

³Statistical analyses using the strict match-to-target scoring method for span showed predictably identical results for unprocessed digits.

Figure 4. Digit spans under each level of speech degradation. CI mean data shown in solid bars, ONH in hatched bars, and YNH in open bars. Top panel contains mean spans scored with respect to the intended target without accounting for systematic perceptual misidentifications. Bottom panel shows spans when listeners were credited for consistently “recalling” digits they had misidentified during shadowing, when recalling the entire list. CI = cochlear implant; ONH = older normal-hearing; YNH = younger normal-hearing; unproc. = unprocessed; ch = channel.

$p = .06$ and $p = .08$, respectively. Thus, younger listeners with NH displayed longer spans than older listeners with NH, and spans were reduced by frequency shift, even when credit was given for correctly recalled misidentifications. When vocoded digits were used, backward spans were not reliably shorter than forward spans, unlike the unprocessed digits case. Furthermore, eight-channel stimuli did not reliably elicit longer spans than four-channel stimuli.⁴

⁴For vocoded digits, statistical analyses using the strict match-to-target scoring method for span showed an additional main effect of Channels, $F(1, 19) = 10.0$, $p < .01$, and two-way Age Group $\times$ Shift interaction ($p < .001$). The three-way Shift $\times$ Channels $\times$ Age Group interaction also reached significance ($p < .01$).

The absence of interactions between Age Group and Recall Direction for either unprocessed or vocoded digit span reflects that, contrary to expectations, we did not find a larger difference between forward and backward recall for ONH, as compared with younger listeners with NH.

Relationships With Age

Figure 5 shows individual memory span scores, plotted as a function of age in years, for unprocessed digits. Forward spans appear in the upper panel and backward spans in the lower panel. CI data are shown in black triangles and NH data in unfilled circles (both ONH and YNH), all plotted as a function of age. Although there is clearly some tendency for spans to drop as a function of

Figure 5. Relationship between listener age and digit span for unprocessed digits. CI = cochlear implant; NH = normal hearing.

age, there is a wide scatter, both for NH and CI groups. There is a much weaker linear relationship between age and span than was observed between age and no-load RT (see Figure 3). For the CI group (filled triangles), older age was only weakly associated with shorter forward, $r(23) = -.26$, *ns*, and shorter backward, $r(22) = -.33$, *ns*, spans. For the listeners with NH (open circles, $n = 21$), the association was more robust; older age was strongly associated with shorter forward spans ($r = -.75$, $p < .001$; point biserial $r = -.73$) and moderately associated with shorter backward spans ($r = -.55$, $p = .01$; point biserial $r = -.51$). Thus, here too, we did not clearly observe any greater sensitivity of backward digit span to advancing age, as compared with the forward digit span measure (see also Füllgrabe, Moore, & Stone, 2015; Hester et al., 2004; Myerson et al., 2003).

Comparisons With Traditional WAIS Digit Span

To determine the degree to which the current memory span measure indexes the same abilities as traditional WAIS Digit Span, we compared performance across the two task formats. Traditional WAIS scoring for the CI group yielded means of 10.1 points for forward and 7.3 points for backward. For the ONH group, the means were 11.3 points for forward and 7.9 points for backward WAIS digit spans. For the YNH group, the means were 11.6 points for forward

and 8.8 points for backward WAIS digit spans. Thus, the overall patterning across listener groups was the same using both digit span measures (i.e., $CI < ONH < YNH$; see also Figure 4). Performance of the ONH group resembled that of the YNH group more closely, however, when the traditional all-verbal digit span measure was used, as compared with the more intensive measure utilizing manual responses.

Correlations across tasks were also examined for each group, despite the limited sample size for the individual NH groups. All correlations were calculated using performance in the unprocessed speech condition. For the CI group, the simple correlation between forward digit span and the WAIS forward points score was $r(22) = +.66$ ($p < .01$, $n = 24$) and between backward digit span and WAIS backward points was $r(21) = +.52$ ($p = .01$, $n = 23$). The ONH group displayed a weak correlation, $r(8) = +.37$, *ns*, between the two forward span measures and a moderately strong, nearly significant correlation, $r(8) = +.63$, $p = .053$, between the two backward span measures. In the YNH group, we saw a negligible correlation between forward span measures, $r(9) = +.15$, *ns*, and a small correlation between backward span measures, $r(9) = +.41$, *ns*.

