Verbal Working Memory in Older Adults: The Roles of Phonological Capacities and Processing Speed

Susan Nittrouer, Joanna H. Lowenstein, Taylor Wucinich, and Aaron C. Moberly

Purpose: This study examined the potential roles of phonological sensitivity and processing speed in age-related declines of verbal working memory.

Method: Twenty younger and 25 older adults with age-normal hearing participated. Two measures of verbal working memory were collected: digit span and serial recall of words. Processing speed was indexed using response times during those tasks. Three other measures were also obtained, assessing phonological awareness, processing, and recoding.

Results: Forward and reverse digit spans were similar across groups. Accuracy on the serial recall task was poorer for older than for younger adults, and response times were slower. When response time served as a covariate, the age effect for accuracy was reduced. Phonological capacities were equivalent across age groups, so we were unable to account for differences across age groups in verbal working memory. Nonetheless, when outcomes for only older adults were considered, phonological awareness and processing speed explained significant proportions of variance in serial recall accuracy.

Conclusion: Slowing in processing abilities accounts for the primary trajectory of age-related declines in verbal working memory. However, individual differences in phonological capacities explain variability among individual older adults.

Advancing age is associated with pervasive declines in physical and cognitive functioning. These declines can significantly and deleteriously affect the quality of life for individuals as they age, because of the diminished capacity to participate in physical activities and to engage with other people, including family and friends. Consequently, it would be helpful if interventions could be developed that would slow these age-related declines. However, interventions will be useful only if they are designed to address the real source or sources of the declines. The study described here was designed to investigate declines in one cognitive function, verbal working memory, from two possible sources: declines in sensitivity to phonological structure and slower processing speeds.

There is strong and reliable evidence of age-related declines in cognitive functioning (e.g., Craik & Salthouse, 2000; Hedden & Gabrieli, 2004; Logie & Morris, 2015). Age-related declines have been reported for most cognitive functions, including language comprehension (Wingfield & Grossman, 2006) and general working memory (Myerson, Emery, White, & Hale, 2003). In fact, working memory capacity accounts for much of the variance in preserved language comprehension (Carpenter, Miyaki, & Just, 1994; DeCaro, Peelle, Grossman, & Wingfield, 2016; Norman, Kemper, Kynette, Cheung, & Anagnopoulos, 1991). However, several literature reviews (e.g., Bopp & Verhaeghen, 2005; Hale et al., 2011) have noted that most investigations demonstrating a relationship between language comprehension and working memory used experimental designs that treated working memory as a single construct, without distinguishing between visuospatial or verbal domains. The focus of the work reported here was with documented age-related declines in verbal working memory only.

The construct of verbal working memory refers to the mechanism by which sensory information for speech is stored for a short time and processed in the service of other mental operations (Baddeley, 1995). A frequently cited account of this mechanism describes it as a dual-component system in which verbal information is stored in a durable phonological code, recovered via the phonological loop, and subsequently processed by a central executive (Baddeley et al., 1992, 2007; Baddeley & Hitch, 1974). Studies of short-term
recall for lists of words that either rhyme or do not rhyme provide support for the proposition that phonological structure plays an essential role in storage of verbal material by demonstrating that lists of nonrhyming words are more accurately recalled than lists of rhyming words (Baddeley, 1966; Conrad & Hull, 1964; Nittouer & Miller, 1999; Salame & Baddeley, 1986). This finding necessitates that a new perspective be taken on the relationship between language processing and working memory. Rather than asking if working memory capacity affects language processing, the question can be posed as to whether an individual’s ability to process linguistic signals—in the service of recovering detailed phonological structure—influences verbal working memory capacity. In the study reported here, the hypothesis was tested that the ability to recover detailed phonological structure from the acoustic speech signal should affect the capacity to store sequences of verbal material.

Further motivation for posing this hypothesis comes from findings of children with dyslexia, a developmental disorder strongly associated with deficits in the nature of phonological representations (Liberman & Shankweiler, 1985; Snowling, 1998; Wagner & Torgesen, 1987). Children with dyslexia commonly perform more poorly than children without dyslexia on tasks of sentence comprehension but only for sentences with complex syntax (e.g., Smith, Mann, & Shankweiler, 1986; Stein, Cairns, & Zurif, 1984). For a long time, the accepted explanation for this deficit was that these children experience a developmental lag in acquiring complex syntax, compared with simple syntax, which they acquire in an age-appropriate manner. However, a few investigators questioned that explanation, proposing instead that the problem arose precisely from the phonological deficit and the load it imposed on verbal working memory (Brady, Shankweiler, & Mann, 1983; Crain, 1989; Mann & Liberman, 1984). If capacity for storage in a short-term memory buffer is affected by the phonological distinctiveness of the material being stored, the argument went, then individuals with degraded phonological representations should be at a disadvantage when it comes to working memory. Because syntactically complex sentences are long, children with these phonological deficits might consequently be impaired in their abilities to store and comprehend these types of sentences but show no difficulty with simple sentences.

