J Am Acad Audiol 17:506–518 (2006)

Sequencing versus Nonsequencing Working Memory in Understanding of Rapid Speech by Older Listeners

Nancy Vaughan*
Daniel Storzbach*
Izumi Furukawa*

Abstract

The goal of this study was to identify specific neurocognitive deficits that are associated with older listeners' difficulty understanding rapid speech. Older listeners performed speech recognition tests comprised of time-compressed sentences with and without context, and on a neurocognitive battery aimed specifically at testing working memory, processing speed, and attention. A principle component analysis identified three main cognitive components as follows: a sequencing working memory (WM-S) component, a nonsequencing working memory (WM-NS) component, and a processing speed (PS) component. Each of the cognitive component scores was divided into high, mid, and low categories. Sentence performance of the cognitive subgroups was compared within each component. The results showed that, with hearing loss and age accounted for, the cognitive score groups differed similarly on the sentence condition scores also at 50 and 60% time compression, particularly on the subgroups of the WM-S component. The results suggest that deficits in a separate working memory function identified as sequencing were associated with differences in ability to understand time-compressed speech in this study.

Key Words: Aging, cognition, memory, speech recognition

Abbreviations: PS = processing speed; WM-NS = working memory—nonsequencing; WM-S = working memory—sequencing

Sumario

La meta de este estudio fue identificar deficiencias neurocognitivas específicas que se asocian con la dificultad de los sujetos mayores para endender el lenguaje rápido. Estos oyentes mayores realizaron pruebas de reconocimiento de lenguaje que incluían frases con compresión temporal, con y sin contexto, y una batería neurocognitiva orientada específicamente a evaluar memoria de trabajo, velocidad de procesamiento y atención. Un análisis de componente principal identificó tres componentes cognitivos: un componente de memoria secuencial de trabajo (WM-S), un componente de memoria no-secuencial de trabajo (WM-NS), y un componente de velocidad de procesamiento (PS). Cada uno de los puntajes de dichos componentes cognitivos fue dividido en

*National Center for Rehabilitative Auditory Research, Portland Veterans Affairs Medical Center, Portland, OR

Nancy Vaughan, PhD, National Center for Rehabilitative Auditory Research, Portland VA Medical Center, Portland, OR 97239; Phone: 503-220-8262, ext. 56030; Fax: 503-273-5021; E-mail: nancy.vaughan@med.va.gov

This work is supported by the Veterans Affairs Rehabilitation Research and Development (#C2631R).

categorías alta, media y baja. El desempe_o con las frases en los subgrupos cognitivos se comparó con cada componente. Los resultados mostraron que, considerando la hipoacusia y la edad, los grupos de puntaje cognitivo tuvieron diferencias similares en las frases a compresiones temporales de 50 y 60%, particularmente en los sub-grupos del componente WM-S. Los resultados de este estudio sugieren que las deficiencias en una función de memoria de trabajo separada, identificada como secuencial, se asocian con diferencias en la capacidad de entender lenguaje con compresión temporal.

Palabras Clave: Envejecimiento, cognición, memoria, reconocimiento del lenguaje

Abreviaturas: PS = velocidad de procesamiento; WM-NS = memoria de trabajo – sin secuenciación; WM-S = memoria de trabajo – secuenciación

It is generally agreed that the speech understanding difficulties of older adults are not consistently or adequately explained by the degree and configuration of hearing loss, and in fact, there are instances where these difficulties exist even in the absence of clinically significant hearing loss. This phenomenon has led to the hypothesis that age-related central auditory and cognitive processing deficits likely contribute to speech understanding difficulties experienced by older adults. The cognitive constraints most often identified as likely contributors are reduced working memory function and slowed speed of cognitive processing, functions that also decline with age (Strayer et al, 1987; Verhaeghen, 1999; Wingfield et al, 2000; Brebion, 2001). Of course, it is recognized that hearing loss affects the early stages of peripheral processing, thereby providing an impoverished signal for later stages of auditory and cognitive processing. If, in older adults, these later stages are compromised by age-related cognitive declines, speech-understanding deficits may result even if hearing loss is mild or even absent.

A characteristic feature of speech is the rapid rate at which it occurs. Speech is a complex acoustic signal in which the incoming acoustic and linguistic information proceeds on average at a rate between 200 and 270 words per minute (Wingfield et al, 1985), with considerable variability in speaking rates between and even within speakers (Miller et al, 1984). Listeners must therefore both store and process the incoming information, and these simultaneous functions must keep pace with the speech input in order to understand verbal communication. Lectures, movies, newscasts, and other types of ongoing speech place even more demands on cognitive processing because of the cumulative nature of the incoming information. Time compression is a method of speeding up speech while preserving both the pitch and prosody of the utterance. Time-compressed speech therefore shows promise as an effective tool to investigate the effects of declines, with age, in working memory and in processing speed. Older adults do not perform as well as younger adults on time-compressed speech recognition tests even when hearing loss is taken into account (Vaughan and Letowski, 1997; Gordon-Salant and Fitzgibbons, 2001). One study compared both uniform time compression and selective time compression of consonants within sentences and found that the poor performance of older listeners on rapid speech recognition tests was due to the reduction of consonant cues produced by the time-compression algorithm (Gordon-Salant and Fitzgibbons, 2001). However, to date, cognitive contributions to speech understanding deficits of older listeners have not been thoroughly investigated in association with time-compressed speech. Recent studies have shown the effects of cognitive limitations on benefits received from sophisticated signal processing strategies in modern hearing aids that inadvertently create temporal distortions in the aided speech signal.

The primary hypothesis of the current study was that declines in working memory contribute to speech understanding deficits in older listeners. Secondary hypotheses were that changes in attention and processing speed interact with working memory and contribute to speech understanding deficits in older adults. Time-compressed speech was used because it increases the working memory load and decreases time available for processing. It has been suggested that when speed of processing slows with aging, working memory deficits appear (Salthouse, 1992; Fry and Hale, 2000). In addition, age-related attentional deficits have been shown to be important in a variety of performance tasks in older adults (Morris et al, 1988; Tun et al, 1992; Giambra, 1993; Sullivan, 1993).

