

## Temporal Factors and Speech Recognition Performance in Young and Elderly Listeners

**Sandra Gordon-Salant**  
*University of Maryland  
College Park*

**Peter J. Fitzgibbons**  
*Gallaudet University  
Washington, DC*

This study investigated factors that contribute to deficits of elderly listeners in recognizing speech that is degraded by temporal waveform distortion. Young and elderly listeners with normal hearing sensitivity and with mild-to-moderate, sloping sensorineural hearing losses were evaluated. Low-predictability (LP) sentences from the Revised Speech Perception in Noise test (R-SPIN) (Bilger, Nuetzel, Rabinowitz, & Rzeczowski, 1984) were presented to subjects in undistorted form and in three forms of distortion: time compression, reverberation, and interruption. Percent-correct recognition scores indicated that age and hearing impairment contributed independently to deficits in recognizing all forms of temporally distorted speech. In addition, subjects' auditory temporal processing abilities were assessed on duration discrimination and gap detection tasks. Canonical correlation procedures showed that some of the suprathreshold temporal processing measures, especially gap duration discrimination, contributed to the ability to recognize reverberant speech. The overall conclusion is that age-related factors other than peripheral hearing loss contribute to diminished speech recognition performance of elderly listeners.

**KEY WORDS:** age effects, reverberant speech, time-compressed speech, interrupted speech, auditory temporal processing, hearing impairment

The speech recognition problems of elderly listeners are well-known but poorly understood. Considerable attention has been paid to the speech recognition deficits experienced by elderly listeners in noisy backgrounds (Bergman, 1971; Blumenfeld, Bergman, & Millner, 1969; Jokinen, 1973; Dubno, Dirks, & Morgan, 1984; Findlay & Denenberg, 1977; Goetzing, Proud, Dirks, & Embrey, 1961; Gordon-Salant, 1987; Townsend & Bess, 1980; Yanz & Anderson, 1984). In contrast, relatively little is known about elderly listeners' ability to recognize speech that has been degraded by some form of temporal waveform distortion. Temporal waveform distortions characterize reverberant speech, interrupted speech, and the speech of fast talkers. Elderly listeners may have particular difficulty processing temporally distorted speech because aging is thought to be accompanied by a generalized slowing of mental processes (Birren, Woods, & Williams, 1980) or diminished speed of sensory and perceptual encoding (Salthouse, 1982, 1985; Wingfield, Poon, Lombardi, & Lowe, 1985). A slowing of auditory temporal processing in an elderly person may delay or disrupt processing of the rapid, complex, and somewhat impoverished acoustic cues available in temporally distorted speech.

A few studies have reported diminished recognition performance of elderly listeners with temporally distorted speech signals. For example, Nabelek and Robinson (1982) showed that subjects aged 64 years and older exhibited significantly greater performance deficits than younger subjects at relatively long reverberation times (RT = 0.8 sec and 1.2 sec). Using time-compression techniques to simulate the speech of fast talkers, Konkle, Beasley, and Bess (1977) demonstrated that recognition of

time-compressed speech deteriorated with increasing age among subjects aged 54 to 84 years. Finally, Bergman, Blumenfeld, Cascardo, Dash, Levitt, and Margulies (1976) found that speech interrupted at the rate of 8 times per sec produced a sharp performance decline that increased with age above 40 years.

One limitation of these earlier studies was that the elderly listeners were characterized as "normally hearing for their age," whereas the young subjects had audiometrically normal hearing. Because decline in high frequency hearing sensitivity often increases with age, the poorer performances of the older listeners could have been related as much to high frequency hearing loss as to age effects. The presence of hearing impairment in younger listeners does reduce recognition of reverberant speech (Nabelek & Mason, 1981), interrupted speech (Korsan-Bengsten, 1973), and time-compressed speech (Grimes, Mueller, & Williams, 1984; Harris, Haines, & Myers, 1963). Thus, it is not yet known to what extent age, independent of hearing loss, contributes to speech recognition problems of elderly listeners for temporally degraded speech signals. One purpose of the present investigation was to examine this issue.