To examine this hypothesis, children with dyslexia were tested on comprehension of sentences with relative clauses (e.g., Bar-Shalom, Crain, & Shankweiler, 1993; Smith, Macaruso, Shankweiler, & Crain, 1989). In some conditions, these clauses were structures that children typically acquire early, such as when the subject of the relative clause is also the subject of the main clause (The lady who held an umbrella kissed the man). In other conditions, these clauses were structures that children typically acquire late, such as when the object of the relative clause is the subject of the main clause (The lady whom the man kissed held an umbrella). Results showed that children with dyslexia had poorer comprehension overall than children without dyslexia, but the pattern of errors across sentence types was similar for both groups of children. This last finding supported the contention that children with dyslexia were just as knowledgeable about syntactic structure as children without dyslexia, but their limited working memory capacities constrained their abilities to retain sentences that are long. Nittouer (1999) was able to demonstrate specifically that comprehension difficulties were related to poorer verbal working memory, especially for words that should be phonologically distinct.

That pattern of results, in turn, raises the question of whether a similar relationship between the nature of phonological representations and verbal working memory might help explain observed decrements in language comprehension for older adults. On one hand, it would seem unlikely, given that phonological representations are cognitive structures developed over the course of childhood. Having become established as cognitive structures, it might be expected that they would remain intact throughout the life span. However, at least one group of aging adults appears to have phonological deficits: older adults with adventitious hearing loss. Several investigators specifically examined the nature of long-term phonological representations by asking older adults with adventitious, severe hearing loss to judge phonological similarity among written words (Andersson, 2002; Classon, Rudner, & Rönnberg, 2013; Lyxell, Andersson, Borg, & Ohlsson, 2003). These adults with hearing loss performed more poorly than adults with normal hearing matched by age, so it was concluded that phonological representations were not well preserved for these individuals with hearing loss. It may be that continuous, high-quality sensory input is needed to maintain cognitive structures, such as phonological categories. It may also be that phonological representations simply degrade with advancing age. Although that alternative hypothesis is hard to support with those studies because age-matched controls were tested, some aspects of those studies make it difficult to dismiss it completely. For example, the experiments were done in Swedish using written text, and Swedish orthography matches the phonology more directly than in many other languages (Seymour, Aro, & Erskine, 2003). Furthermore, those studies did not compare outcomes for older adults with normal hearing to younger adults with normal hearing. Thus, although the older adults with normal hearing had better sensitivity to phonological structure than the older adults with severe hearing loss, it may not have been identical to that of younger adults. The current study was designed to investigate the possibility that phonological representations may deteriorate for older adults generally and to explore whether the nature of phonological representations affects verbal working memory in older adults.

Hearing loss could be expected to degrade phonological representations, based on the fact that impaired auditory systems provide only impoverished spectral structure to the central nervous system, a problem that is compounded by the signal processing of cochlear implants. One study involving children who use cochlear implants provided support for this proposed effect of hearing loss on phonological structure and subsequently on verbal working
implants. Overall, children with hearing loss
tative was just as well developed for the children with cochlear
ingly, it was suggested that processing in the central execu-
ding. Finally, it was found that response times for the chil-
dren with cochlear implants matched those of the peers
with normal hearing. It was further observed that accuracy of
recall for the nonrhyming words correlated strongly with
scores on the task of phonological sensitivity (i.e., aware-
ness), and this relationship was stronger for the children
with cochlear implants than their peers with normal hear-
ing. Finally, it was found that response times for the chil-
dren with cochlear implants matched those of the peers
with normal hearing, despite their poorer recall. Accord-
ingly, it was suggested that processing in the central execu-
tive was just as well developed for the children with cochlear
implants. Overall, children with hearing loss—at least
those with losses severe enough to require cochlear implants—
appear to have degraded phonological representations,
arising from that hearing loss and subsequent implantation.
Those degraded representations, in turn, impede the quality
of the verbal material being stored in a working memory
buffer. However, processing speed was age appropriate for
these children.

That last finding for children with cochlear implants is
relevant because it contrasts with what might be expected
for older adults. Age-related declines in processing speed for
working memory tasks have reliably been observed (Salthouse,
1996; Salthouse & Babcock, 1991; Schneider, Daneman, &
Murphy, 2005; Stine, Wingfield, & Poon, 1986; Tun, 1998).
Slower responses allow more time for the memory trace to
decay. Thus, the current study examined the contributions
of both phenomena—phonological representations and
processing speed—to verbal working memory capacity in
older adults.

Current Study

Declines in verbal working memory related to aging
have been documented, but the bases of these declines are
not fully understood. Using a dual-component model of
working memory, this study examined two candidate pro-
cesses as potential sources of age-related decline in verbal
working memory: (a) deterioration in the quality of phono-
logical representations and (b) declines in the ability to
quickly and efficiently process sensory information. Under-
standing the factors that are responsible for declines in
working memory that accompany the aging process should
help us develop better strategies for forestalling these de-
clines, whether related to normal aging or clinical processes.

Two measures of verbal working memory were in-
cluded in this study: digit span and serial recall of lists of
established words. The first task was included because it is
used clinically, and the second was included because its
design could help identify whether any age-related differ-
ences were due to the nature of phonological representations
or to processing speed. Specifically, a picture-pointing task
was used in which the listener heard a string of words and
needed to point to pictures in the order recalled. 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 transparently
related to pictures. It was predicted that recall accuracy
would be poorer for the rhyming nouns than the non-
rhyming words, because rhyming words are phonologically
confusable. Response times were measured, and it was pre-
dicted that these times would be slower for the adjectives
than the nouns, because more processing is involved in
matching adjectives to pictures.