METHODS

Subjects

Qualified participants were 176 native speakers of American English, aged 50 through 75 years, with pure-tone thresholds of 40 dB HL or better between 250 Hz and 1000 Hz, 60 dB HL or better at 2000 and 3000 Hz, and not greater than 70 dB HL at 4000 Hz. Normal hearing was defined as less than or equal to 25 dB HL at octave and interoctave frequencies from 250 through 4000 Hz. Participants with conductive hearing loss (defined as air bone gaps of >15 dB and abnormal tympanograms; American Speech-Language-Hearing Association, 1990) were excluded. Other exclusion criteria were based on screening tests to rule out changes in cognitive function inconsistent with normal aging, that is, significantly below normal general intellectual ability (IQ), memory, and the presence of significant clinical depression. Reduced intellectual ability was defined by a Wechsler Abbreviated Scale of Intelligence (WASI) IQ score less than 70. Reduced memory was defined by a California Verbal Learning Test-II (CVLT-II) uncued delayed recall z-score of $\leq$-2. Significant depression was defined by Beck Depression Inventory-II (BDI-II) score of >19. A limited battery of auditory processing tests was administered to be included in future analysis, but none was used to exclude participants.

Assessment Instruments

Neurocognitive Tests

As described above, the BDI-II, CVLT-II, and the WASI were used to exclude participants who were inappropriate for the study goals. The primary neurocognitive battery listed below was composed of tests of working memory, speed of processing, and attention.

Visual Working Memory Tests

1. N-back (Cohen et al, 1994; Cohen et al, 1997). This is a computerized test that has been frequently used to assess working memory in functional Magnetic Resonance Imaging (fMRI) studies of brain function. The test requires the participant to identify, from a series of letters appearing one-by-one on the screen, each letter that has previously appeared either one or three letters previously (e.g., one-back, three-back). It increases in difficulty with increased number of intervening letters between repeated letters.
2. Self-Ordered Pointing Test (SOP) (Petrides and Milner, 1982; Petrides, 1994). This test requires the participant to point to one of 8 or 10 abstract designs that appear simultaneously in an array on the computer screen. The participant is presented with a series of trials with the designs rearranged at the beginning of each trial at different positions within the array. Trials are presented in three blocks each of either eight trials for the eight-design array or ten trials for the ten-design array. For each trial, the participant must point to one of the designs without pointing to any design already selected in a previous trial in the same block.

Verbal Working Memory Tests

1. Wechsler Adult Intelligence Scale-III (WAIS-III) Digit Span subtest (Wechsler, 1997). Both the Digit Span Forward and Backward were used.

Digit Span Forward requires the participant to repeat a series of numbers in exactly the same order. The Digit Span Backward increases demand on working memory by requiring the participant to repeat a series of numbers in reverse order.

2. WAIS-III Letter-Number Sequencing (LNS) subtest (Wechsler, 1997). A test of auditory working memory. A series of letters and numbers are presented in random order, and the participant must reorder the numbers and letters by first repeating the numbers in ascending order and then the letters in alphabetical order.

Speed of Processing Tests

1. WAIS III Digit Symbol Coding test (Wechsler, 1997). A timed test presented visually. The participant is presented with a table matching numbers 1–9 with simple symbols and a random array of numbers without symbols. The task is to copy the appropriate symbol for each number in a box below the number as quickly as possible without making mistakes. The task is terminated after 120 seconds. Number correct is scored.
2. Choice Reaction Time has been associated with processing speed in speech recognition tests (van Rooij and Plomp, 1990; Jerger et al, 1991). We used a simple task in each of the visual and auditory modalities. The participant must push either a right or a left button upon hearing or seeing the appropriate word ("RIGHT," "LEFT").

Tests of Attention

1. The Brief Test of Attention (BTA) (Schretlen et al, 1996a, 1996b). For each trial of this test, participants are presented with sequences of letters and numbers of varying length. During the first condition, participants are required to count how many numbers are presented in each trial while ignoring the letters. During the second condition, participants are required to count how many letters are presented in each trial while ignoring the numbers. No recall is involved.
2. The Conners' Continuous Performance Test (Conners, 2000) is a computerized test of attention that requires sustained attention over a long time period (14 minutes). Participants are presented a series of single letters on the computer screen at varying interstimulus intervals and are instructed to press the space bar for every letter except "x."

Speech Recognition Tests

Two types of speech materials were chosen for this study to provide two levels of difficulty based on differences in contextual cues. The IEEE sentences (Rothauser et al, 1969) were selected because these sentences are both semantically and syntactically correct, but predictability is not high. Anomalous sentences, syntactically correct but devoid of contextual cues, were computer generated from the vocabulary of the IEEE sentences to be of similar length and phonemic characteristics as the IEEE sentences and were designated "Anomalous" (ANOM) (e.g., "Hang the wheel in the stupid air"). The sentences were equated for difficulty under time-compression conditions in the NCRAR laboratory with young, normal-hearing listeners. Sentence scores were averaged over time-compression rates, and based on the distribution of scores, sentences with scores outside 1.5 standard deviations were considered too easy or too difficult and were excluded from the final corpus. All experimental speech recognition materials were time compressed by a custom software algorithm at four different rates (40%, 50%, 60%, and 65%). These time-compression rates were chosen based on sentence recognition performance-rate functions derived from previous research with normally hearing older listeners (Vaughan and Letowski, 1997). Some lower rates of time compression were included in this study in anticipation of the effects of hearing loss on sentence recognition.

