
Sources of Age-Related Recognition Difficulty for Time-Compressed Speech

Sandra Gordon-Salant
University of Maryland
College Park

Peter J. Fitzgibbons
Gallaudet University
Washington, DC

Older people frequently show poorer recognition of rapid speech or time-compressed speech than younger listeners. The present investigation sought to determine if the age-related problem in recognition of time-compressed speech could be attributed primarily to a decline in the speed of information processing or to a decline in processing brief acoustic cues. The role of the availability of linguistic cues on recognition performance was examined also. Younger and older listeners with normal hearing and with hearing loss participated in the experiments. Stimuli were sentences, linguistic phrases, and strings of random words that were unmodified in duration or were time compressed with uniform time compression or with selective time compression of consonants, vowels, or pauses. Age effects were observed for recognition of unmodified random words, but not for sentences and linguistic phrases. Analysis of difference scores (unmodified speech versus time-compressed speech) showed age effects for time-compressed sentences and phrases. The forms of time compression that were notably difficult for older listeners were uniform time compression and selective time compression of consonants. Indeed, poor performance in recognizing uniformly time-compressed speech was attributed primarily to difficulty in recognizing speech that incorporated selective time compression of consonants. Hearing loss effects were observed also for most of the listening conditions, although these effects were independent of the aging effects. In general, the findings support the notion that the problems of older listeners in recognizing time-compressed speech are associated with difficulty in processing the brief, limited acoustic cues for consonants that are inherent in rapid speech.

KEY WORDS: aging, time-compressed speech, hearing loss, speech recognition, presbycusis

Older listeners have difficulty understanding speech, particularly when the speech is presented at a rapid rate. Wingfield, Poon, Lombardi, and Lowe (1985) posited that age-related changes in understanding naturally produced rapid speech reflect a decline in rapid information processing. Moreover, this decline in processing speed interacts with the linguistic redundancy in the spoken message, as evidenced by more substantial age-related deficits for recognition of items with minimal syntactic and semantic information compared to items containing multiple syntactic and semantic cues. These findings overall are consistent with cognitive theories of aging that suggest there is an overall decline in the speed of mental perceptual processing with increasing age (Birren, Woods, & Williams, 1980; Salthouse, 1985) and a decline in available processing resources with increasing age, which becomes apparent for complex stimuli or tasks

requiring multiple levels of processing (Craik & Byrd, 1982).

Age-related decline has been observed also for recognizing time-compressed speech, in which presentation rate is increased through mechanical or digital techniques using the sampling method (Letowski & Poch, 1995, 1996; Vaughan & Letowski, 1997). The problems in recognizing time-compressed speech that are attributed to age are independent of those attributed to hearing loss (Gordon-Salant & Fitzgibbons, 1993). A series of recent studies indicated that among a range of temporally distorted speech signals, those that incorporate time compression are notably difficult for older individuals (Gordon-Salant & Fitzgibbons, 1995, 1999). However, younger adults with hearing loss also show substantial decrements in recognizing time-compressed speech compared to age-matched individuals with normal hearing. Cognitive decline does not characterize younger individuals; thus, the findings for these listeners suggest that the reduced duration of acoustic information in the time-compressed speech signal may exceed the capacity of listeners with hearing loss to use brief acoustic cues (Dubno, Dirks, & Schaefer, 1987; Turner, Smith, Aldridge, & Stewart, 1997).

Because time-compressed speech appears to be particularly difficult for elderly listeners to recognize, and because rapid speech (>200 words per minute) is likely to be encountered in everyday listening situations (Wingfield et al., 1985), this study attempted to examine the possible mechanisms that may account for older listeners' difficulty in understanding this form of temporally altered speech. Two hypotheses are considered. One hypothesis postulates that the older listener's difficulty in recognizing time-compressed speech is attributed primarily to a slowed rate of information processing. This decline in processing rate would challenge the older person's ability to label chunks of acoustic information, process it serially, and derive meaning within a brief period of time. A decline in processing rate would be revealed most prominently for speech signals with a reduction in overall duration, but in which the principal acoustic cues are minimally compromised. Stimuli developed to assess this hypothesis include selective time compression of pauses between words, which only reduces the duration of silent intervals between words, and selective time compression of vowels, which reduces the steady-state portion of the vocalic segment without altering transition information. The alternative hypothesis is that older listeners experience more difficulty processing transient acoustic cues in speech than do younger listeners. The most obvious problem would be a reduced ability by older listeners to process very brief consonant cues, including bursts, transitions, and frication. Although time compression of consonant bursts, transitions, and frication segments represents a minimal reduction in the

overall duration of a sentence-length signal, it could result in a substantial reduction in recognition performance because of the importance of consonants for conveying the salient information in speech (Pickett, 1999). Thus, a finding that age-related problems are observed for time-compressed speech that incorporates selective time compression of consonants would suggest that the principal problem is bottom-up processing of an impoverished acoustic signal rather than a limitation associated with speed of information processing.