Several factors that contribute to speech recognition problems of elderly listeners have been identified recently. Van Rooij and Plomp (1990) investigated auditory and cognitive factors that are related to measured speech-reception thresholds (SRTs) of young and elderly listeners in quiet and noise. The SRTs of the elderly listeners were predicted primarily by reduced auditory sensitivity in the high frequencies and, secondarily, by reduced memory performance. Jerger, Jerger, and Pirozzolo (1991) also found that degree of hearing loss had the strongest relation to scores on four monotic speech recognition procedures. General cognitive status contributed secondarily to performance on two dichotic speech measures. Similarly, Humes and Christopherson (1991) found that auditory sensitivity was the primary component contributing to recognition of nonsense syllables presented in unfiltered, filtered, and filtered + reverberant conditions by elderly listeners with hearing loss. Additional psychoacoustic measures obtained from duration discrimination tasks, embedded test-tone tasks, and frequency discrimination tasks contributed somewhat less to the variance of unfiltered (undistorted) nonsense syllable recognition scores.

For younger listeners with hearing impairment, important factors contributing to speech recognition performance in quiet and noise include temporal difference limen (Tyler, Summerfield, Wood, & Fernandes, 1982) and gap detection (Dreschler & Plomp, 1985; Tyler et al., 1982). Gap detection is also related significantly to recognition of reverberant speech in noise by young listeners with hearing loss (Irwin & McCauley, 1987). Auditory temporal processing measures have not proven to be important predictors of speech recognition performance in elderly listeners for degraded speech materials (c.f. Van Rooij & Plomp, 1990). However, methodological constraints, such as use of nonsense syllables, measures of speech reception threshold rather than supra-threshold speech recognition, or use of very demanding measures of temporal resolution, may have limited such findings. Although one recent study (Moore, Peters, & Glasberg, 1992) concluded that reduced temporal gap resolution

does not accompany aging, other evidence (Fitzgibbons & Gordon-Salant, 1992) indicates that age can have a significant effect on auditory temporal resolution measures, independent of the effects of peripheral hearing impairment. Thus, auditory temporal processing deficits may be an important source of diminished speech recognition performance by elderly listeners, especially for speech signals distorted in the temporal domain.

The purpose of the present study was to investigate the sources of recognition deficits of elderly listeners for temporally distorted speech signals. The first hypothesis addressed was that age contributes independent of hearing loss to performance deficits in recognizing reverberant, interrupted, and time-compressed speech. A series of experiments was conducted in which each type of temporal waveform distortion was imposed in order to generate psychometric functions that related speech recognition performance to degree of degradation. Independent and interactive effects of age and hearing loss were determined by comparing the performance of young and elderly listeners with normal hearing and with significant hearing loss. The second hypothesis examined was that auditory temporal processing deficits of elderly listeners contribute to problems in recognizing temporally distorted speech. Eight auditory temporal processing measures were compared to the speech recognition measures obtained for all subjects, in an effort to identify subtle relations between specific parameters of auditory temporal processing and specific forms of temporal speech distortion.

## Methods

### Subjects

Four subject groups with 10 subjects each participated in the experiments. Group 1 was composed of elderly listeners, aged 65–76 years with normal hearing (pure tone thresholds  $\leq 15$  dB HL, re: ANSI, 1989, 250–4000 Hz). Group 2 was composed of young listeners, aged 20–40 years with normal hearing (pure tone thresholds  $\leq 15$  dB HL, re: ANSI, 1989, 250–4000 Hz). Group 3 consisted of elderly listeners (65–76 years, with mild-to-moderate, sloping sensorineural hearing losses. These subjects had a negative history for otologic disease, noise exposure, familial hearing loss, and ototoxicity. Thus, the presumed etiology of the loss for these subjects was presbycusis. The subjects in Group 4 were young (20–40 years) and had mild-to-moderate, sloping sensorineural hearing losses. Each subject in Group 4 was matched as closely as possible to a subject in Group 3 on the basis of pure tone thresholds between 250 and 4000 Hz and on suprathreshold word recognition scores. Mean pure tone thresholds and threshold ranges of the four subject groups are shown in Table 1. The young and elderly subjects with normal hearing showed mean threshold differences between 3.5 and 10.5 dB across frequency; however, the threshold ranges showed considerable overlap between age groups at each frequency. Similarly, threshold differences between young and elderly subjects with hearing loss were between 1 dB and 6 dB with comparable threshold ranges across frequency.