PROCEDURES

Audiometric Procedures

All audiometric testing including speech recognition tests took place in a double-walled isolated chamber (Acoustic Systems model 19701A) adjoined to a single-walled control room. A Grason-Stadler (GSI 61) clinical audiometer calibrated to meet current standards (ANSI, 1996) was used for threshold testing and word-recognition testing through insert earphones (Etymotic ER3A). A Grason-Stadler (GSI 33, version 2) Middle Ear Analyzer was used to obtain tympanometric measurements. Speech-recognition materials were prerecorded on compact discs and delivered via the master hearing aid module of the Otowizard (MedRx Inc.) audiometer monaurally to the better ear through the insert earphones. In the case of symmetrical hearing sensitivity, the right ear was chosen as the test ear.

Time-Compression Procedures

In this study the time-compression rate refers to the amount of compression applied to the signal. For example, a 60% time-compression rate means that the processed speech has a duration of 40% of its original time. A custom software program was used to time compress all the speech materials for this study. This algorithm is based on a modified synchronized-overlap-add (SOLA) method adapted from Hardam (1990), originally proposed by Roucos and Wilgus (1985).

Audiometric Tests

Pure-tone air-conduction threshold testing was conducted at octave and interoctave frequencies from 250 through 8000 Hz and bone-conduction thresholds at octave frequencies up to and including 4000 Hz.

Neuropsychological Test Procedures

The neurocognitive tests were administered by a neuropsychologist or by psychometrists trained and supervised by the neuropsychologist. Standardized protocols for each test were strictly adhered to, and scoring was according to the test standardized procedures. A standard desktop computer with a large monitor for easier viewing was used to administer computerized neurocognitive tests. Other tests used pencil and paper or live speech presentation with verbal responses.

Speech Recognition Test Procedures

It has been shown in the NCRAR laboratory that three or four sentences are sufficient to familiarize the listener with time-compressed speech. In this study, ten sentences were used for practice at each time-compression rate. Feedback was provided for the first five practice sentences. Each practice set was administered just prior to the actual test at each corresponding rate to prepare participants to listen and respond to speech at that rate. For the experimental sessions, listeners were asked to repeat individual sentences as accurately as possible in each speech recognition test described above. Responses were recorded for later scoring by the examiner and one additional scorer. Scoring was based on percent correct key word recognition. For each test, ten different sentences were presented at each of the test rates during the test session. Each speech recognition test began with ten sentences at the normal rate of speech (0% time compression) for a baseline measurement followed by either two or three time-compression rates. Order of presentation of the compression rates was determined by an adaptive procedure. All participants were initially tested at the 50% time-compression rate (TC50). Those who failed to score better than 60 percent correct were then tested at 40% compression rate (TC40) so that only two rates were tested. Those who scored above 60% correct were tested at the 60% rate (TC60) and moved on to the 65% TC rate (TC65) if they again scored 60% correct or better. If the score was less than 60% correct, they were tested at TC40 in order to provide responses at three rates. Order of Anomalous (ANOM) sentences and IEEE (IEEE) sentences were alternated across sessions.

All sentences were presented at 90 dB SPL monaurally through an ER3A insert earphone. The National Acoustics Laboratory NAL formula (Byrne and Dillon, 1986) for linear hearing aid fitting was applied through the master hearing aid function of the

Otowizard computer/audiometer for participants whose hearing loss exceeded 25 dB HL at any frequency in the test ear. The objective of the master hearing aid procedure was to individualize the 90 dB SPL presentation level appropriately for hearing-impaired listeners by adding gain appropriate to the degree and configuration of hearing loss.

RESULTS

The data reported here are from a three-year study at the Veterans Affairs National Center for Rehabilitative Auditory Research (NCRAR) at the Portland, Oregon, Veterans Affairs Medical Center. This research was approved by the Institutional Review Board of the Portland VA Medical Center, and all participants signed an informed consent form. After hearing loss criteria were satisfied, an additional 22 participants were excluded from the analyses. Twenty-one were excluded based on BDI-II scores of greater than 19, and one was excluded for CVLT-II z-score less than -2. None was excluded for a WASI IQ score of less than 70. The data were analyzed for 176 participants.

Table 1 displays the mean sentence recognition scores and standard deviations for each time-compression rate and for the normal rate for each type of sentence (ANOM and IEEE). Number of participants (n) who participated in each sentence condition (type and rate) varied. The results of a paired T-test confirmed that the IEEE overall mean scores were significantly better than the ANOM at all rates (p = .000) except at 65% time compression (p = .443). At the 65% rate, average scores for both types of sentences were below 50%, and scores were slightly higher on the ANOM sentences than on the IEEE.

The adaptive nature of the speech recognition testing resulted in fewer scores at the lowest time-compression (TC) rate (40%) because only listeners who were unable to achieve a 60% correct score or better at 50% and 60% were tested at the slower 40% rate. Fewer listeners also were able to progress to the (most difficult) 65% time-compression rate, and scores at that rate may be no better than chance. For this report, further analyses were carried out only for the 50% and 60% rates resulting in four sentence conditions, two sentence types based on context, and two rates of speech: IEEE 50%, IEEE 60%, ANOM 50%, ANOM 60%. Figure 1 displays the means and standard errors for only these four sentence conditions. It can readily be seen that the rate of speech (50% or 60%) had a greater effect on speech recognition than type of sentence (IEEE, ANOM) for these speech materials.