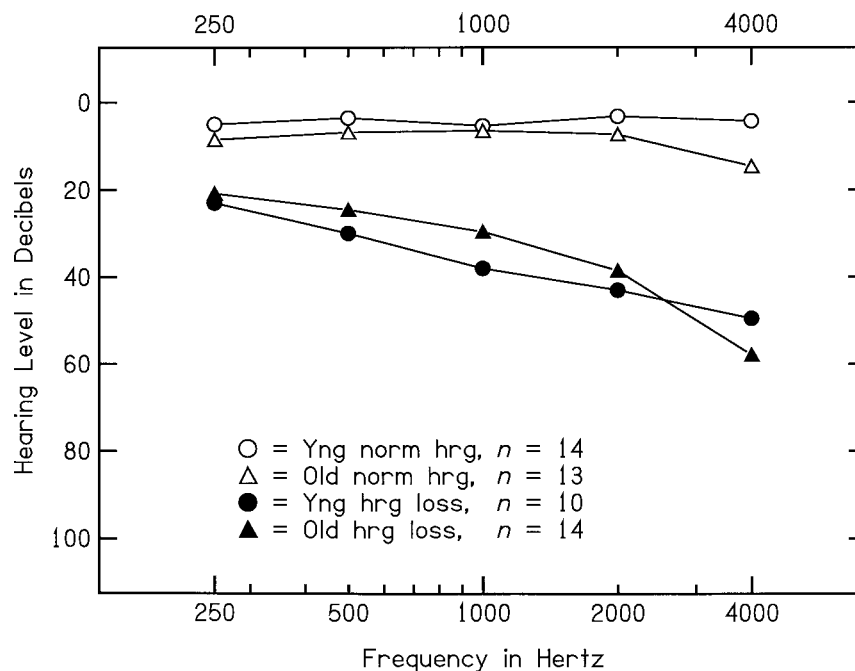
The goals of the present study were to examine the effects of age and hearing loss on recognition of time-compressed speech, and to determine if the age-related decline in performance could be attributed primarily to a decline in processing speed or to a decline in processing brief acoustic cues. A related aim was to examine the influence of the availability of linguistic cues on the age-related patterns of performance with time-compressed speech. To that end, recognition performance by younger and older listeners with and without hearing loss was examined for sentence-length material that was either time compressed uniformly or that incorporated selective time compression of pauses, vowels, or consonants. Additionally, three forms of sentence-length speech stimuli were used: sentences, syntactic sets, and random-order words that were presented to listeners in order to examine the added effects of syntactic cues. Sentence-length materials were used as stimuli, rather than individual words, to better approximate the duration of a spoken message, as well as to have a sample of speech of sufficient length to manipulate the temporal cues of interest.

Method

Participants

Four groups participated in the experiments. They included a group of 14 young listeners (21–34 years of age, $M = 25.36$ years) with normal hearing (YNH: pure-tone thresholds from 250 to 4000 Hz between 0 and 15 dB HL, re: ANSI, 1996), a group of 13 older listeners (65–72 years of age, $M = 67.31$ years) with normal hearing (ONH), a group of 10 younger listeners (19–40 years, $M = 31$ years) with hearing loss (YHL), and a group of 14 older listeners (65–75 years, $M = 70.73$ years) with hearing loss (OHL). The listeners in the two hearing loss groups had gradually sloping sensorineural hearing losses of mild-to-moderate degree and were matched for pure-tone thresholds across frequency on a paired basis. Mean audiograms of the four participant groups are shown in Figure 1. For the groups with hearing loss, the range of pure-tone thresholds was approximately ± 10 dB of the average thresholds at each frequency. Additional audiometric criteria for participant selection

Figure 1. Mean pure-tone air-conduction thresholds for the four listener groups: Yng Norm Hrg (younger listeners with normal hearing), Old Norm Hrg (older listeners with normal hearing), Yng Hrg Loss (younger listeners with hearing loss), and Old Hrg Loss (older listeners with hearing loss). Also shown is the size of the sample in each listener group.



were monosyllabic word recognition scores in quiet $\geq 80\%$ (Northwestern University Auditory Test No. 6); tympanograms indicating a normal pressure peak, peak admittance, tympanometric width, and volume; acoustic reflex thresholds elicited at levels within the 90th percentile range for equivalent pure-tone thresholds (Silman & Gelfand, 1981); and negative findings for acoustic reflex adaptation. The etiology of hearing loss for listeners in the YHL group included heredity and noise exposure. The etiology of hearing loss for listeners in the OHL group was assumed to be presbycusis, based on an absence of a significant otologic history and a gradual onset and progression of hearing loss during the 6th or 7th decade of life. Most listeners in the OHL and YHL groups had bilateral, symmetrical hearing losses.

In addition to the audiologic criteria, participants were required to be native speakers of English, pass a screening test of cognitive function (Pfeiffer, 1975), and possess sufficient motor control to provide a written response to the speech stimuli in a timely manner. The Short Portable Mental Status Questionnaire (Pfeiffer, 1975) is a simple screening test of general cognitive awareness, with 10 questions querying the respondent's knowledge of personal information and current events. Students at the University of Maryland were recruited to serve as participants in the YNH group. Clients of the University of Maryland Hearing Clinic and their

family members were invited by letter to participate in the other three listener groups.

Stimuli

Sentences from the eight forms of the Revised Speech Perception in Noise Test (R-SPIN; Bilger, Nuetzel, Rabinowitz, & Rzezchowski, 1984) formed the corpus of original speech stimuli from which the test stimuli were derived. The low-probability R-SPIN sentences, which contain minimal semantic contextual cues (e.g., "She wants to talk about the crew"), were selected. These stimuli were digitized onto a laboratory computer for waveform editing to create three stimulus forms: sentences, syntactic sets, and random-order words. The sentences (SENT) were essentially the original sentence waveforms of the R-SPIN test. The syntactic sets (SS) were created by separating each sentence into three phrases: noun phrase, verb phrase, and object phrase. The word boundaries that separated these three phrases were identified auditorily and visually, and then marked on the sentence stimulus waveform at a zero-crossing. The three phrases were then re-ordered randomly, into one of five possible orders, for each sentence (i.e., all possible phrase orders except 1-2-3). An example of these stimuli is "Wants to talk she about the crew." The random-order words (ROW) stimuli were developed in

a manner similar to that of the syntactic sets, except that the word boundaries were identified between each pair of sequential words in the original sentences, and then the order of the words was randomized across sentences. An example of a ROW stimulus is "To crew talk the wants she about." The final set of baseline stimuli consisted of eight lists of 25 sentences for each of the three stimulus modes (SENT, SS, ROW). Thus, these sentence-length stimuli varied primarily in the availability of syntactic (grammatical) cues.