**TABLE 1.** Mean pure tone thresholds and threshold ranges (in parentheses) in dB HL for the four subject groups.

Group	Frequency (Hz)				
	250	500	1000	2000	4000
YNH	4.0 (0-10)	0.5 (-5-10)	2.0 (0-10)	2.0 (0-10)	3.0 (0-15)
ENH	7.5 (0-15)	7.5 (0-15)	7.0 (0-15)	7.0 (0-15)	13.5 (5-15)
YHI	17.5 (0-45)	22.0 (0-50)	28.5 (5-55)	44.5 (15-60)	51.5 (35-70)
EHI	24.0 (10-40)	26.5 (15-40)	29.5 (5-50)	38.0 (25-50)	56.0 (40-80)

Note. YNH = young normal hearing; YHI = young hearing impaired; ENH = elderly normal hearing; EHI = elderly hearing impaired.

All subjects exhibited word recognition scores (Northwestern University Test #6, Tillman & Carhart, 1966) in quiet exceeding 80%. Immittance measures confirmed the presence of normal middle ear function in each subject.

In addition to the audiometric and age characteristics, all subjects were native speakers of English, had a minimum high school education and good hand-motor control, and were in good general health. Each elderly subject passed the Short Portable Mental Status Questionnaire (Pfeiffer, 1975), a screening measure for cognitive function.

### Speech Materials

One criterion for stimulus selection was to use sentence-length materials to provide a stable sample of language in which to embed the test items and to implement the desired temporal waveform distortions. The speech stimuli selected were the low-predictability (LP) sentences of the Revised Speech Perception in Noise test (R-SPIN) (Bilger et al., 1984). The eight lists of LP R-SPIN sentences were selected specifically because they provide minimal semantic and linguistic cues to aid recognition of the final test word of the sentence.

The sentence stimuli were presented in undistorted form and in three forms of temporal distortion. There were four degrees of distortion for each type of temporal distortion. These degrees of distortion were determined as a result of pilot testing with subjects with normal hearing to sample performance between approximately 70 and 100% correct. Sets of temporally distorted speech stimuli were created digitally in the following manner. The stimuli were digitized onto a laboratory microcomputer (10 kHz sampling rate), processed with one of three different algorithms, then converted back to analog form, low-pass filtered (5000 Hz nominal cut-off, 104 dB/octave attenuation rate), and recorded onto digital-audio tape (DAT).

The first form of distortion was time-compression. Time-compression techniques produce speeded speech in a manner that avoids spectral distortion. The software developed for this purpose examines the entire speech waveform for epochs of between 5 and 15 msec duration. In the case of voiced segments, these epochs correspond to actual pitch periods associated with fundamental frequencies in the range between 200 and 67 Hz. All quasialternate epochs

(depending upon the time-compression ratio) are extracted from both voiced and unvoiced segments at zero-crossing points and written to a new waveform file. Several points before and after each extracted segment are then overlapped with those of adjacent segments at the juncture points in the new waveform file. A weighting function is also applied to these overlapped points so that the rise-fall time between sequential speech samples is gradual. The LP sentences were time-compressed at 30%, 40%, 50%, and 60% time-compression ratios. These time-compression ratios represent the percentage reduction in the total duration of the original sentences that were implemented in the time-compressed sentences.

The second form of distortion was reverberation, which modifies the waveform characteristics of speech. The software to create the reverberant speech was based on the image method for simulating small room acoustics (Allen & Berkley, 1979). The room model assumes a rectangular enclosure with a source-to-receiver impulse response, or transfer function, calculated using a time-domain image expansion method. The subroutine to calculate the impulse response requires as input parameters the number of impulse response points described, the source location, the receiver location, the room dimensions, and the reflection coefficients of each of the six wall surfaces. Once a simulated room impulse response has been calculated, an energy decay curve for the impulse response is used to estimate the reverberation time of a signal produced in this room. A reverberant speech sample is created by using a Fast Fourier Transform (FFT) method to convolve an anechoic (unreverberant) speech sample with the calculated impulse response. In this experiment, the LP R-SPIN stimuli were processed to simulate four reverberation times: 0.2 sec, 0.3 sec, 0.4 sec, and 0.6 sec.