These results suggest that the two memory span measures tap some of the same skills for CI users, but for listeners with NH, there is more overlap between tasks in the backward recall case. This might be related to aspects of visual-spatial processing demands of the present span task with its manual response format and the manipulation of order relations demanded by reverse recall tasks. Listeners with CIs might rely more heavily on additional visual inputs in forward recall tasks than listeners with NH.

We also found that traditional digit span was less sensitive to participant age than our manual response-based digit span measure: There was no correlation between age and either WAIS digit span score for the CI group, and among listeners with NH, only small correlations were observed between age and forward WAIS span ($r = -.14$, *ns*) and backward WAIS span ($r = -.25$, *ns*). Nevertheless, given the expectation of interrelationships and the presence of covariance in some conditions, we calculated the partial correlations between tasks with age statistically adjusted for. As shown in Table 5, some of these partial correlations are slightly larger than the simple bivariate correlations, but the overall pattern and interpretation remain the same.

Relationship Between Span and RT

We additionally examined the degree to which RTs predicted span. CI listeners with longer simple no-load ID times displayed significantly shorter forward, $r(23) = -.57$, $p < .01$, and backward, $r(22) = -.55$, $p < .01$, spans. These relationships did not diminish in size when statistically controlling for age. ID times under load did not show this same relationship with span, with only weak nonsignificant associations present ($p > .05$). Among listeners with NH, the relationship was similar in strength; no-load RT was correlated, $r(19) = -.61$, $p < .01$, with forward span,

Table 5. Partial correlations (controlling for age).

Manual response task	Traditional forward digit span			Traditional backward digit span		
	CI	ONH	YNH	CI	ONH	YNH
Forward Digit Span	.71***	.29	.16	.64**	.67*	.48
Backward Digit Span	.68***	-.12	.05	.55**	.58	.66*

Note. CI = cochlear implant; ONH = older normal-hearing; YNH = younger normal-hearing.

* $p < .05$. ** $p < .01$. *** $p < .001$.

and $r(19) = -.56$, $p < .01$, with backward span. Again, longer RTs were associated with shorter spans. For the NH group, however, unlike the CI listeners, these relationships disappeared once age was statistically controlled for. RTs under load did not associate reliably with spans for any listener group.

Discussion

Previous studies of digit span in CI users have not fully disentangled perceptual influences from memory assessment because stimulus ID has rarely been checked during target list presentation. In this study, to measure the contributions of perception, we utilized a digit span task that incorporated shadowing during list presentation. We found that both CI listeners and listeners with NH were highly accurate at closed-set ID of auditory-only digits when heard in isolation and also when under memory load. Span differences between listeners, therefore, could be more validly attributed to factors other than perceptual ability.

The present digit span measure used a manual response method that avoided demands on speech production and examiner perception and correlated only modestly with traditional all-verbal digit span. There was more overlap between the manual and traditional digit span measures for the CI listeners, and more overlap between the present task and traditional backward, than with traditional forward digit span tasks, for listeners with NH (see Table 5). Spatially related processing of order relations in reverse digit span may contribute to these results. Further comparisons using other response methods would be useful.

WAIS digit span can be administered quickly but does not generate a fine-grained measure of memory span. Furthermore, WAIS digit span scores typically take on only a few different point values within a limited range. Nonstandard WAIS scoring using the longest list length correct (e.g., as used in Nittrouer, Lowenstein, Wucinich, & Moberly, 2016) can further limit the distribution of spans. For tracking changes in memory function over time or for examining relatively small effects on memory, more intensive measures, such as used here, involving multiple test trials, are useful. Testing significantly more trials at each list length resulted in a more continuously distributed span measure across the observed range of scores.

Using this more intensive method, we found age to be associated cross-sectionally with a modest decline in

span scores across the NH group and with a small decline in scores in the CI group (see Figure 5). No clear effects of age on traditional digit spans were observed for either group. This might be because the coarse-grained nature of WAIS digit span scoring obscures such relatively small effects unless large numbers of listeners are available across the age range (as in, e.g., Gregoire & Van der Linden, 1997; Myerson et al., 2003).

The association between age and simple ID RT was, in contrast, strongly evident in all data (see Figure 3). For CI listeners, these provide some baseline data from a large response set, against which the ID times for other less basic speech tokens can be compared. The finding of slightly larger relative increases in RT per decade of life in the CI group as compared with the NH groups is suggestive of a possible difference in developmental trajectory when hearing is compromised. Well-established gradual changes in RT associated with aging after young adulthood could perhaps be steeper in such a clinical population. Future comparisons between CI listeners and age-matched listeners with NH would be helpful to better assess this possibility. In particular, more data from middle-aged adults with NH would be required.