Three measures of phonological abilities were used:
(a) awareness of phonological structure in heard words,
(b) processing of phonological structure in heard words,
and (c) phonological recoding for words that are read. The
third task was included because it did not involve auditory
word recognition. Consequently, if older adults were found
to have poorer phonological abilities with the heard mate-
rials, but performed similarly with the read materials, it
could be concluded that their mental lexicons still consisted
of phonological representations and they still processed
linguistic materials with phonological codes but had diffi-
culty recovering a detailed phonological representation
from the acoustic speech signal.

Method

Participants

Forty-five adults participated in this experiment: 20
were between 18 and 32 years of age, and 25 were between
60 and 80 years of age. All participants in the younger
adult group were undergraduate students at the Ohio State
University. All participants in the older adult group had
taken at least some college courses, and most had 4-year
degrees. Mean ages and standard deviations at the time of
testing are shown in the first row of Table 1, and a t test re-
vealed that the difference in age was significant, t(43) = 36.25,
p < .001. All participants reported negative histories of
speech, language, and hearing problems. In particular,
none reported ever having any concerns regarding hearing,
and none used amplification.

Several kinds of additional data were collected and
used both as criteria for inclusion and as independent vari-
ables. Standardized measures of vocabulary and word read-
ing were obtained using the Expressive One-Word Picture
Vocabulary Test–Fourth Edition (Martin & Brownell,
2011) and the Word Reading subtest of the Wide Range
Achievement Test 4 (Wilkinson & Robertson, 2006). Par-
ticipants had to have standard scores better than 90 to
participate. Means and standard deviations for these values

are shown in the middle two rows of Table 1. When t tests were computed on these values, the difference in vocabulary scores was significant, \( t(43) = 2.17, p = .036 \), and the difference in word reading was close to significant, \( t(43) = 1.98, p = .055 \). These outcomes indicate that older adults had slightly better language abilities than younger adults. Those differences are noteworthy because it means that if older adults were found to have poorer verbal working memory than younger adults, that outcome could not be attributed to poorer language skills.

Auditory thresholds for the octave frequencies between 250 Hz and 8000 Hz were measured, and better-ear pure-tone averages (PTAs) were computed for the frequencies of 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz. To participate, thresholds at 2000 Hz needed to be better than 45-dB hearing level, which is roughly the mean threshold at that frequency for older men (Lee, Matthews, Dubno, & Mills, 2005); thresholds for older women are slightly better. Means and standard deviations for the better-ear, four-frequency PTAs are shown in the last row of Table 1. When t tests were computed on these values, a significant age effect was found, \( t(43) = 7.71, p < .001 \), indicating that younger adults had better hearing than older adults. Nonetheless, all older listeners had PTAs better than published means (Lee et al., 2005).

### Equipment

All testing was completed in a sound-treated booth. Pure-tone thresholds were measured with a Welch Allyn TN262 audiometer using TDH-39 headphones. Stimuli were stored on a computer and presented through a Creative Labs Soundblaster soundcard using a 44.1-kHz sampling rate and 16-bit digitization. That signal was then passed through a Samson headphone amplifier and AKG-K141 headphones. This system has a flat frequency response and low noise. Custom written software presented all stimuli. For the two memory tasks, responses were collected via keyboard. For this experiment, in this way, the rate and inflection of digit presentation was consistent across participants, and response time could be precisely measured. All digits were recorded by a male talker, and five samples of each were obtained. Samples matching best in duration, fundamental frequency, and intonation were used. They were matched in RMS amplitude. Stimuli were presented at a rate of one per second. Responses consisted of tapping on digits shown on a computer monitor, so verbal responses were not required. This response mode was implemented to eliminate the need to plan and orchestrate articulatory gestures, the speed of which could differ for younger and older adults.

To administer, the participant sat in front of the computer monitor with the hand he or she would use to respond on the table in front of the monitor. As a pretest, all digits were shown on the monitor, and the participant heard each one presented separately. The participant needed to tap each digit as it was heard. Participants were instructed not to say or even mouth the digits as they were heard. This pretest simply ensured that all participants were able to match the digits they heard. No participant had difficulty doing so. During testing, the digits were not shown on the monitor. The participant heard the sequence, and then the digits appeared at the top of the monitor. The participant needed to tap them in the order heard, as quickly as possible. As they did, each

---

### Table 1. Means and standard deviations for audiologic measures and language abilities of younger and older adults.

<table>
<thead>
<tr>
<th></th>
<th>Younger adults</th>
<th>Older adults</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M )</td>
<td>( SD )</td>
</tr>
<tr>
<td>Age (years)</td>
<td>20</td>
<td>3</td>
</tr>
<tr>
<td>Expressive Vocabulary</td>
<td>103</td>
<td>7</td>
</tr>
<tr>
<td>Word Reading</td>
<td>104</td>
<td>7</td>
</tr>
<tr>
<td>Better-ear PTA (dB)</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

Note. Better-ear pure-tone average (PTA) is the PTA for the frequencies 500 Hz, 1000 Hz, 2000 Hz, and 4000 Hz, presented as decibel hearing level. Expressive Vocabulary and Word Reading scores are standardized, with population means of 100 and standard deviations of 15.

---

Nittrouer et al.: Verbal Working Memory in Older Adults
digit moved to the vertical middle of the monitor, ordered left to right in the order tapped. Again, they were instructed not to say or mouth the digits. Each sequence length was presented twice with different digit orders, and the one just preceding the length at which both presentations resulted in incorrect responses was taken as the participant’s digit span. Both forward and reverse digit spans served as dependent measures, as well as response time per digit, for correct responses.