Time compression of the stimuli was accomplished by a software algorithm created by H. T. Bunnell of Dupont Labs, and described elsewhere (Gordon-Salant & Fitzgibbons, 1993). A time-compression ratio (TCR) of 50% was selected on the basis of prior results with the sentence stimuli indicating that all listener groups would show a significant score decrement (re: 0% TCR) without creating a floor effect (Gordon-Salant & Fitzgibbons, 1993). Four types of time compression were applied to each of the three stimulus forms described above. Before implementing these types of time compression, the beginning and end of each consonant, vowel, and pause were identified auditorily and visually on the stimulus waveform and marked as closely as possible to a zero-crossing. The first type of time compression was uniform time compression (UNI TC), in which the time-compression algorithm was applied throughout the entire duration of the sentence-length stimulus. That is, each consonant, vowel, and pause within a sentence was compressed by 50% in duration in order to accomplish the overall sentence time compression of 50%. The second type of time compression was selective time compression of pauses (STC-P). Each of the pause segments from a sentence-length stimulus was excised, written to a new file, time compressed by 50% of the original duration, and placed back in the stimulus. It should be noted that a silent interval was considered a pause only if it was not part of a stop consonant and its duration exceeded 20 ms. Thus, each pause in a sentence-length stimulus was time compressed by 50%, resulting in 50% time compression for the total duration of all pauses in the sentence. The third type of time compression was selective time compression of vowels (STC-V). The same process was used to create these stimuli as was used for the selective time compression of pauses, except that only vocalic segments were excised, time compressed, and placed back in the original waveform. Finally, selective time compression of consonants (STC-C) involved excising, compressing, and appending the consonant segments of each sentence-length stimulus. All stimuli were scaled in amplitude to be equivalent in overall RMS level.

A total of 120 stimulus lists were created, including 96 time-compressed lists (8 lists \times 3 forms \times 4 types of time compression) and 24 undistorted stimulus lists,

each with 25 sentence-length stimuli. The stimulus lists were recorded onto digital audiotape (DAT) following conversion into analog form and anti-aliasing filtering. Stimuli were recorded with a 12-second interstimulus interval to provide sufficient time for a written response.

Procedures

The experimental procedures included five conditions, each presented in the three stimulus forms (SENT, SS, ROW). The five conditions were undistorted speech, UNI TC, STC-P, STC-V, and STC-C. There was random assignment of stimulus list to condition, and the order of the conditions was randomized over listeners.

During the experiments, the recorded stimuli were played on the DAT player, amplified (Crown D75), attenuated (Hewlett-Packard 350D), amplified through an audio mixer-amplifier (Colbourn S82-24), and transduced through a monaural insert earphone (Etymotic ER3A). The stimulus level was uniform for all participants at 90 dB SPL to ensure audibility for the listeners with hearing loss and to equate stimulus levels across groups. The test ear was arbitrarily chosen as the right ear for listeners with normal hearing. For listeners with hearing loss, the test ear was the ear with better thresholds (to provide the best stimulus audibility) or the ear with better speech recognition score (in the case of symmetrical audiograms). The listener's task was to write the entire sentence or sentence-length stimulus that he or she perceived. Listeners were encouraged to guess at any words that weren't clear or intelligible.

Stimuli were calibrated daily so that the level of a 1000-Hz tone, recorded at the beginning of each tape and equal in RMS level to the overall RMS level of the stimuli on each tape, was equal to 90 dB SPL. All participants were tested in a double-walled, sound-attenuating chamber that meets ANSI standards for maximum permissible noise levels for ears-covered testing (ANSI, 1991). Testing was completed in two test sessions of approximately 2 hours each, scheduled one week apart.

Participant responses were scored for the number of content words accurately identified in each sentence list. Content words included nouns, verbs, adjectives, prepositions, and adverbs. Function words, such as articles, were not scored for accuracy. Spelling errors were not counted as speech recognition errors. A percent correct score was derived for each list based on the total number of content words correct relative to the total number of content words in the list as a whole.

Results

The first level of analysis examined the effect of age and hearing loss on recognition of the three stimulus

forms, without time compression. Mean percent recognition scores of the four listener groups for sentences, syntactic sets, and random-order words are shown in Figure 2. Examination of this figure suggests that listeners with hearing loss performed more poorly than listeners with normal hearing, and that all listeners performed more poorly as the amount of syntactic redundancy was reduced (SENT > SS > ROW). Individual scores were analyzed using analysis of variance (ANOVA) with a split-plot factorial design with two between-subjects factors (age, hearing loss) and one within-subjects factor (stimulus form). Individual scores were arcsine transformed before data analysis. Results showed significant main effects of hearing status [$F(1, 46) = 11.33, p < .01$] and stimulus [$F(2, 92) = 748.21, p < .01$], and a significant interaction between age and stimulus [$F(2, 92) = 6.15, p < .01$]. Listeners with hearing loss showed significantly poorer performance than listeners with normal hearing in all stimulus conditions. Simple effects analyses revealed that the source of the interaction was a significant age effect for the ROW stimuli, but not for either of the other two stimulus forms. For the ROW stimuli, older subjects demonstrated poorer recognition performance than younger subjects. For all listener groups, recognition performance was significantly higher for SENT than for SS and ROW, and significantly higher for SS than for ROW (Scheffe, $p < .01$).

The effect of time compression was analyzed separately for each of the three stimulus forms because the effect of stimulus form for undistorted speech was significant. Additionally, because there were significant differences between listener groups for the undistorted stimuli, subsequent examination of the data involved using difference scores as the dependent variable. Difference scores were calculated for each participant as the difference between performance for the undistorted speech condition and performance for one of the time-compression conditions. One of the younger listeners with hearing loss exhibited skewed performance that was several orders of magnitude poorer than that of other listeners in the same group. As a consequence, this individual's data were removed from subsequent analyses. For all analyses, a repeated measures ANOVA was conducted with the between-subjects factors of age and hearing loss and a within-subjects factor of time-compression type (four levels).