The goal of this study was to identify age-related cognitive changes that are associated with speech-understanding deficits in an older population. It was expected that time-compressed speech would be sensitive to the effects of specific cognitive functions, with working memory as the primary contributor to speech recognition deficits, and processing speed and attention as secondary factors. The neurocognitive tests were divided into three categories a priori: five tests of working memory (three visual and two auditory), three tests of processing speed, and two tests of attention. Raw overall scores from this battery of ten tests were subjected to a principal components analysis (PCA) with varimax rotation in order to reduce the number of cognitive variables and to identify appropriate groupings for those variables. The PCA resulted in three major cognitive components with eigenvalues greater than one (Table 2) accounting for approximately 60% of the total variance: "Component One,"

Table 1. Means and Standard Deviations (in parenthesis) of IEEE and ANOM Sentence Scores for the Four Time-Compression Rates Used in This Study and for the Normal Rate (0%)

	0%	40%	50%	60%	65%
IEEE	99.2 (1.8) n = 176	93.9 (6.9) n = 69	85.5 (14.6) n = 176	72.9 (16.8) n = 166	43.5 (17.3) n = 129
ANOM	96.6 (4.9) n = 176	89.8 (10.4) n = 82	82.5 (12.7) n = 176	68.2 (17.7) n = 164	46.7 (14.4) n = 114

Note: Number of participants (n) varies across sentence conditions due to the adaptive method of test.

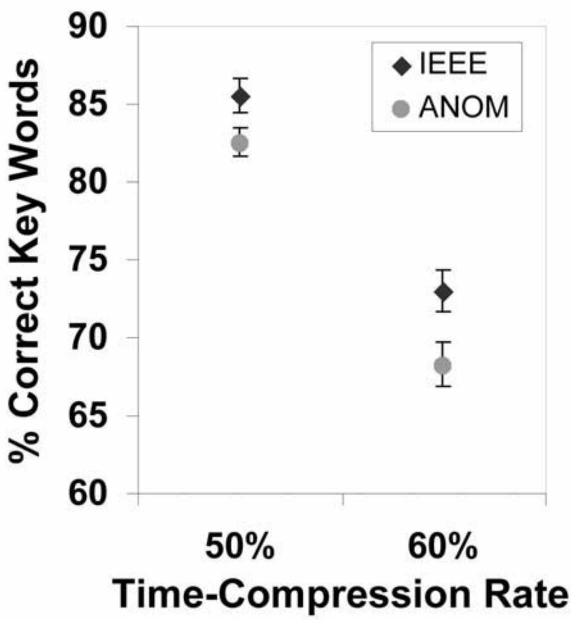

Figure 1. Group means and standard errors for each sentence type (IEEE vs. ANOM) at two time-compression rates (50% and 60%).

labeled as "nonsequential working memory" (WM-NS), accounting for 22.4% of the variance; "Component Two," labeled as "processing speed" (PS), accounting for 19.5% of the variance; and "Component Three," labeled as "sequential working memory" (WM-S), accounting for 19.2% of the variance. As indicated, two separate working memory components were identified by the PCA in addition to a processing-speed component.

The WM-S component, at first glance, appears to be an auditory working memory variable, since all three of the tests included in that component are verbal tests while the working memory tests in WM-NS include computerized visual tests. However, there is another unique characteristic that distinguishes the working memory tests in the WM-S component from those in the WM-NS component. Performance on each of the working memory tasks in WM-S critically depends on sequencing of stimulus and response items. The visually presented working memory tests (N-Back and SOP) in WM-NS do not involve sequential ordering of stimuli, although both are putative working memory tasks (Daigneault and Braun, 1993; Petrides et al, 1993; West et al, 1998; Ross and Segalowitz, 2000), and the N-Back has been associated with prefrontal cortex activity on fMRI (Honey et al, 2000; Jansma et al, 2000).

WM-NS consists of two computerized (N-back and SOP) working memory tasks, one computerized (CPT), and one auditory (BTA) attention/vigilance task. Of the two working memory tests, the N-back test requires continuous monitoring of ongoing stimuli (letters) while tracking a changing target (any letter that occurred one back or three back). The subject-ordered pointing (SOP) test is a tracking task also, using abstract drawings that are rearranged on each consecutive screen. The task is to choose a different picture each time in spite of the changing locations. There are 8, 10, or 12 drawings per screen depending on the version of the test. The auditory attention (BTA) and computerized vigilance (CPT) tasks included in WM-NS are easily reconciled with the focused attention required for the N-Back and SOP WM tests. Note that the Digit Symbol Coding (DSC) test shared similar variances with all three components. It is a complex paper and pencil test that is known to be influenced by all three variables of interest: working memory, processing speed,

Table 2. Loadings of 10 Cognitive Measures on Three Components Derived from Principle Component Analysis

Measure	Component One WM-NS	Component Two PS	Component Three WM-S
N-back	.<u>723</u>	-.077	-.047
SOP	-.<u>696</u>	-.086	-.120
BTA	.<u>623</u>	-.342	.249
CPT	.<u>545</u>	.027	.099
Digit Symbol	.522	-.478	.325
Visual RT	.021	.<u>914</u>	-.050
Auditory RT	-.032	.<u>820</u>	-.142
Digits Forward	-.110	-.015	.<u>868</u>
Digits Backward	.259	-.262	.<u>702</u>
LNS	.437	-.120	.<u>677</u>

and attention (Lezak, 1995).

Participants were divided into “LOW,” “MID,” and “HIGH” subgroups using cognitive component scores on each of the three components from the PCA. The LOW groups were comprised of component scores at and below the 25th percentile, and the HIGH groups of scores were at and above the 75th percentile for the two working memory components. Since the processing-speed component reflects response latencies, the faster responses (HIGH group) were represented by the 25th percentile and below while the slower responses (LOW group) were comprised of the scores at and above the 75th percentile. The interquartile range (middle 50% scores) constituted the MID groups of all three component scores. Table 3 displays mean ages and pure-tone averages (.5, 1, 2, 4 kHz) for each subgroup for each cognitive component. Note that for the WM-NS subgroups, the HIGH cognitive performers are, as one might expect, younger than the LOW and MID performers. However, this same pattern does not hold for WM-S and PS subgroups.

It was hypothesized that poor cognitive performers would also perform poorly on the time-compressed speech recognition tests. Figures 2a and 2b show a consistent pattern for sentence scores of subgroups within both working memory (WM-S and WM-NS) components. The HIGH cognitive groups are shown to have the best sentence recognition scores, and LOW cognitive groups reveal the poorest sentence recognition scores. However, there is no consistent sentence score pattern for the three subgroups within the PS component (Figure 2c).