The third form of distortion was interruption. Speech interruption was created digitally by multiplying the digitized speech waveform by a square wave. An algorithm was used to smooth the signal amplitude at the interruption junctures. The method effectively turned the speech signal on and off at a 50% speech-time ratio. The four rates of interruption used were 12.5/sec, 25/sec, 50/sec, and 100/sec based on pilot testing with three subjects with normal hearing who obtained recognition scores between 68 and 92% correct for this range of interruption frequencies. Moreover, this range of interruption frequencies produced scores of between approximately 60 and 90% correct among naive subjects with normal hearing in a classic study by Miller and Licklider (1950).

### Psychoacoustic Measures

The measures of auditory temporal processing were those described in another paper (Fitzgibbons & Gordon-Salant, 1992). Briefly, there were four temporal processing tasks, each measured at two test frequencies, 500 Hz and 4000 Hz, to examine expected regions of minimal and maximal hearing loss, respectively. In order to avoid potential age-related effects of response bias, discrimination thresholds were measured using a three-interval forced-choice paradigm

**TABLE 2.** Summary of speech recognition measures and psychoacoustic measures used in the experiments.

Speech recognition measures		Psychoacoustic measures
Undistorted LP SPIN	(SPIN)	Duration Discrimination 500 Hz (DD5) 4000 Hz (DD4)
Time-Compressed LP SPIN		Gap Duration Discrimination 500 Hz (GD5) 4000 Hz (GD4)
30% Time Compression	(TC30)	
40% Time Compression	(TC40)	
50% Time Compression	(TC50)	
60% Time Compression	(TC60)	
Reverberant LP SPIN		Gap Detection—Same Tones 500 Hz (SAME5) 4000 Hz (SAME4)
0.2 sec Reverb. Time	(REV2)	
0.3 sec Reverb. Time	(REV3)	
0.4 sec Reverb. Time	(REV4)	
0.6 sec Reverb. Time	(REV6)	
Interrupted LP SPIN		Gap Detection—Different Tones 500 Hz (DIFF5) 4000 Hz (DIFF4)
12.5 Interruptions/Sec	(INT12)	
25 Interruptions/Sec	(INT25)	
50 Interruptions/Sec	(INT50)	
100 Interruptions/Sec	(INT100)	

(3IFC) and an adaptive rule for stimulus change that converged on 70.7% performance level (Levitt, 1971). The subject's response was simply to select the trial interval that sounded different from the other two, regardless of experimental condition and stimulus. Stimulus presentation, event timing, and collection of responses were controlled by a laboratory computer. The first task was duration discrimination, which measured difference limens (DLs) for duration increments in tone bursts having a standard duration of 250 msec. The second task was gap duration discrimination, which measured the DL for silent intervals between tone bursts. The silent-interval DL was assessed for a 250-msec standard interval that was marked by 250-msec tone bursts of equal frequency. The third and fourth tasks were measures of gap detection, which examine the listener's ability to detect a brief silent interval between successive tone bursts. Stimuli preceding and following the gap were 250 msec in duration with 5-msec rise-fall envelopes. In the third task, the temporal gap was preceded and followed by tonal stimuli of the same frequency; in the fourth task, the temporal gap was surrounded by two different stimulus frequencies.

### Apparatus

During the speech experiments, the LP-SPIN sentences were played back on the DAT recorder/player (SONY PCM-2500A), amplified (Crown D-75), attenuated (Hewlett-Packard 350D), amplified again (Colbourn S82-24), and delivered monaurally to an Etymotic ER-3A insert earphone at 90 dB SPL. The stimuli for the psychoacoustic experiments were played from a Crescent microcomputer system, routed to a programmable attenuator system (Tucker-Davis Technologies PA3 and HBUF3), amplified (Colbourn S82-24), attenuated (Hewlett-Packard 350D), and delivered to an insert earphone (Etymotic ER-3A) at 85 dB SPL. For subjects with hearing loss, the test ear was always the better ear; for normal hearing subjects, the test ear was the preferred ear. The nontest ear was occluded with an EAR foam plug to

mimic standard monaural measurement techniques. Subjects were seated in a soundproof booth during all test procedures.