Mean difference scores of the four listener groups for the sentence stimuli are shown in Figure 3. ANOVA results indicated significant main effects of compression [$F(3, 138) = 109.96, p < .01$], age [$F(1, 46) = 5.50, p < .05$], and hearing status [$F(1, 46) = 14.75, p < .01$], and significant interactions of Compression \times Hearing [$F(3, 138) = 7.67, p < .01$] and Compression \times Age [$F(3, 138) = 9.15, p < .01$]. Simple effects analyses revealed that the

Figure 2. Percent correct recognition scores (means and standard deviations) of the four listener groups for normal speech rate stimuli (three forms).

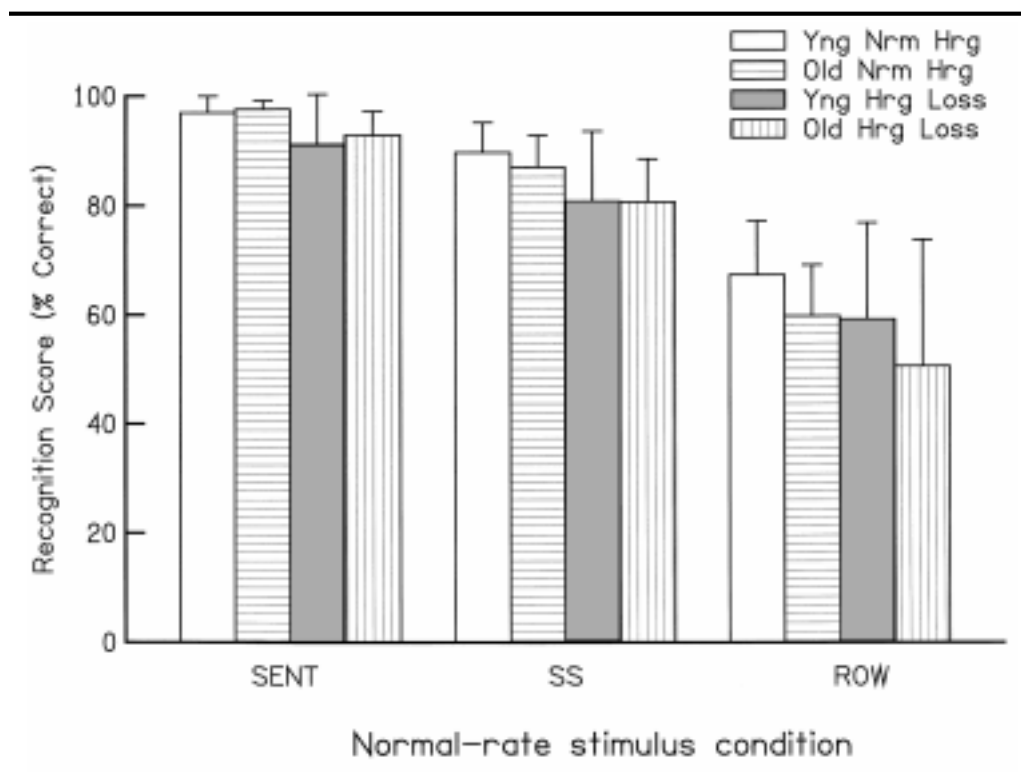
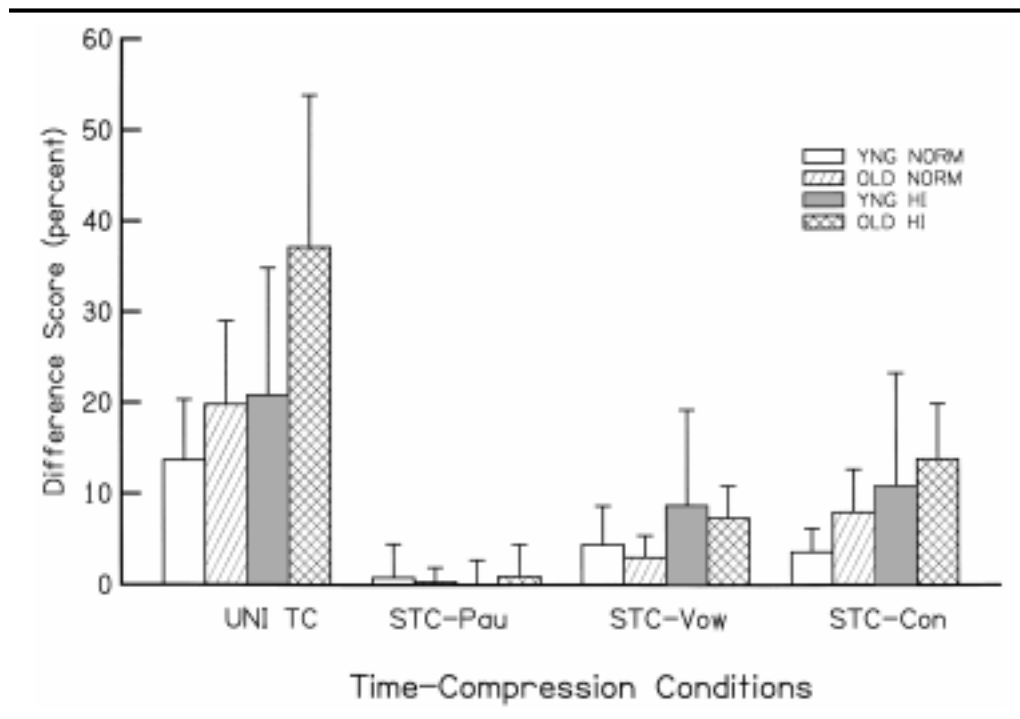


Figure 3. Difference scores of percent correct performance (means and standard deviations) for undistorted sentences versus time-compressed sentences, for four types of time compression: uniform 50% time compression (UNI TC), selective time compression of pauses (STC-Pau), selective time compression of vowels (STC-Vow), and selective time compression of consonants (STC-Con).