Analyses of covariance (Table 4) were carried out to compare performance of HIGH, MID, and LOW cognitive groups on the sentence recognition tests within each of the 50% and 60% time-compression sentence conditions. Subscript numbers indicate the results of post hoc tests for each pair of comparisons. Main effects are adjusted for pure-tone threshold average hearing loss (500 Hz, 1000 Hz, 2000 Hz, 4000 Hz) with aided thresholds for those who were tested with the Master Hearing Aid (MHA). This initial analysis was not adjusted for age either by using age as a covariate or by using age-adjusted cognitive test scores within this limited range of ages.

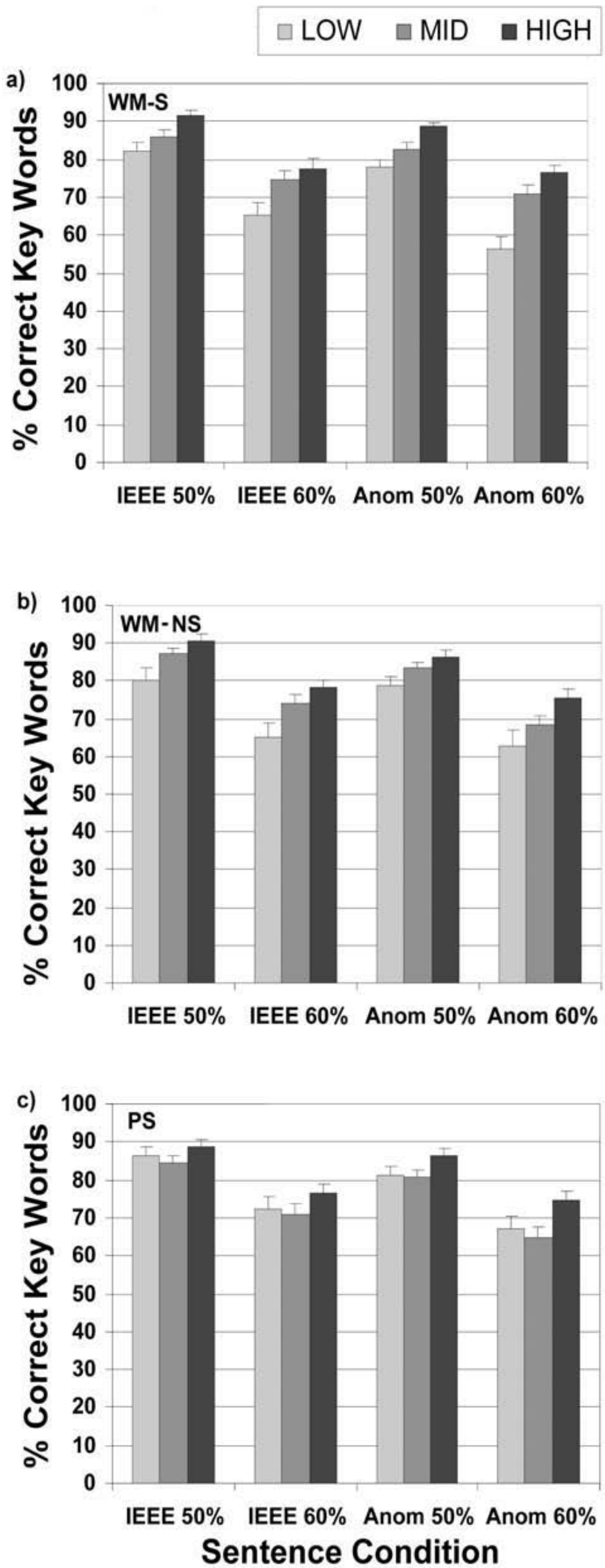

Figure 2. WM-S = sequential working memory *(a)*; WM-NS = nonsequential working memory *(b)*; and PS = processing-speed subgroup sentence scores for each sentence condition (type by rate) *(c)*.

Table 3. Mean Ages in Years and Pure-Tone Averages (PTA: .5, 1, 2, 4 kHz) in dB HL with Standard Errors (in parentheses) for Each Subgroup within Each of the Cognitive Components

	Sequential Working Memory (WM-S)			Nonsequential Working Memory (WM-NS)			Processing Speed (PS)		
	N	AGE	PTA	N	AGE	PTA	N	AGE	PTA
LOW	28	60.9 (1.2)	24.6 (1.7)	29	65.1 (1.2)	25.9 (1.9)	39	59.4 (1.0)	20.8 (1.6)
MID	59	62.2 (.90)	21.9 (1.2)	58	61.2 (.80)	20.8 (1.1)	47	64.9 (.70)	24.6 (1.1)
HIGH	28	60.7 (1.2)	19.8 (1.4)	28	58.4 (1.0)	20.7 (1.6)	29	62.6 (1.3)	20.6 (1.0)

The ANCOVAs were conducted using the SPSS V.13 statistical package with LSD post hoc comparisons. There were no significant differences among any of the subgroups from the Processing Speed component of the cognitive tests. The differences among subgroups of the Sequential Working Memory (WM-S) component were significant for all of the sentence conditions, although the differences were greater for the ANOM than for the IEEE sentence types. The subgroups of the Nonsequential Working Memory (WM-NS) differed significantly for the IEEE sentence performance but not for the ANOM sentences. When the ANCOVAs were repeated adjusting for both hearing loss and age, the significant differences among the groups for WM-S persisted, but the differences among the WM-NS subgroups were eliminated. This change in outcomes after adjusting for age suggested that the effect of WM-NS performance was age related.