### Procedures

There were 21 experimental conditions presented to each subject. These conditions are summarized in Table 2. The speech experiments consisted of 13 different conditions corresponding to the 12 forms of distorted speech (3 types of distortion  $\times$  4 degrees of distortion) and undistorted speech. These conditions were presented in random order to the subjects, and random assignment was made of SPIN list to experimental condition. Subjects were asked to write the final test word of each sentence.

The eight psychoacoustic conditions (four tasks  $\times$  2 stimulus frequencies) were presented in randomized order to the subjects. Three measures of threshold corresponding to 70.7% correct discrimination were obtained in each condition. The mean of these measures was taken as the final threshold estimate.

The speech conditions were presented first, followed by the psychoacoustic conditions for half of the subjects in each group, and were presented in reverse order for the other half of the subjects.

The entire test procedure was completed in 12 to 15 hours, scheduled in 2-hour sessions at 1-week intervals. Subjects were reimbursed for their participation in the experiments.

## Results

### Speech Recognition Data

Mean percent correct scores and standard deviations for the undistorted LP-SPIN sentences are shown in Figure 1. Each subject group exhibited good-to-excellent speech recognition scores. However, a one-way ANOVA on arc-sine

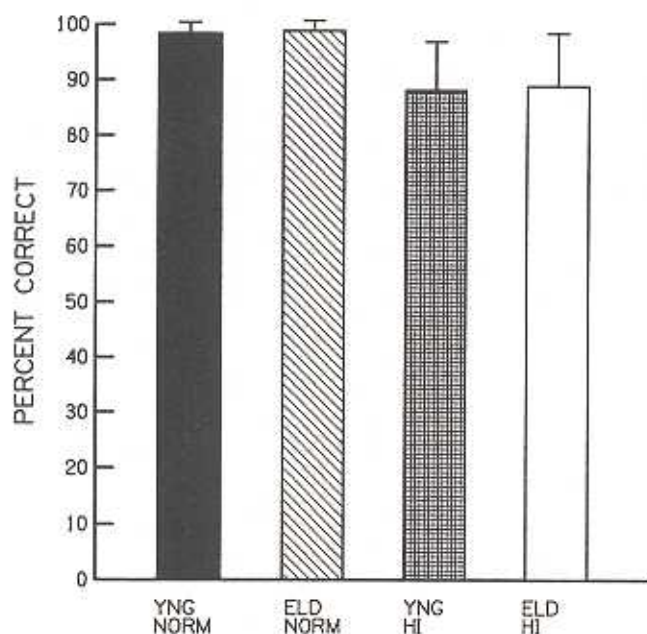


FIGURE 1. Mean percent-correct recognition scores and standard deviations of the four subject groups for the undistorted low-predictability SPIN sentences presented in quiet.

transformed scores revealed a significant effect of subject group ( $F = 7.70, p < .001$ ). Post hoc multiple comparison testing (Student-Newman-Keuls) revealed that the two groups of subjects with normal hearing performed significantly better than the two groups with hearing loss. There were no significant differences in performance between the two groups with normal hearing, indicating that small sensitivity differences between the young and elderly listeners did not play an important role in suprathreshold speech recognition performance. Tests for homogeneity of vari-

ances (Cochran's C, Bartlett-Box F, and Hartley's F-Max) failed to reach significance in this and all subsequent analyses of differences between means, indicating that the assumption of homogeneity of variances was met in these analyses.

Figure 2 presents the mean performance of the four subject groups for the time-compressed LP-SPIN stimuli. An Analysis of Covariance (ANCOVA) was conducted on the arc-sine transformed scores, to remove the systematic differences between groups revealed in undistorted speech processing. A randomized block design (Kirk, 1968) with two between-subjects factors (age and hearing status) and one within-subjects factor (degree of distortion) was applied for this and all subsequent ANCOVA's. The results of the ANCOVA on the time-compressed speech data showed a significant main effect of age ( $F = 28.45, p < .001$ ), hearing status ( $F = 132.01, p < .001$ ), and time-compression ratio ( $F = 57.60, p < .001$ ). Interaction effects were not significant.