Compression \times Age interaction derived from significant age effects for the uniform time-compression condition only. The Compression \times Hearing interaction was associated with hearing-loss effects in three of the time-compression conditions (UNI TC, STC-C, STC-V). There were no hearing loss effects or age effects for selective time compression of pauses.

A somewhat similar pattern was observed for the syntactic sets, as shown in Figure 4. ANOVA revealed significant main effects of age [$F(1, 46) = 7.51, p < .01$], hearing [$F(1, 46) = 8.63, p < .01$], and compression [$F(3, 138) = 155.50, p < .01$], and significant interactions between compression and age [$F(3, 138) = 6.62, p < .01$], and between compression and hearing status [$F(3, 138) = 6.70, p < .01$]. Simple effects analyses showed that the age effect was significant for two conditions: uniform time compression and the selective time compression of consonants ($p < .01$). The effect of hearing loss was also significant for these two conditions only.

For random-order words (shown in Figure 5), analyses revealed a significant main effect of compression [$F(3, 138) = 224.18, p < .01$]. Other main effects and interactions between effects were not significant. Multiple comparison tests showed that for all subjects, there was a progressive and significant decline in score (that is, the difference score increased significantly), relative to the

STC-P, for the following conditions: STC-V, STC-C, and UNI TC.

The final data analysis sought to determine the extent to which performance decline (relative to undistorted recognition performance) in the three selective time-compression conditions predicted performance in the uniform time-compression condition. Separate multiple regression analyses were performed for the three stimulus forms (SENT, SS, ROW). For SENT and ROW, the only significant predictor variable was STC-C, with an associated $r = .73$ (SENT) and $r = .61$ (ROW). For SS, significant predictor variables were STC-C and STC-V ($r = .73$).

Discussion

The overall objective of this study was to examine possible sources of older listeners' difficulty in recognizing time-compressed speech. To that end, the effects of age and hearing loss were examined as well as the significance of linguistic contextual cues. Also assessed was the extent to which difficulty in understanding time-compressed speech could be attributed to a decline in processing speed or to a decline in processing brief acoustic cues in consonants.

Figure 4. Difference scores of percent correct performance (means and standard deviations) for undistorted syntactic sets versus time-compressed syntactic sets, for four types of time compression: uniform 50% time compression (UNI TC), selective time compression of pauses (STC-Pau), selective time compression of vowels (STC-Vow), and selective time compression of consonants (STC-Con).

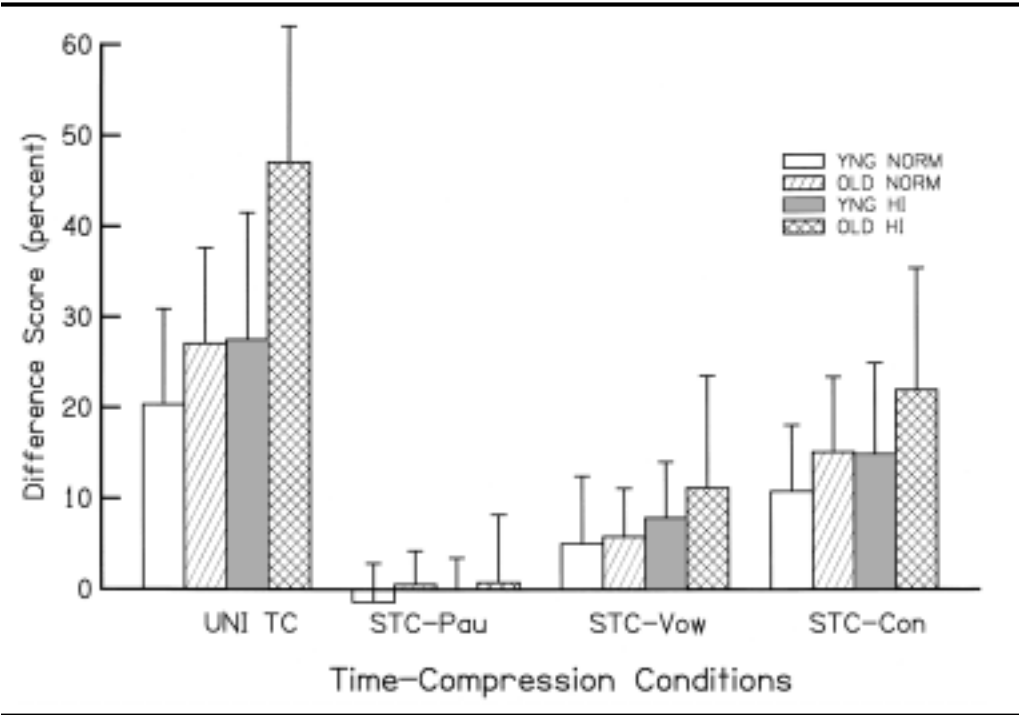
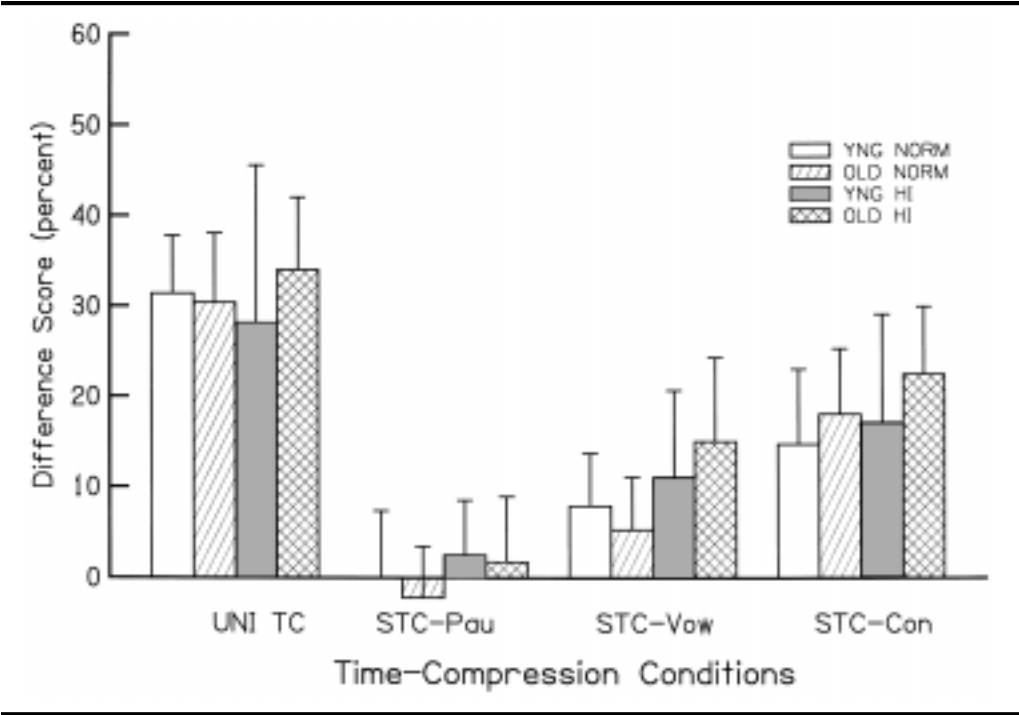