Serial Recall of Words

This was the other working memory task. Stimuli that have been used previously were used in this experiment, because it has been demonstrated that recall accuracy correlates with phonological sensitivity and response times differ, as would be expected based on listener age (Nittouer et al., 2013; Nittouer & Miller, 1999). Regarding the latter effect, young children with normal hearing were found to respond more slowly than young adults with normal hearing. These stimuli consisted of three sets of eight words each: nonrhyming nouns, rhyming nouns, and nonrhyming adjectives. The nonrhyming nouns were ball, coat, dog, ham, pack, rake, seed, and teen. The rhyming nouns were bat, cat, hat, mat, gnat, Pat (represented by a picture of a woman), rat, and var. The adjectives 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), neat (a neat desk), sad (a crying child), thin (a very thin man), and wet (a wet cat). All words were spoken by a male talker. Five tokens of each word were recorded, and individual tokens that matched best on duration, fundamental frequency, and intonation were used. It was not possible to equate words 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. Nonetheless, mean frequency of occurrence per one million words was obtained for each word using counts of Brysbaert and New (2009). Means were 51 for the nonrhyming nouns, 26 for the rhyming nouns, and 156 for the adjectives.

In addition to the eight-item lists, participants were also tested with six-item lists. This shorter list length was included because it was not known prior to testing whether or not eight items would result in performance by the older adults so poor as to be uninterpretable (i.e., near the floor). Items removed from the eight-item lists to create those shorter lists were those with the lowest frequency of occurrence. The words teen and seed were removed from the set of nonrhyming nouns, var and gnat were removed from the set of rhyming nouns, and neat and thin were removed from the set of adjectives.

Prior to testing with this task, the participant saw a series of blue squares (six or eight, depending on the list length) and was required to tap the squares in order from left to right as quickly as possible. Five trials were completed, and the average time across those trials was used to normalize response times to the test items. Next, testing was started. 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 prior to and subsequent to testing as a way of verifying that the participant recognized the words. During testing, words were presented at a rate of one per second, and 10 trials of each condition were included. Both response accuracy and response time, normalized for general response time, were used as dependent measures. Participants were instructed not to talk or move their mouths between hearing the words and responding.

Final Consonant Choice

This task assessed listeners’ sensitivity to or awareness of phonological structure. It consisted of 48 test trials and six practice trials and has been reported previously (e.g., Nittouer, Shune, & Lowenstein, 2011). All test items were recorded by a male talker. For each trial, the participant heard a target word and repeated it. They were provided with up to three opportunities to repeat the target correctly. However, more than one opportunity was rarely needed, indicating that all participants had good speech recognition. Three words were then presented, and the participant had to select which one ended in the same sound as the target word. All responses were audio-video recorded and scored later. Percentage correct responses served as the dependent measure.

Backwards Words

This task examined phonological processing: Listeners not only needed to recognize phonological structure but also needed to manipulate that structure. The task consisted of 48 test trials and six practice trials. These items are listed in the Appendix. Again, all word samples were recorded by a male talker. In this task, the participant heard a target word and repeated it. Next, the participant needed to say the word that resulted when the order of phonemes was reversed (e.g., nips becomes spin). All responses were audio-video recorded and scored later. Percentage correct responses served as the dependent measure.

Lexical Decision

This task examined phonological recoding. This was a reading task, consisting of 160 one- and two-syllable words and nonwords. These items were divided into five categories of 32 items each: (1) high-frequency, phonemically regular real words (e.g., dog, father, song); (2) low-frequency, phonemically less regular words (e.g., aisle, ewe, ostrich); (3) homophones of real words with a wide range of frequencies (e.g., fraun, oshin, toste); (4) nonwords that are somewhat phonemically regular (e.g., drint, kalife, snald); and (5) nonwords that are best described as letter strings (e.g., cifkr, pljuf, zchnm). Mean frequency of occurrence according to Brysbaert and New (2009) was 182 (range from 42 to 774) for the Category 1 words, 2 (range from
For the Category 2 words, and 217 (range from 1 to 5721) for the words corresponding to the homophones of Category 3.

During testing, the items appeared on the computer monitor in large letters, one item at a time. The participant’s task was to decide as quickly as possible if the item was a real word or not. If the decision was that it was a real word, the participant hit one key, marked in green. If the decision was that it was not a real word, the participant hit a different key, marked in red. Which sides of the keyboard the green and red keys were on was randomized across participants. Response time served as the dependent measure of primary interest, but percentage correct responses were also examined.

**General Procedures**

All procedures were approved by the Ohio State University Institutional Review Board. Participants were tested in two sessions. In the first session, hearing thresholds were measured first. Four tasks were then administered. First, either one of the serial recall tasks (eight or six items) or two of the phonological tasks (final consonant choice and backwards words) were administered. Whether listeners were given a serial recall task or the phonological tasks first was randomized across participants in each age group. Similarly, whether participants were administered the eight- or six-item serial recall tasks in this first session was also randomized. The second task administered in that first session was either a serial recall task or both phonological tasks, depending on what had been administered first. The last tasks administered in the first session were the word reading and vocabulary measures. In the second session, the lexical decision task was administered first, followed by the second serial recall task (eight or six items) and the digit span tasks. The order of administration of the serial recall and digit span tasks was randomized across participants in each age group in this second session.