With the current dataset, we cannot rule out possible contributing factors, such as general familiarity with mouse click navigation of a computer screen or manual motor dexterity, which also correlate with advancing age. However, when taken together with measures of span that do not incorporate such features, a more complete picture of short-term and working memory function in CI and ONH populations can be obtained.

Simple ID RTs were reliably correlated with digit span within the CI group, even when shared variance with chronological age was statistically controlled for, which suggests that a speed-related factor (e.g., see Salthouse, 1996) during perception may influence digit span. This sounds plausible (faster ID times might leave more time for mental repetition and rehearsal); however, contrary to expectation, this relationship was only visible in the NH data when age was left in as a covariate and disappeared when age was removed.

A further complication is that, although we found that memory load did indeed slow RTs in all listeners, load did not reduce the probability of successful ID for any listener group. Furthermore, listeners who identified digits more quickly under load did not display longer spans than

listeners who had slower RTs under load. We may need to examine how listeners choose to spend their time while a new digit is presented with an existing list held in memory. Perhaps some listeners, but not others, for example, mentally repeat the existing list and then press a button to identify the new item when shadowing under load. Imposing further task constraints might clarify this.

Auditory stimulus degradation via vocoding does not appear to significantly affect memory spans obtained using such vocoded digits, unless an effect on simple ID is also observed (see Table 3 and Figure 4). This finding runs counter to some older but often referenced findings suggesting that mild auditory degradation can impair short-term memory (e.g., Rabbitt, 1966, 1968, 1991). The present results also do not fit well with more recent investigations using slightly different methodologies and content. For example, it has been reported that memory for arbitrary word pairings and memory for whether a word appeared in a studied auditory word list are impaired by having low levels of noise (that do not prevent correct ID) present during the initial exposure and intervening before testing (Kjellberg, Ljung, & Hallman, 2008; Murphy, Craik, Li, & Schneider, 2000). Research with older listeners with mild hearing loss has also suggested that recall, even for very small and recently presented three-item word sets, is impaired beyond what would be expected from simple ID accuracy for the items (McCoy et al., 2005).

Our results are, however, in agreement with other reports examining learning and memory under conditions of low-to-moderate non-speech-based environmental background noise (see discussion in Klatte, Bergström, & Lachmann, 2013). If a stimulus can be identified, it can potentially be remembered, even if the initial presentation was not optimal. Indeed, most speech signals are heard under nonoptimal conditions, and our brains are equipped to deal with this. Mild degradation of speech seems not to consume cognitive resources needed to ensure later recall.

Given the small response set used in this study, we were somewhat surprised by how disruptive the 6-mm frequency shift was for many older listeners with NH and some of the younger listeners with NH, even after training (see Figure 2 and Table 3). Variability in understanding frequency-shifted vocoded words has been shown in young adults and in children 8–10 years of age (Waked, Dougherty, & Goupell, 2017). In the frequency-shifted stimuli, some of the high-frequency cues to consonant identity, though audible, may have been difficult to perceive due to the slightly worse thresholds at the higher frequencies that are typical past middle age (see Table 2). However, this may not be the main cause of the aging effect. First, the older listeners with NH had relatively good thresholds at 8 kHz, and the speech information should have been mostly audible. Second, it appears that little frequency information for the shifted vocoded stimuli is presented at these high frequencies (see Figure 1). Alternatively, aspects of how the training and feedback were structured might have been less than optimal for older listeners. For example, longer pauses between training feedback and the next training trial (of a different

item) might allow for better learning in older listeners. Future studies will need to disentangle these possible interpretations.

In conclusion, memory function can be studied relatively independent of perception in adult CI users, if care is taken to measure perceptual contributions. Our training and shadowing results suggest that there are age-related differences in how vocoded speech is perceived, particularly vocoded speech that incorporates a frequency shift. When this distortion is extreme enough to disrupt ID, spans measured using such stimuli will reflect perceptual contributions and will underestimate the listener's actual memory capabilities. Finally, aging is associated not only with changes at the auditory periphery but also with how quickly common speech stimuli can be classified and, to some degree, how accurately arbitrary orderings of increasing length can be retained in short-term and working memory. These relationships seem to hold true not only for listeners with NH but also, to an extent, for CI listeners of quite heterogeneous backgrounds.