To establish the association between the age range used in the current study and each of the three cognitive components, Pearson correlations were calculated. The correlation between age and WM-NS ($r = -448$) was significant ($p < .0001$), and that component showed the strongest correlation with age. Processing speed (PS) showed a weaker correlation with age ($r = .186$) although it was significant also ($p = .047$). The correlation between age and WM-S was very small ($r = -.078$) and was not statistically significant ($p = .410$). To further pursue the differences in cognitive performance due to aging, the participants were divided into age groups as follows: 50- to 59-year-old (middle-aged) and 60+ years (older). A comparison of these two age groups on performance on each of the cognitive components of the PCA revealed significant differences on both WM-NS ($F = 17.846$, $p < .0001$) and PS ($F = 4.427$, $p = .038$), but the two age groups performed similarly on the WM-S tasks ($F = 1.7779$, $p = .185$).

DISCUSSION

It has been suggested that, in order to identify the role of cognitive components in speech understanding difficulties, it is necessary to use cognitive measures that are specific to speech processing (Sommers, 1997; Hallgren et al, 2001). In the current study, the cognitive tests were targeted to identify associations of specific age-related cognitive deficits such as working memory, processing speed, and attention with the ability to process rapid speech for sentence recognition. Principal component analysis identified two

Table 4. ANCOVA Table for Comparison of Sentence Performance by Cognitive HIGH, MID, and LOW Subgroups within Each Sentence Condition

PCA Components	IEEE 50%		IEEE 60%		ANOM 50%		ANOM 60%	
	F	p-value	F	p-value	F	p-value	F	p-value
WM-S	3.078	**.050**[1]	3.344	**.039**[1,2]	5.710	**.004**[1,3]	10.225	**<.001**[1,2]
WM-NS	3.135	**.047**[1]	3.604	**.031**[1]	1.698	.188	2.876	.061
PS	.434	.625	.537	.492	2.261	.109	2.380	.098

Note: WM-S = sequential working memory; WM-NS = nonsequential working memory; PS = processing speed. Significant p-values are bolded. Significant post hoc comparisons are shown as superscripts: [1] = LOW vs. HIGH, [2] = LOW vs. MID, [3] = HIGH vs. MID.

separate types of working memory that were associated with verbal speech processing. One of these, the cognitive component labeled "sequential working memory" (WM-S), showed the most robust association with speech recognition. This component loaded quite clearly with tests that were assigned as working memory tests a priori. Another working memory component labeled "nonsequential working memory" was comprised of two working memory tests and two attention tests. This component had a weaker association with speech recognition. The remaining component labeled "processing speed" was comprised of visual and auditory choice reaction time tests and was not associated with time-compressed speech recognition.

The nonsequential working memory component (WM-NS) was associated with age as well as with speech recognition, consistent with our primary hypothesis. However, the WM-NS component included attention tests in addition to two working memory tests. An alternative interpretation of this component is that it reflects primarily an attention or a focused attention factor. Two working memory tests (n-back and SOP) that loaded on this component require continuous monitoring of successive stimuli with a changing target. At the neural level, the function of both attention and working memory in spoken language processing is to sustain activation of neural representation of information. It has been suggested that they are essentially the same process but that they differ in that attention is externally focused, while working memory is internally focused (Fuster, 2003). This dual working memory and attention component (WM-NS) showed a significant correlation both with age and with IEEE sentence recognition.

The WM-S (sequential working memory) showed a more robust association with the speech recognition performance but was not associated with age. This is in contrast to a recent study by Humes and Floyd (2005) that found age effects on other sequencing measures including sequence memory and sequence learning. The results showed that young normal-hearing participants were able to both learn and correctly repeat longer sequences than older participants. The sequence memory task was mildly to moderately correlated with the WAIS-R digit-span scores associated with working memory. The authors found mild-to-moderate correlations also between the sequencing scores and measures of aided and unaided speech recognition performance in elderly listeners. Contrary to the findings of the current study, their results suggest that poorer sequencing skills that relate to speech understanding are age related. However, it should be noted that Humes and Floyd compared young (21- and 22-years-old) and elderly (62- to 86-years-old) listeners while, in the current study, ages ranged from 50 (middle-aged) to 75 years. Thus, the age comparison in the current study shows performance differences only between middle-aged (50- to 59-years-old) and older (60- to 75-years-old) participants. As discussed earlier, both changes in cognitive function and in processing of temporally altered speech begin much earlier than "old" age. We are currently testing young adults (18–30 years old) to investigate the role of cognitive function in rapid speech recognition across a wider age span.

Sequencing ability is associated with higher cognitive function (Marshuetz, 2005) but involves also lower level temporal auditory processes. A recent study by Surprenant and Watson (2001) found that intellectual/academic measures obtained from archival academic records (GPA/SAT) were associated with temporal order test scores from a psychoacoustic test battery (Test of Basic Auditory Capabilities; Watson et al, 1982). However, those same tests were not associated with the results of speech recognition tests. They concluded that temporal ordering skills (i.e., sequencing) were related to general cognitive skills but not to speech processing although specific cognitive skills or functions were not tested in this study. Rather, cognitive status was inferred from the academic scores on previously administered standardized tests. The association of speech processing with working memory in the current study was likely due to the fact that more complex sentence-level materials at increased rates of speech required higher-level cognitive processing. The sequences of pure tones and of nonsense syllables used in the temporal ordering tests (Surprenant and Watson, 2001) may not provide enough cognitive load to reveal associations with more complex speech stimuli. Contrary to the Surprenant et al findings, Tallal et al (1985, 1996) has shown that temporal processing deficits are related to speech perception disorders in children.

In addition to the differences involving the effects of age and attention between the two working memory components in the current study, a third difference was the effect of sentence type (IEEE and ANOM) on sentence recognition performance. The WM-NS cognitive subgroups differed significantly on sentence recognition performance only for IEEE sentences but not for ANOM sentences. In contrast, among the WM-S subgroups, the significant differences in sentence recognition performance for ANOM sentences were more robust than those for IEEE sentences in either working memory component. Although Figure 1 suggested that rate of speech had a greater effect than type of sentence overall, when differences in cognitive functions were considered, the sequencing working memory functions were more important in the ability to understand speech that lacked contextual information even though proper syntax was maintained. One potential theory may be constructed based on electrophysiologic evidence of timing in speech processing.