**Results**

All data were screened for normal distributions and homogeneity of variances. An α of .05 was set, but precise p values are reported when p < .10. When p > .10, outcomes are reported as not significant.

**Working Memory**

**Digit Span**

Table 2 shows mean forward and reverse digit span and response times per digit for correct items. Only response time for reverse digits was significant between age groups, t(43) = 2.46, p = .018, with an effect size of .95 when given as a Cohen’s d. This outcome indicates that it took older adults longer to respond in this condition than younger adults. However, their digit spans were similar. The mental operation of reversing the order of digits heard requires greater processing than simply recalling the order originally heard. The significant age effect for response time suggests that processing demands were greater for the older adults (e.g., Cooper-Martin, 1994; Piolat, Olive, & Kellogg, 2005).

**Serial Recall of Words**

Figure 1 shows the mean percentage correct for words presented in each list position for each of the three conditions.
conditions. Outcomes for eight-word lists are shown on top, and outcomes for six-word lists are shown on the bottom. Table 3 shows outcomes of three-way, repeated-measures analyses of variance (ANOVAs) done on data from each list length, with condition and list position as the repeated measures and age group as the between-subjects measure. These results indicate that the list condition had a significant effect for both list lengths. Performance was best for the nonrhyming nouns and poorest for the rhyming nouns, with performance for adjectives intermediate. However, the magnitude of differences was greater for the eight-item lists. These outcomes are shown in Figure 2. List position also had a significant effect on participants’ abilities to recall accurately what word was in that position, for both list lengths, and this effect is apparent in Figure 1. The anticipated primacy and recency effects can be seen, although, again, the magnitude of the effect is somewhat greater for the eight-item lists rather than the six-item lists. These two within-subjects effects (i.e., condition and position) also demonstrated a significant two-way interaction for both list lengths. This interaction likely arose because participants did not demonstrate as strong recency effects for the rhyming nouns as for the other two conditions.

Turning attention to the between-subjects factor of age, this main effect was significant for both list lengths, as well: Younger adults were more accurate than older adults. These outcomes are shown in Figure 3 for each list length.

Regarding two-way interactions involving age, participants in both groups demonstrated similar condition effects, so the Condition × Age interaction was not significant. However, the Position × Age interaction was significant for the six-item lists and almost significant for the eight-item lists. These outcomes indicate that younger adults showed shallower recognition functions across list position, especially for six-item lists, for which they performed much closer to ceiling than the older adults.

Response times were analyzed next. To do this, calibration times were examined first. These were the times it took for participants to tap on the string of blue squares shown at the top of the computer monitor in order from left to right. Mean calibration times in seconds are shown in Table 4 and reveal that younger adults were quicker to respond than older adults. This impression was confirmed by significant outcomes of t tests performed on times for each list length: eight-item strings, t(43) = 3.41, p = .001, and six-item strings, t(43) = 3.04, p = .004. Because these calibration times showed significant differences between the two groups, response times during testing were corrected using calibration times. That is, response times were obtained by computing the mean time across the 10 trials in each condition and subtracting the calibration time for

Table 3. Outcomes of three-way, repeated-measures analyses of variance performed on percentage correct recognition scores for serial recall of words.

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>df</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Eight-Item lists</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>24.67</td>
<td>2.86</td>
<td>&lt;.001</td>
<td>.365</td>
</tr>
<tr>
<td>Position</td>
<td>165.97</td>
<td>7.301</td>
<td>&lt;.001</td>
<td>.794</td>
</tr>
<tr>
<td>Age group</td>
<td>11.13</td>
<td>1.43</td>
<td>&lt;.001</td>
<td>.206</td>
</tr>
<tr>
<td>Two-way interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition × Position</td>
<td>4.55</td>
<td>14.602</td>
<td>&lt;.001</td>
<td>.096</td>
</tr>
<tr>
<td>Condition × Age</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Position × Age</td>
<td>1.84</td>
<td>7.301</td>
<td>.079</td>
<td>.041</td>
</tr>
<tr>
<td>Three-way interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition × Position × Age</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td><strong>Six-Item lists</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Main effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>7.57</td>
<td>2.86</td>
<td>.001</td>
<td>.150</td>
</tr>
<tr>
<td>Position</td>
<td>72.99</td>
<td>5.215</td>
<td>&lt;.001</td>
<td>.629</td>
</tr>
<tr>
<td>Age group</td>
<td>16.84</td>
<td>1.43</td>
<td>&lt;.001</td>
<td>.281</td>
</tr>
<tr>
<td>Two-way interactions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition × Position</td>
<td>4.11</td>
<td>10.430</td>
<td>&lt;.001</td>
<td>.087</td>
</tr>
<tr>
<td>Condition × Age</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>Position × Age</td>
<td>5.77</td>
<td>5.215</td>
<td>&lt;.001</td>
<td>.118</td>
</tr>
<tr>
<td>Three-way interaction</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Condition × Position × Age</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
</tbody>
</table>

Note. ns = not significant.
that participant. Figure 4 shows mean response times for each group across conditions. The effect of condition is clearly evident in this figure: For both list lengths, it took longer for participants to respond to adjectives than to nonrhyming or rhyming nouns. Outcomes of a two-way, repeated-measures ANOVA supported this observation in that the main effect of condition was significant for list lengths of six items, \(F(2, 86) = 21.07, p < .001, \eta^2 = .329\), as well as of eight items, \(F(2, 86) = 17.99, p < .001, \eta^2 = .295\).