Figure 5. Difference scores of percent correct performance (means and standard deviations) for undistorted random-order words versus time-compressed random-order words, for four types of time compression: uniform 50% time compression (UNI TC), selective time compression of pauses (STC-Pau), selective time compression of vowels (STC-Vow), and selective time compression of consonants (STC-Con).



Effects of Syntactic Cues

Three stimulus forms—sentences, syntactic sets, and random-order words—were used in all of the experiments in order to examine the effect of reducing the syntactic cues on recognition of time-compressed speech. Speech materials with minimal semantic cues were chosen specifically for the experiments so that semantic context (i.e., meaning) would not be a confounding factor. The original sentences include multiple syntactic cues, the syntactic sets preserve some but not all of the syntactic cues, and the random-order words are completely devoid of these cues. It was hypothesized that age-related deficits would be more prominent for speech stimuli with minimal syntactic cues. Direct comparisons of recognition performance for the three stimulus forms were made for the stimuli presented at a normal speech rate. For all listener groups, there were significant differences in performance for each of the stimulus forms, with recognition of sentences exceeding recognition of syntactic sets, and recognition of syntactic sets exceeding recognition of random-order words. Thus, all listener groups experienced increasing difficulty with reduced syntactic cues. The interaction effect between age and stimuli indicated that elderly listeners were poorer at recognizing random-order words than younger listeners, although the two age groups performed similarly for the sentences and syntactic sets. For time-compressed stimuli, age effects were observed for the UNI TC condition for the sentence stimuli, and were observed for both the UNI TC and STC-C conditions for the syntactic sets. Uniform time compression of speech represents the most difficult listening condition for each of the sentence forms, because the time-compression algorithm was applied to every phoneme and silent interval throughout the stimulus, not just selective phonemes as in STC-C and STC-V. Age effects were not observed for the time-compressed ROW stimuli, perhaps because the analysis of difference scores removed the age-related difficulty in recognizing these stimuli observed at the normal speech rate. Nevertheless, all listeners experienced considerable difficulty in correctly identifying the uniformly time-compressed ROW stimuli (group mean score decrements ranging from 30% to 42% relative to recognition scores for normal-rate sentences; see Figure 1). These findings generally support the notion that age-related problems in speech recognition become more substantial for stimuli that incorporate reduced linguistic (i.e., syntactic) cues. Additionally, the effects of time compression interact with the linguistic redundancy of the speech material, suggesting that older listeners are more affected by time compression when contextual information in the spoken message is limited.

Wingfield and his colleagues (Wingfield et al., 1985; Wingfield, Tun, & Rosen, 1995) have also reported an

age-related decline for recognition of rapid speech when the amount of linguistic information was reduced systematically. These results can be interpreted according to a cognitive theory of aging, which stipulates that as people age, the amount of available processing resources diminishes (Craik & Byrd, 1982). Tasks that require multiple levels of effortful processing, including perceptual and cognitive demands, may overburden the limited processing resources of older listeners. For example, Pichora-Fuller, Schneider, and Daneman (1995) reported that word-recall performance declined significantly more for older listeners than for younger listeners in conditions that combined reduced acoustic information (background noise) and an added memory load. The present results suggest that tasks in which listeners resolve rapid speech cues and recall strings of words or phrases also tax multiple levels of processing and are particularly difficult for older listeners. Recollection of strings of words or phrases is a demanding cognitive task, because listeners must rely exclusively on the acoustic information in each individual word rather than benefiting from syntactic and semantic meaning derived from the sentence as a whole. Recognition of time-compressed speech may reflect a perceptual processing demand, in which rapid speech cues need to be perceived in bottom-up processing, or a cognitive demand, in which speech information must be processed at a rapid rate.