Appropriate sequences of acoustic and phonetic elements are fundamental to both written and spoken language. It has been demonstrated that sequence processing is important for vocabulary acquisition and language production (Dell et al, 1997; Burgess and Hitch, 1999), but it may also be critical for speech understanding. Speech sequences occur at the word and sentence level as well as at the phoneme level, and processing must be fast enough to keep pace with typical speech rates of 200 to 270 words per minute (200 to 300 msec per word). Studies have shown that event-related potentials (ERPs), neural responses used to study the cognitive effects of certain changes in speech stimuli, occur as late as 400 msec to 800 msec poststimuli (Hagoort et al, 1993; Hagoort and Brown, 2000; Kutas and Iragui, 1998). This suggests that the rate of word presentation is faster than some higher-level language processing, thus the need for working memory to store information while simultaneously processing additional incoming information. When speech rate is increased, the demands on working memory are increased because the delay between the peripheral perception and processing is increased. The same is true when processing is slowed because of aging. The ANOM sentences in this study may have further increased working memory demands due to incongruous semantic and syntactic information (Gunter et al, 1995) added to the already heavy demands of the rapid speech.

The present findings have some limitations. The reported interim results have a limited focus and do not address all possible issues. For instance, the complex multivariate nature of our data suggests several alternate methods of analysis. Also, our participants' ages encompass a 25-year span, and although we accounted for age and hearing loss in the ANCOVA, there may be differences between youngest and oldest members of the sample that current analyses do not reveal. It has also not been possible, from these data, to determine which of the time-compressed speech recognition tests may be most sensitive to the cognitive issues raised in this report. We currently are obtaining data from participants younger than age 50. With an expanded sample from our continuing study in the future, it will be possible to assess age-related differences and to investigate the sensitivity of the speech recognition tests.

Cognitive measures that are specific to speech processing should be assessed in relation to understanding of different types of distorted speech. These findings have taken a step toward that goal. Future research should specifically address the issue of sequencing-dependent working memory and determine whether difficulties caused by rate of speech involve different cognitive processes than the difficulties encountered in noise for older listeners. With the advance of sophisticated signal processing technology in hearing aids, it is important to determine how this technology can be best used for older clients. Recent research involving signal processing strategies has shown that listeners who demonstrate better cognitive performance are better able to benefit from and also prefer nonlinear processing of aided speech using rapid time constants such as that in wide dynamic range hearing aids (Gatehouse et al, 2006a, 2006b). Others with poorer cognitive function prefer linear amplification. Continued research is planned at the NCRAR to investigate the use of time-compressed speech as a clinical tool to improve the accuracy of assessment of hearing aid benefit for older listeners based on potential cognitive changes. Further development of time-compressed speech tests may provide the clinician with simple auditory tests for specific cognitive deficits that affect speech recognition, particularly speech altered by temporal distortions.

Acknowledgments. The authors are grateful for the dedicated work of psychometrists Rebecca Hoffenberg and Renee Hohimer.

REFERENCES

ANSI. (1996) *American National Standard Specifications for Audiometers, S3.6*. New York: American National Standards Institute.

American Speech-Language-Hearing Association. (1990) Guidelines for screening for hearing impairments and middle ear disorders. *Asha* 32(Suppl.):17–24.

Brebion G. (2001) Language processing, slowing, and speed/accuracy trade-off in the elderly. *Exp Aging Res* 27:137–150.

Burgess N, Hitch GJ. (1999) Memory for serial order: a network model of the phonological loop and its timing. *Psychol Rev* 106:551–581.

Byrne D, Dillon H. (1986) The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear* 7:257–265.

Cohen JD, Forman SD, Braver TS, Casey BJ, Servan-Schreiber D, Noll DC. (1994) Activation of the prefrontal cortex in a nonspatial working memory task with functional MRI. *Hum Brain Mapp* 1:293–304.

Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE. (1997) Temporal dynamics of brain activation during a working memory task. *Nature* 386:604–608.

Conners CK. (2000) *Conners' Continuous Performance Test II: Technical Guide*. Toronto: Multi-Jealth Systems.

Daigneault S, Braun CM. (1993) Working memory and the Self-Ordered Pointing Task: further evidence of early prefrontal decline in normal aging. *J Clin Exp Neuropsychol* 15:881–895.

Dell GS, Burger LK, Svec WR. (1997) Language production and serial order: a functional analysis and a model. *Psychol Rev* 104:123–147.

Fry AF, Hale S. (2000) Relationships among processing speed, working memory, and fluid intelligence in children. *Biol Psychol* 54:1–34.

Fuster JM. (2003) *Cortex and Mind: Unifying Cognition.* New York: Oxford University Press.

Gatehouse S, Naylor G, Elberling C. (2006a) Linear and non-linear hearing aid fittings - 1. Patterns of benefit. *Int J Audiol* 45(3):130–152.

Gatehouse S, Naylor G, Elberling C. (2006b) Linear and non-linear hearing aid fittings - 2. Patterns of candidature. *Int J Audiol* 45(3):153–171.

Giambra LM. (1993) Sustained attention in older adults: performance and processes. In: Cerella J, Raybash JM, eds. *Adult Information Processing: Limits on Loss.* San Diego: Academic Press, 259–272.

Gordon-Salant S, Fitzgibbons PJ. (2001) Source of age-related recognition difficulty for time-compressed speech. *J Speech Lang Hear Res* 44:709–719.

Gunter TC, Jackson JL, Mulder G. (1995) Language, memory, and aging: a electrophysiological exploration of the N400 during reading of memory-demanding sentences. *Psychophysiology* 32:215–229.

Hagoort P, Brown CM. (2000) ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia* 38:1531–1549.