It is also clear from Figure 4 that even though response times were corrected for general differences between groups in speed of tapping the pictures (by subtracting calibration times), older adults were nonetheless slower than younger adults. Again, this was found to be the case for lists of six items, \(F(1, 43) = 20.70, p < .001, \eta^2 = .325\), as well as of eight items, \(F(1, 43) = 26.95, p < .001, \eta^2 = .385\). Finally, the Condition × Age interaction was significant for both list lengths: for six-item lists, \(F(2, 86) = 3.87, p = .024, \eta^2 = .083\), and for eight-item lists, \(F(2, 86) = 3.64, p = .030, \eta^2 = .078\).

Logarithmic transformations of these response times were computed and also used to examine these effects, as suggested by other investigators (e.g., Faust, Balota, Spieler, & Ferraro, 1999). Use of these transformed values in two-way ANOVAs did not change outcomes for the main effects of condition and age. However, the Condition × Age interaction ceased to be significant for either six- or eight-item lists. Nonetheless, clearly there remained evidence of generally slower processing for the older adults.

In light of the finding that older adults were slower to respond than younger adults, even when general age-related slowing of movement was taken into consideration, the main effect of age in response accuracy was reevaluated, using response time as a covariate. This analysis was performed to assess whether the poorer performance of older adults could be attributed to their slow responding, which would mean that the memory trace would degrade to a greater extent before they had a chance to respond. Specifically, the analyses reported in Table 3 were run again for six and eight items, using response times for nonrhyming nouns, rhyming nouns, and adjectives as covariates. For both list lengths, it was found that the magnitude of the main effect of age was reduced but not necessarily eliminated: for eight-item lists, \(F(1, 40) = 5.01, p = .031, \eta^2 = .111\); for six-item lists, \(F(1, 40) = 3.05, p = .089\).

### Relationship of Digit Span to Serial Recall of Words

The analyses described above indicated that digit span did not differentiate working memory performance for younger and older adults very well, but the serial recall of words did distinguish between these two groups. Thus, the two tasks differ in sensitivity to age effects. Nonetheless, it seemed valuable to examine the strength of the relationship between performance on the various measures of working memory to see where commonalities may rest. To accomplish this goal, Pearson product–moment correlation coefficients were computed between forward and reverse digit span, and each of the six measures obtained of serial recall of words: percentage correct recall and response time for each of the conditions of nonrhyming nouns, rhyming nouns, and adjectives. This was done for outcomes of both the six-item and eight-item lists. Of these 24 analyses, eight resulted in correlation coefficients with \(p < .10\). These are listed in Table 5 and reveal that only percentage correct scores for nonrhyming nouns and adjectives correlated with digit span scores. No significant or close-to-significant

---

**Table 4.** Mean time (seconds) for participants to tap on strings of objects on computer monitor.

<table>
<thead>
<tr>
<th></th>
<th>Younger adults</th>
<th>Older adults</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(M)</td>
<td>(SD)</td>
</tr>
<tr>
<td>Eight-item strings</td>
<td>1.8</td>
<td>0.3</td>
</tr>
<tr>
<td>Six-item strings</td>
<td>1.4</td>
<td>0.3</td>
</tr>
</tbody>
</table>

**Figure 4.** Response time (in seconds) for each age group for each condition. Eight-word lists are shown on the top panel, and six-word lists are shown on the bottom panel.
correlations were obtained for rhyming nouns or for measures of response time.

**Phonological Capacities**

The working memory measures revealed that the older adults were less skilled at serial recall of words, and that finding seemed to be partially explained by their slower processing. However, the phonological loop is a critical component of most models of working memory, and it could be that diminished sensitivity to or processing of phonological structure could explain the diminished working memory performance of older adults. To examine this possibility, three measures of phonological skills were obtained: awareness, processing, and recoding.

Figure 5 shows mean scores for both groups of participants, for the final consonant choice task (phonological awareness) and the backwards words task (phonological processing). Although it appears that older adults performed more poorly than younger adults, these differences were not statistically significant.

Turning to the lexical decision task (phonological recoding), Figure 6 shows response accuracy for participants in the two groups. In this case, accuracy was defined by the percentage of responses in which the word was judged correctly either as a real word (high-frequency [HF] and low-frequency [LF] words) or as a nonword (homophones [HP], phonologically regular nonwords [PR], and phonologically irregular nonwords [PI]). Although there is some variability in response accuracy, mean performance was better than 90% correct in all conditions. A two-way, repeated-measures ANOVA performed on these scores showed a significant effect of condition, $F(4, 172) = 15.58$, $p < .001$, $\eta^2 = .278$. However, neither the main effect of age nor the Condition $\times$ Age interaction was significant. Consequently, it can be concluded that younger and older

---

**Table 5.** Significant Pearson product–moment correlation coefficients between digit span and percentage correct scores for serial recall of words.