Acknowledgments

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Number R01AG051603, awarded to Matthew J. Goupell. This study was also partially supported by a seed grant from the College of Behavioral and Social Sciences and a Tier 1 seed grant from the Office of the Vice President for Research at the University of Maryland–College Park awarded to Matthew J. Goupell. The authors gratefully acknowledge the help and advice of Maureen Shader and Arifi Waked in the data collection. The authors thank the University of Maryland Medical School (Nicole Nguyen, David Eisenman, and Ronna Hertzano) and Walter Reed National Military Medical Center (Gerald Schuchman) for their help with recruiting listeners. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Abbas, P. J., & Brown, C. J. (2015). Assessment of responses to cochlear implant stimulation at different levels of the auditory pathway. *Hearing Research*, 322, 67–76.
- American Speech-Language-Hearing Association. (2005). *Guidelines for manual pure-tone threshold audiometry*. Retrieved from <http://www.asha.org/policy>
- AuBuchon, A. M., Pisoni, D. B., & Kronenberger, W. G. (2015a). Short-term and working memory impairments in early-implanted, long-term cochlear implant users are independent of audibility and speech production. *Ear and Hearing*, 36, 733–737.
- AuBuchon, A. M., Pisoni, D. B., & Kronenberger, W. G. (2015b). Verbal processing speed and executive functioning in long-term cochlear implant users. *Journal of Speech, Language, and Hearing Research*, 58, 151–162.
- Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2015). *Memory* (2nd ed.). London, United Kingdom: Psychology Press.
- Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: A meta-analysis. *Journal of Gerontology: Psychological Sciences*, 60, P223–P233.