Hagoort P, Brown C, Groothusen J. (1993) The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Lang Cogn Process* 8:439–483.

Hallgren M, Larsby B, Lyxell B, Arlinger S. (2001) Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons. *J Am Acad Audiol* 12:357–370.

Hardam E. (1990) High quality time scale modification of speech signals using fast synchronized-overlap-add algorithms. Paper presented at the International Conference on Acoustics, Speech, and Signal Processing, IEEE, Albuquerque, NM.

Honey GD, Bullmore ET, Sharma T. (2000) Prolonged reaction time to a verbal working memory task predicts increased power of posterior parietal cortical activation. *Neuroimage* 12:495–503.

Humes L, Floyd S. (2005) Measures of working memory, sequence learning, and speech recognition in the elderly. *J Speech Lang Hear Res* 48:224–235.

Jansma JM, Ramsey NF, Coppola R, Kahn RS. (2000) Specific versus nonspecific brain activity in a parametric n-back task. *Neuroimage* 12:688–697.

Jerger J, Jerger S, Pirozzolo F. (1991) Correlation analysis of speech audiometric scores, hearing loss, age, and cognitive abilities in the elderly. *Ear Hear* 12:103–109.

Kutas M, Iragui V. (1998) The N400 in a semantic categorization task across 6 decades. *Electroencephalogr Clin Neurophysiol* 108:456–471.

Lezak MD. (1995) *Neuropsychological Assessment.* New York: Oxford.

Marshuetz C. (2005) Order information in working memory: an integrative review of evidence from brain and behavior. *Psychol Bull* 131:323–339.

Miller JL, Grosjean F, Lomanto C. (1984) Articulation rate and its variability in spontaneous speech: a reanalysis and some implications. *Phonetica* 41:215–225.

Morris R, Gick M, Craik F. (1988) Processing resources and age differences in working memory. *Mem Cogn* 16(4):362–366.

Petrides M. (1994) Frontal lobes and behavior. *Curr Opin Neurobiol* 4:207–211.

Petrides M, Alivisatos B, Meyer E, Evans AC. (1993) Functional activation of the human frontal cortex during the performance of verbal working memory tasks. *Proc Natl Acad Sci USA* 90:878–882.

Petrides M, Milner B. (1982) Deficits on subject-ordered tasks after frontal- and temporal-lobe lesions in man. *Neuropsychologia* 20:249–262.

Ross P, Segalowitz SJ. (2000) An EEG coherence test of the frontal dorsal versus ventral hypothesis in n-back working memory. *Brain Cogn* 43:375–379.

Rothauser EH, Chapman WD, Guttman N, Nordby KS, Silbiger HR, Urbanek GE, Weinstock M. (1969) IEEE recommended practice for speech quality measurements. *IEEE Trans Audio Electroacoust* 17:225–246.

Roucos S, Wilgus A. (1985) High quality time-scale modification for speech. Paper presented at the International Conference on Acoustics, Speech, and Signal Processing, IEEE.

Salthouse TA. (1992) Influence of processing speed on adult age differences in working memory. *Acta Psychol* 79:155–170.

Schretlen D, Bobholz JH, Brandt J. (1996a) Development and psychometric properties of the Brief Test of Attention. *Clin Neuropsychol* 10:80–89.

Schretlen D, Brandt J, Bobholz JH. (1996b) Validation of the Brief Test of Attention in patients with Huntington's Disease and amnesia. *Clin Neuropsychol* 10:90–95.

Sommers M. (1997) Speech perception in older adults: the importance of speech-specific cognitive abilities. *J Am Geriatr Soc* 45:633–637.

Strayer DL, Wichkens CD, Braune R. (1987) Adult age differences in the speed and capacity of information processing: II. An electrophysiological approach. *Psychol Aging* 2:99–110.

Sullivan MP, Faust ME. (1993) Evidence for identity inhibition during selective attention. *Psychol Aging* 8:589–598.

Surprenant AM, Watson CS. (2001) Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners. *J Acoust Soc Am* 110:2085–2095.

Tallal P, Miller S, Bedi G, Byma G, Wang X, Nagarajan S, Schreiner C, Jenkins W, Merzenich M. (1996) Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science* 271:81–84.

Tallal P, Stark R, Mellits E. (1985) Identification of language-impaired children on the basis of rapid perception and production skills. *Brain Lang* 25:314–322.

Tun PA, Wingfield A, Stine EA, Mecsas C. (1992) Rapid speech processing and divided attention: processing rate versus processing resources as an explanation of age effects. *Psychol Aging* 7:546–550.

van Rooij JCGM, Plomp RJ. (1990) Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses. *J Acoust Soc Am* 88:2611–2624.

Vaughan N, Letowski T. (1997) Effects of age, speech rate, and type of test on temporal auditory processing. *J Speech Lang Hear Res* 40:1192–1200.

Verhaeghen P. (1999) The effects of age-related slowing and working memory of asymptotic recognition performance. *Aging Neuropsychol Cogn* 6:201–213.

Watson CS, Johnson DM, Lehman JR, Kelly WJ, Jensen JK. (1982) An auditory discrimination test battery. *J Acoust Soc Am* 71:S73.

Wechsler D. (1997) *Wechsler Adult Intelligence Scale® - Third edition (WAIS® - III)*. San Antonio: The Psychological Corporation.

West R, Ergis AM, Winocur G, Saint-Cyr J. (1998) The contribution of impaired working memory monitoring to performance of the self-ordered pointing task in normal aging and Parkinson's disease. *Neuropsychology* 12:546–554.

Wingfield A, Lindfield KC, Goodglass H. (2000) Effects of age and hearing sensitivity on the use of prosodic information in spoken word recognition. *J Speech Lang Hear Res* 4:915–925.

Wingfield A, Poon LW, Lombardi L, Lowe D. (1985) Speed of processing in normal aging: effects of speech rate, linguistic structure, and processing time. *J Gerontol* 40(5):579–585.