<table>
<thead>
<tr>
<th></th>
<th>Forward digit span</th>
<th>Reverse digit span</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$r$</td>
<td>$p$</td>
</tr>
<tr>
<td><strong>Eight-item serial recall</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonrhyming nouns</td>
<td>.358</td>
<td>.016</td>
</tr>
<tr>
<td>Adjectives</td>
<td>.341</td>
<td>.022</td>
</tr>
<tr>
<td><strong>Six-item serial recall</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonrhyming nouns</td>
<td>.344</td>
<td>.021</td>
</tr>
<tr>
<td>Adjectives</td>
<td>.264</td>
<td>.080</td>
</tr>
</tbody>
</table>

---

**Figure 5.** Percentage correct responses for each age group on the final consonant choice and backwards words tasks.

**Figure 6.** Percentage correct responses for each age group on the lexical decision task. HF = high-frequency real word; LF: low-frequency real word; HP: homophone of a real word; PR: phonologically regular nonword; PI: phonologically irregular nonword.
adults were similar in their abilities to decide the lexical status of the test items.

Figure 7 shows response times for each condition. Adults in both groups were fastest for the items that were most clearly real words or nonwords (HF and PI conditions) and slowest for the homophones. A two-way, repeated-measures ANOVA performed on these scores showed a significant effect of condition, $F(4, 172) = 36.46$, $p < .001$, $\eta^2 = .459$, indicating that response times were different across conditions. However, it appears from Figure 7 that response times may not necessarily have been slower for the homophones than for other nonwords that are somewhat phonologically regular. If so, that would mean that these adults were not specifically recoding the items into phonological structures that formed words but instead were just recognizing the phonological regularity of the items. As a check on this possibility, a paired-samples $t$ test was done on scores for the homophones and the phonologically regular nonwords. A significant outcome was obtained, $t(44) = 2.34$, $p = .024$, so it can be concluded that response times were slower for the homophones than for the phonologically regular nonwords.

Returning to outcomes for the two-way ANOVA, the main effect of age was also significant, $F(1, 43) = 10.59$, $p = .002$, $\eta^2 = .198$, but the Condition $\times$ Age interaction was not. That means that older adults were generally slower at responding than younger adults, but patterns of response times across conditions did not differ for the two groups. As was observed for phonological awareness and processing, younger and older adults seem to have engaged in phonological recoding to a similar extent. In general, older adults do not appear to have diminished phonological capacities.

**Explaining Variance**

Older adults were found to have slower response times than younger adults on the serial recall task, and when response times were used as covariates, it was discovered that the magnitude of the age effect diminished for the serial recall task at list lengths of both eight and six items. Despite that finding, however, it is still possible that the ability of older adults to bring their phonological capacities to bear on the storage of words in a working memory buffer could affect their serial recall performance. It may be that processing speed explains the group differences observed—the main course of age-related declines in verbal working memory—but phonological capacities might explain variance within the group of older adults, thus affecting the performance of individuals relative to the group. Consequently, analyses were undertaken on accuracy outcomes only of older adults to examine the contributions of response times, which represent processing speeds, and phonological capacities. To achieve this goal, six stepwise linear regressions were conducted: one for each of the three conditions at each list length. In each case, three measures served as predictor variables: response time for that condition and percentage correct scores for the final consonant choice and backwards words tasks. Scores for the phonological recoding task were not included because it was unclear which single metric would indicate this phenomenon.

For the eight-item lists, a unique solution was obtained for the nonrhyming nouns, in which it was found that performance on the backwards words task explained a significant amount of variance, standardized $\beta = .449$, $p = .024$. For the other two conditions with eight-item lists, none of the three predictor variables explained a significant portion of the variance. For the six-item lists, solutions were obtained for all three conditions. For the nonrhyming nouns, performance on the backwards words task was found to be the only score to explain a significant amount of variance, standardized $\beta = .410$, $p = .042$, as had been the case for eight-item lists. For rhyming nouns, the solution obtained included both performance on the final
consonant choice task, standardized β = .456, p = .013, and response time, standardized β = −.357, p = .046. For adjectives, the solution obtained included performance on the backwards words task, standardized β = .399, p = .027, and response time, standardized β = −.417, p = .021. These findings indicate that even though slowing of response time accounted for the poorer performance of older adults, relative to that of younger adults, on the serial recall task, phonological capacities explained much of the within-group variance.

Discussion

Advancing age is associated with declines in a number of cognitive functions, including verbal working memory. Deficits in verbal working memory for older adults likely have multiple sources, and this experiment examined two potential sources. First, it was considered possible that older adults would demonstrate degradation in the nature of their phonological representations or their abilities to access those representations. Such declines could explain poorer verbal working memory performance for older adults. Second, declines in verbal working memory for older adults might be explained by general slowing of cognitive functioning. To test these hypotheses, older and younger adults with age-appropriate hearing thresholds were tested using two measures of verbal working memory (digit span and serial recall of words), along with three measures of phonological skills (awareness, processing, and recoding).

In agreement with previous studies of verbal working memory, the older adults in this experiment demonstrated poorer verbal working memory performance and slower responses than younger adults. Importantly, these age-related effects were seen even though the older adults had what is considered normal hearing. In particular, the serial recall tasks were more sensitive than digit span to the age effect.

The first hypothesis tested was that older adults would demonstrate poorer phonological skills than younger adults, and these skills would explain poorer working memory performance. This hypothesis was only partially supported. Older adults did not demonstrate phonological deficits. They were able to access phonological structure upon listening to spoken language, and they recoded orthographic materials into phonological forms. However, those results do not mean that phonological skills did not contribute to working memory abilities. Within the group of older participants, phonological capacities explained a large portion of the variance in working memory performance.