- Burkholder, R., Pisoni, D. B., & Svirsky, M. (2005). Effects of a cochlear implant simulation on immediate memory in normal-hearing adults. *International Journal of Audiology*, 44, 551–558.
- 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, 395–411.
- 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.
- Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. (H. A. Ruger & C. E. Bussenius, Trans.). New York, NY: Teachers College, Columbia University.
- Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., & Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America*, 107, 2704–2710.
- Faulkner, A., Rosen, S., & Stanton, D. (2003). Simulations of tonotopically mapped speech processors for cochlear implant electrodes varying in insertion depth. *The Journal of the Acoustical Society of America*, 113, 1073–1080.
- Firszt, J. B., Chambers, R. D., Kraus, N., & Reeder, R. M. (2002). Neurophysiology of cochlear implant users I: Effects of stimulus current level and electrode site on the electrical ABR, MLR, and N1-P2 response. *Ear and Hearing*, 23, 502–515.
- Füllgrabe, C., Moore, B. C., & Stone, M. A. (2015). Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition. *Frontiers in Aging Neuroscience*, 6, 1–25.
- Gold, T. (1980). Speech production in hearing impaired children. *Journal of Communication Disorders*, 13, 397–418.
- Goupell, M. J., Gaskins, C. R., Shader, M. J., Walter, E. P., Anderson, S., & Gordon-Salant, S. (2017). Age-related differences in the processing of temporal envelope and spectral cues in a speech segment. *Ear and Hearing*, 38, e335–e342.
- Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *The Journal of the Acoustical Society of America*, 87, 2592–2605.
- Gregoire, J., & Van der Linden, M. (1997). Effect of age on forward and backward digit spans. *Aging, Neuropsychology, and Cognition*, 4, 140–149.
- Hester, R. L., Kinsella, G. J., & Ong, B. (2004). Effect of age on forward and backward span tasks. *Journal of the International Neuropsychological Society*, 10, 475–481.
- Kidd, G., Best, V., & Mason, C. R. (2008). Listening to every other word: Examining the strength of linkage variables in forming streams of speech. *The Journal of the Acoustical Society of America*, 124, 3793–3802.
- Kjellberg, A., Ljung, R., & Hallman, D. (2008). Recall of words heard in noise. *Applied Cognitive Psychology*, 22, 1088–1098.
- Klatte, M., Bergström, K., & Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Frontiers in Psychology*, 4, 578. <https://doi.org/10.3389/fpsyg.2013.00578>
- Landsberger, D. M., Svrakic, M., Roland, J. T., & Svirsky, M. (2015). The relationship between insertion angles, default frequency allocations, and spiral ganglion place pitch in cochlear implants. *Ear and Hearing*, 36, e207–e213.
- Loizou, P. (2006). Speech processing in vocoder-centric cochlear implants. *Advances in Otorhinolaryngology*, 64, 109–143.
- McCoy, S. L., Tun, P. A., Cox, L. C., Colangelo, M., Stewart, R. A., & Wingfield, A. (2005). Hearing loss and perceptual effort: Downstream effects on older adults' memory for speech. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 58(1), 22–33.
- Miller, G. A., & Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *The Journal of the Acoustical Society of America*, 22, 167–173.
- Moberly, A. C., Harris, M. S., Boyce, L., & Nitttrouer, S. (2017). Speech recognition in adults with cochlear implants: The effects of working memory, phonological sensitivity, and aging. *Journal of Speech, Language, and Hearing Research*, 60, 1046–1061.
- Moore, B. C. J. (1997). *An introduction to the psychology of hearing* (4th ed.). San Diego, CA: Academic Press.
- Murphy, D. R., Craik, F. I. M., Li, K. Z. H., & Schneider, B. A. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychology and Aging*, 15, 323–334.
- Myerson, J., Emery, L., White, D., & Hale, S. (2003). Effects of age, domain, and processing demands on memory span: Evidence for differential decline. *Aging Neuropsychology and Cognition*, 10, 20–27.
- Nickerson, R. S. (1975). Characteristics of the speech of deaf persons. *Volta Review*, 77, 342–362.
- Nitttrouer, 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.
- Nitttrouer, S., & Lowenstein, J. H. (2014). Separating the effects of acoustic and phonetic factors in linguistic processing with impoverished signals by adults and children. *Applied Psycholinguistics*, 35, 333–370.
- Nitttrouer, 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.
- Park, D. C., & Festini, S. B. (2017). Theories of memory and aging: A look at the past and a glimpse of the future. *Journal of Gerontology: Psychological Sciences*, 72, 82–90.
- 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(Suppl.), 106S–120S.
- Rabbitt, P. (1966). Recognition memory for words correctly heard in noise. *Psychonomic Science*, 6, 383–384.
- Rabbitt, P. (1968). Channel-capacity, intelligibility and immediate memory. *Quarterly Journal of Experimental Psychology*, 20, 241–248.
- Rabbitt, P. (1991). Mild hearing loss can cause apparent memory failures which increase with age and reduce with IQ. *Acta Otolaryngology Supplement*, 476, 167–176.
- Rosen, S., Faulkner, A., & Wilkinson, L. (1999). Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants. *The Journal of the Acoustical Society of America*, 106, 3629–3636.
- Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103, 403–428.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303–304.
- Sherbecoe, R. L., & Studebaker, G. A. (2004). Supplementary formulas and tables for calculating and interconverting speech recognition scores in transformed arcsine units. *International Journal of Audiology*, 43, 442–448.
- Shub, D. E. (2013). Compensating for identification errors in memory recall tasks. *JASA Express Letters*, EL293–EL299.

-
- Siegel, L.** (1994). Working memory and reading: A life-span perspective. *International Journal of Behavioral Development*, 17, 109–124.
- Souza, P., & Rosen, S.** (2009). Effects of envelope bandwidth on the intelligibility of sine- and noise-vocoded speech. *The Journal of the Acoustical Society of America*, 126, 792–805.
- Uchanski, R. M., & Geers, A. E.** (2003). Acoustic characteristics of the speech of young cochlear implant users: A comparison with normal-hearing age-mates. *Ear and Hearing*, 24(Suppl.), 90S–105S.
- Verhaeghen, P.** (2014). *The elements of cognitive aging*. Oxford, United Kingdom: Oxford University Press.
- Waked, A., Dougherty, S., & Goupell, M. J.** (2017). Vocoded speech understanding with simulated shallow insertion depths in adults and children. *The Journal of the Acoustical Society of America*, 141, EL45–EL50.
- Wechsler, D.** (1991). *The Wechsler Intelligence Scale for Children—Third Edition (WISC-III)*. San Antonio, TX: The Psychological Corporation.
- Wechsler, D.** (1997). *The Wechsler Adult Intelligence Scale—Third Edition (WAIS-III)*. San Antonio, TX: The Psychological Corporation.
- Wichmann, F.A., & Hill, N. J.** (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Attention, Perception, & Psychophysics*, 63, 1293–1313.
- Wingfield, A., Stine, E. A., Lahar, C. J., & Aberdeen, J. S.** (1988). Does the capacity of working memory change with age? *Experimental Aging Research*, 14, 103–107.