Effects of Time Compression

To identify the primary source of the older person's difficulty in recognizing time-compressed speech, we implemented four forms of time compression: uniform time compression, selective time compression of consonants, selective time compression of vowels, and selective time compression of pauses. One hypothesis was that the older listener's difficulty in processing time-compressed speech was attributed primarily to a decline in the ability to process information presented at a rapid rate rather than a decline in the ability to process impoverished, brief acoustic cues for speech. The ANOVA results comparing the difference scores for each form of time compression relative to 0% time compression revealed that older listeners showed significantly poorer recognition of 50% time-compressed sentences (UNI TC) than any other form of time compression. However, for syntactic sets, older listeners showed significantly poorer performance for both uniform time compression and selective time compression for consonants than for the other forms of time compression. For the random-order words, all listeners demonstrated essentially no performance decline for selective time compression of pauses, and a statistically significant, systematic decline in performance with selective time compression for vowels, consonants, and uniform time compression. Thus, it

appears that increasing the rate of speech presentation through selective time compression of vowels had a relatively minor effect on performance, particularly for older listeners. In contrast, listeners with hearing loss performed more poorly than did listeners with normal hearing for selective time compression of vowels. All listeners were virtually unaffected by selective time compression of pauses.

One reason for the minimal decline in performance with time compression of pauses could have been the limited opportunities for reducing the pause duration in the original sentences. These sentences were originally spoken in a recitation style, and waveform analyses revealed that there were generally only three or four brief pauses (<40 ms) in each sentence. Thus, reducing the pause duration by 50% represented only a slight reduction in presentation rate for the entire sentence. This was not the case for the syntactic sets and random-order words, however, because the digital cutting and pasting methods to create these stimuli also added silent intervals between words that were not inherent in the original sentence stimuli.

STC of vowels also had a relatively minor effect on listeners' recognition of the three types of speech stimuli (mean score decrements of 3–15% across stimuli). Each vowel in the original sentences was approximately 100 to 300 ms in duration. The time-compression algorithm was used to reduce the duration of these relatively long periodic segments for each occurrence throughout a sentence-length stimulus. As a consequence, the overall reduction in sentence duration was considerable, and second only to the uniform TC condition. In this condition, therefore, information was presented at a more rapid rate than for either of the other two STC conditions (pauses and consonants), without dramatically altering the acoustic structure of the vowels. Nevertheless, older listeners did not exhibit substantial decreases in recognition scores for this form of time compression, suggesting that an increase in the presentation rate is not the sole source of the processing deficit observed for time-compressed speech.

Older listeners showed a significant performance decline for STC for consonants (compared to STC-P and STC-V) for two stimulus forms (SS and ROW). Consonant duration in the original sentences was approximately 20–35 ms for stops, and 40–100 ms for fricatives, nasals, and glides. The time-compression algorithm reduced the duration of the bursts, transitions, and frication segments that formed these consonants by 50%. Because these acoustic cues are inherently brief and constantly changing, it is possible that the sampling method removed nonredundant information in a manner that resulted in a somewhat distorted sequence of acoustic information. The substantial difficulty that

older listeners exhibited for recognizing speech that incorporated STC for consonants indicates that one source of the age-related deficit in recognizing time-compressed speech is difficulty in perceptual processing of the very brief and possibly impoverished acoustic cues that convey consonant information. Listeners with hearing loss (both young and old) also demonstrated poorest performance among the selective time-compression conditions for STC of consonants. This finding is in agreement with previous results showing that listeners with hearing loss have substantial difficulty perceiving consonant place if transition duration or burst duration is reduced (Dubno et al., 1987; Schum & Collins, 1990; Turner et al., 1997). Time compression of consonant segments may also have altered the relative-amplitude cue for consonant perception that is used by listeners with hearing loss (Hedrick, Schulte, & Jesteadt, 1995). The problem for older listeners in recognizing time-compressed speech therefore appears to be similar to the processing deficit of listeners with hearing loss (including young listeners).

Multiple regression analyses were conducted to investigate this question further. For sentences and syntactic sets, the only significant predictor variable of performance in the UNI TC condition was performance in the STC-C condition. For random-order words, STC-C was identified first, followed by STC-V, as significant predictor variables for performance UNI TC condition, using stepwise multiple regression analysis. This analysis reaffirms the trend observed from the comparisons of means that reduction in the duration of acoustic cues for consonants is a primary source of listeners' processing difficulties for recognition of time-compressed speech. Separate multiple regression analyses were conducted also for each listener group, and analyses revealed trends for each individual group similar to the trends observed for all listeners combined. However, only about 49% of the variance in recognition scores for uniform time-compressed speech was accounted for by performance in the selective time-compression conditions, suggesting that other factors contribute also to listeners' deficits for speech that is time compressed in a uniform manner.

Effects of Hearing Loss

Though not the principal focus of this investigation, hearing loss was examined as a between-subjects variable that potentially could interact with aging effects. All of the statistical analyses demonstrated that hearing impairment was a significant variable affecting listeners' recognition of the speech stimuli, including the normal-rate stimuli. However, for each analysis, the hearing-loss effect was independent of the effect of aging, suggesting that these are two distinct factors that

act to reduce listener performance. The combined effect of these two factors is that older listeners with hearing impairment are at a substantial disadvantage compared to younger listeners and listeners with normal hearing for understanding time-altered speech and speech with reduced linguistic cues. Several different performance patterns were observed for the listeners with hearing loss, including the significant effect of selective time compression of vowels. The specific source of this performance decrement is unknown at present, although it may be associated with possible reductions in the availability of cues to final consonant voicing that are conveyed by vowel duration.

Summary and Conclusions

There are three principal findings in the present investigation of aging, speech form, and time compression. First, age-related deficits in speech recognition are influenced by the amount of available linguistic information, with greater age-related differences revealed for stimuli with reduced linguistic cues. From a practical perspective, this finding underscores the importance of linguistic cues for older listeners in understanding communication. From a theoretical perspective, it also suggests that removing the linguistic redundancy in the message represents an increase in cognitive demand that is sensitive to the effects of aging.