Mean performance scores on the reverberant speech tasks are shown in Figure 3. Again, ANCOVA was performed and showed significant effects of age ( $F = 33.05, p < .001$ ), hearing status ( $F = 69.34, p < .001$ ), and reverberation time ( $F = 70.15, p < .001$ ). There were no significant interactions.

Finally, the interrupted speech data for the four subject groups are shown in Figure 4. ANCOVA of these data revealed a significant main effect of age ( $F = 8.85, p < .01$ ) and hearing status ( $F = 58.69, p < .01$ ). A significance level of .01 was selected per experiment. Using this criterion, the main effect of interruption frequency ( $F = 3.30, p < .05$ ) was not considered significant. It should be noted, however, that the performance of the young listeners with normal hearing is comparable to that reported previously (Miller & Licklider, 1950). The present analysis also failed to reveal any significant interactions among the variables evaluated.

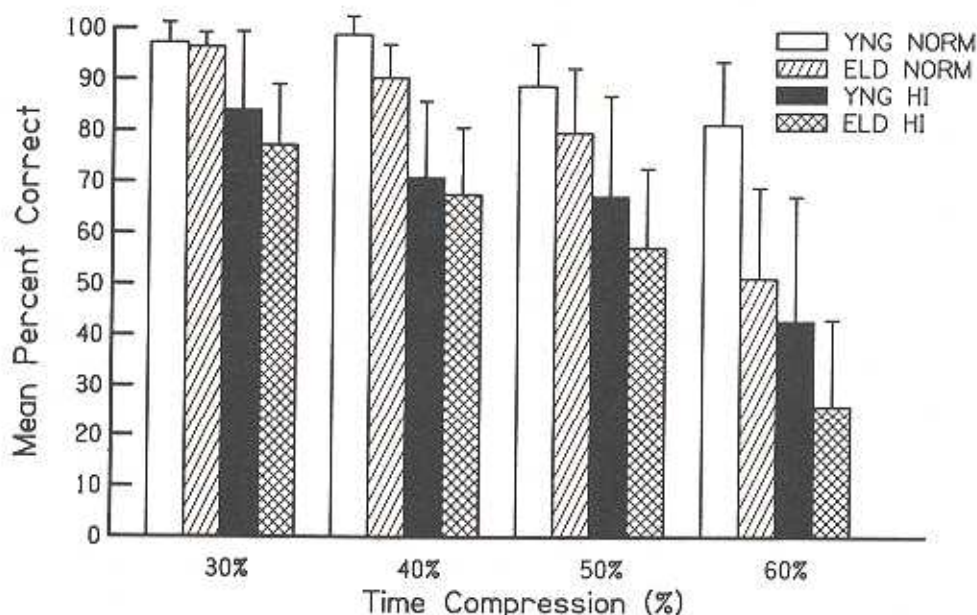


FIGURE 2. Mean percent-correct recognition scores and standard deviations of the four subject groups for the time-compressed low-predictability SPIN sentences, at four time-compression ratios.