Turning to the second hypothesis, that general slowing of cognitive functioning would explain declines in verbal working memory for older adults, ample support was found. Response times during the working memory tasks were longer for older adults than for younger adults, even after accounting for generalized slowing, as measured with calibration times. Moreover, response times on serial recall explained variance in working memory accuracy, at least for the shorter six-word tasks. Thus, slower responses appeared to negatively affect working memory, and this slowing largely explained poorer performance for older versus younger adults. These findings are consistent with the concept that slower processing would result in greater degradation of the memory trace prior to responding, resulting in poorer accuracy.

Taken together, the results of this study reveal the multifactorial effects of aging on cognitive processes, even for adults with age-normal hearing. Both hypotheses tested by this experiment were at least partially supported, suggesting that both phonological sensitivity and cognitive slowing contribute to verbal working memory ability in older adults, extending our previous understanding of this phenomenon. The current study extends findings of earlier studies involving adults with severe hearing loss (Andersson, 2002; Lyxell et al., 2003). In this study, the older adults with age-normal hearing retained the phonological skills of younger adults. Those phonological capacities play a central role in verbal working memory processing, as evidenced by significant within-group correlations for older adults between recall accuracy and phonological measures.

Conclusion

The results of the experiment presented here confirm the highly interactive nature of language processing and cognition, especially as they relate to normal aging. For one cognitive skill in particular, verbal working memory, aging appears to have a deleterious effect on performance primarily as a result of cognitive slowing. However, individual differences in phonological skills accounted for variability in verbal working memory among the older adults. Findings suggest that interventions to maintain phonological capacities could help ameliorate age-related declines in verbal working memory.

Acknowledgments

This work was funded by research Grant R01 DC-000633 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health. ResearchMatch, used to recruit some older participants, is supported by Grant U1LTR001070 from the National Center for Advancing Translational Sciences. The authors thank Lori Altmann for her suggestions on an earlier draft.

References


### Appendix

Backwards Words Task

#### Practice Examples

A. nip pin  
B. eat tea  
C. dice side

Discontinue testing after six consecutive errors.

<table>
<thead>
<tr>
<th>Test Trials</th>
<th>Answer</th>
<th>Test Trials</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. nap</td>
<td>pan</td>
<td>25. chip</td>
<td>pitch</td>
</tr>
<tr>
<td>2. sub</td>
<td>bus</td>
<td>26. make</td>
<td>came</td>
</tr>
<tr>
<td>3. tip</td>
<td>pit</td>
<td>27. pans</td>
<td>snap</td>
</tr>
<tr>
<td>4. cab</td>
<td>back</td>
<td>28. nuts</td>
<td>stun</td>
</tr>
<tr>
<td>5. gum</td>
<td>mug</td>
<td>29. claw</td>
<td>walk</td>
</tr>
<tr>
<td>6. pal</td>
<td>lap</td>
<td>30. loops</td>
<td>spool</td>
</tr>
<tr>
<td>7. meat</td>
<td>team</td>
<td>31. spin</td>
<td>nips</td>
</tr>
<tr>
<td>8. peek</td>
<td>keep</td>
<td>32. swap</td>
<td>paws</td>
</tr>
<tr>
<td>9. pool</td>
<td>loop</td>
<td>33. pets</td>
<td>step</td>
</tr>
<tr>
<td>10. leaf</td>
<td>feel</td>
<td>34. stack</td>
<td>cats</td>
</tr>
<tr>
<td>11. cat</td>
<td>tack</td>
<td>35. sleep</td>
<td>peels</td>
</tr>
<tr>
<td>12. right</td>
<td>tire</td>
<td>36. cans</td>
<td>snack</td>
</tr>
<tr>
<td>13. deer</td>
<td>reed</td>
<td>37. spill</td>
<td>lips</td>
</tr>
<tr>
<td>14. face</td>
<td>safe</td>
<td>38. slim</td>
<td>mills</td>
</tr>
<tr>
<td>15. time</td>
<td>might</td>
<td>39. stone</td>
<td>notes</td>
</tr>
<tr>
<td>16. pass</td>
<td>sap</td>
<td>40. nicks</td>
<td>skin</td>
</tr>
<tr>
<td>17. jab</td>
<td>badge</td>
<td>41. plug</td>
<td>gulp</td>
</tr>
<tr>
<td>18. gas</td>
<td>sag</td>
<td>42. spans</td>
<td>snaps</td>
</tr>
<tr>
<td>19. wall</td>
<td>law</td>
<td>43. stops</td>
<td>spots</td>
</tr>
<tr>
<td>20. peach</td>
<td>cheap</td>
<td>44. snoops</td>
<td>spoons</td>
</tr>
<tr>
<td>21. sob</td>
<td>boss</td>
<td>45. lambs</td>
<td>small/smell</td>
</tr>
<tr>
<td>22. shack</td>
<td>cash</td>
<td>46. spins</td>
<td>snips</td>
</tr>
<tr>
<td>23. name</td>
<td>mane</td>
<td>47. lever</td>
<td>reveal</td>
</tr>
<tr>
<td>24. mile</td>
<td>lime</td>
<td>48. turkeys</td>
<td>secret</td>
</tr>
</tbody>
</table>