The second finding is that poor performance in recognizing time-compressed speech is attributed more to the acoustic alteration of consonants than to the acoustic alteration of vowels and pauses, which in turn affects the overall speed of the message. This result also has practical and theoretical implications. It suggests that by increasing the duration of consonants in rapid speech, there may be an improvement in overall recognition performance for older listeners and for listeners with hearing loss. Turner et al. (1997) attempted to slow the rate of frequency change in formant transitions and measured identification of voiced consonant-vowel syllables that varied in consonant place of articulation. Their findings showed that performance improvement with slowing was inversely related to the degree of hearing loss. Thus, there appears to be some merit in slowing selective aspects of the speech signal for some listeners, although the amount of benefit in everyday speech is as yet unknown. The theoretical import of the relative weight of selective time compression of consonants in contributing to recognition of time-compressed speech is that much of the difficulty experienced by older listeners in processing rapid speech is a result of limited processing capacity for brief consonant cues, rather than decline in the speed of information processing. This confirms Pickett's (1999) observation that in natural speech it is important to compress consonant duration

considerably less than vowel duration in order to preserve speech intelligibility at rapid rates of utterance.

Finally, the results show that older listeners are at a particular disadvantage in conditions that combine time compression of speech with reduced contextual information. These findings support the theory of a decline in processing resources available to older listeners, and extend this theory to include the increased bottom-up processing demand of reduced acoustic information in time-compressed speech, coupled with the top-down processing demand of reduced linguistic cues.

Acknowledgments

This research was supported by a grant from the National Institute on Aging (AG R01 09191). The authors are grateful to Leslie McCreight and Jennifer Lantz for their assistance with the collection of data for this project and to H. Timothy Bunnell for generously providing the time-compression and waveform-editing software.

References

- American National Standards Institute.** (1991). *American national standard maximum permissible ambient noise levels for audiometric rooms* (ANSI S3.1-1991). New York: Author.
- American National Standards Institute.** (1996). *American national standard specification for audiometers* (ANSI S3.6-1996). New York: Author.
- Bilger, R. C., Nuetzel, J. M., Rabinowitz, W. M., & Rzezczkowski, C.** (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research, 27*, 32–48.
- Birren, J. E., Woods, A. M., & Williams, M. V.** (1980). Behavioral slowing with age: Causes, organization, and consequences. In L. W. Poon (Ed.), *Aging in the 1980's: Psychological issues*. Washington, DC: American Psychological Association.
- Craik, F. I. M., & Byrd, M.** (1982). Aging and cognitive deficits: The role of attentional resources. In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes* (pp. 191–221). New York: Plenum.
- Dubno, J. R., Dirks, D. D., & Schaefer, A. B.** (1987). Effects of hearing loss on utilization of short-duration spectral cues in stop consonant recognition. *Journal of the Acoustical Society of America, 81*, 1940–1947.
- Gordon-Salant, S., & Fitzgibbons, P.** (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech and Hearing Research, 36*, 1276–1285.
- Gordon-Salant, S., & Fitzgibbons, P.** (1995). Recognition of multiply degraded speech by young and elderly listeners. *Journal of Speech and Hearing Research, 38*, 1150–1156.
- Gordon-Salant, S., & Fitzgibbons, P.** (1999). Profile of auditory temporal processing in older listeners. *Journal of Speech, Language, and Hearing Research, 42*, 300–311.
- Hedrick, M. S., Schulte, L., & Jesteadt, W.** (1995). Effect

- of relative and overall amplitude on perception of voiceless stop consonants by listeners with normal and impaired hearing. *Journal of the Acoustical Society of America*, 98, 1292–1303.
- Letowski, T., & Poch, N.** (1995). Understanding of time-compressed speech by older adults: Effect of discard interval. *Journal of the American Academy of Audiology*, 6, 433–439.
- Letowski, T., & Poch, N.** (1996). Comprehension of time-compressed speech: Effects of age and speech complexity. *Journal of the American Academy of Audiology*, 7, 447–457.
- Pfeiffer, E.** (1975). A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatric Society*, 23, 433–441.
- Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M.** (1995). How young and old adults listen to and remember speech in noise. *Journal of the Acoustical Society of America*, 97, 593–608.
- Pickett, J. M.** (1999). *The acoustics of speech communication: Fundamentals, speech perception theory, and technology*. Needham Heights, MA: Allyn & Bacon.
- Salthouse, T. A.** (1985). *A theory of cognitive aging*. New York: North-Holland.
- Schum, D. J., & Collins, M. J.** (1990). The time course of acoustic/phonemic cue integration in the sensorineurally hearing-impaired listener. *Journal of the Acoustical Society of America*, 87, 2716–2728.
- Silman, S., & Gelfand, S.** (1981). The relationship between magnitude of hearing loss and acoustic reflex thresholds. *Journal of Speech and Hearing Disorders*, 46, 312–316.
- Turner, C. W., Smith, S. J., Aldridge, P. L., & Stewart, S. L.** (1997). Formant transition duration and speech recognition in normal and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 101, 2822–2825.
- Vaughan, N. E., & Letowski, T.** (1997). Effects of age, speech rate, and type of test on temporal auditory processing. *Journal of Speech, Language, and Hearing Research*, 40, 1192–1200.
- Wingfield, A., Poon, L. W., Lombardi, L., & Lowe, D.** (1985). Speed of processing in normal aging: Effects of speech rate, linguistic structure, and processing time. *Journal of Gerontology*, 40, 579–585.
- Wingfield, A., Tun, P. A., & Rosen, M. J.** (1995). Age differences in veridical and reconstructive recall of syntactically and randomly segmented speech. *Journal of Gerontology, B: Psychological Sciences and Social Sciences*, 50, 257–266.

Received July 10, 2000

Accepted March 6, 2001

DOI: 10.1044/1092-4388(2001/056)

Contact author: Sandra Gordon-Salant, PhD, University of Maryland, Department of Hearing and Speech Sciences, LeFrak Hall, College Park, MD 20742.
E-mail: sgordon@hesp.umd.edu