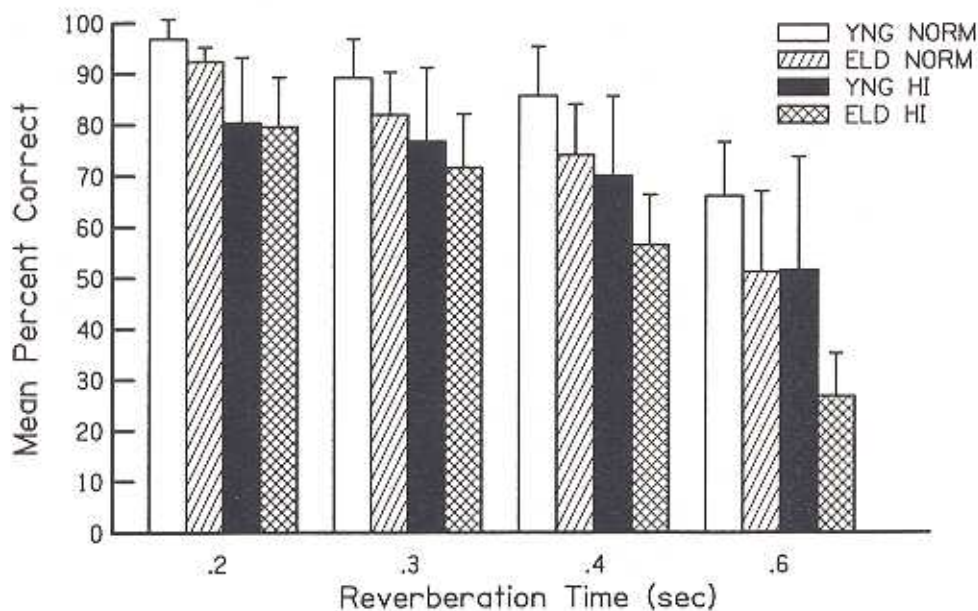


FIGURE 3. Mean percent-correct recognition scores and standard deviations of the four subject groups for the reverberant low-predictability SPIN sentences, at four reverberation times.

### ***Relations Between Psychoacoustic and Speech Measures***

The goal of the second series of analyses was to identify the predictive relationship between the psychoacoustic measures and the speech recognition measures. To that end, a canonical correlation analysis (Hotelling, 1935, 1936) was conducted. It examined the linear relationships between a set of predictor variables and a set of criterial variables. The only restriction in applying canonical correlation is that there be a minimum of two variables in each set (Tucker & Chase, 1980). A second general

assumption is that data for all variables be collected on each subject. These assumptions were met in the present data analysis. The canonical correlation technique finds several linear combinations of the predictor variables and the same number of linear combinations of the criterial variables, such that the correlations between the two sets are maximized. The linear combinations are called the canonical variables, and the correlations between corresponding pairs of canonical variables are called canonical correlations (SAS, 1989). The coefficients of the linear combinations are canonical coefficients or canonical weights. Each canonical variable is uncorrelated with all the other canon-

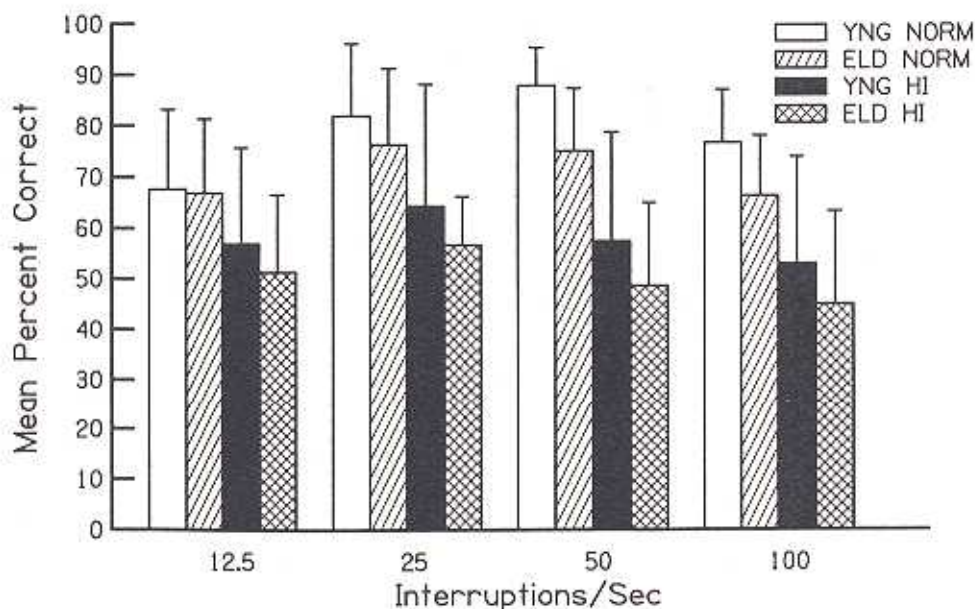


FIGURE 4. Mean percent-correct recognition scores and standard deviations of the four subject groups for the interrupted low-predictability SPIN sentences, at four interruption ratios.

ical variables of either set, except for the one corresponding canonical variable of the opposite set.

The SAS CANCERR procedure was conducted in the present analysis. Initially, the full set of predictor variables included the eight scores from the psychoacoustic measures, individual pure tone thresholds from 250 to 4000 Hz, three calculations of pure tone average, and age. The threshold measures (individual frequencies and averages) were included because of previous research demonstrating that the most significant predictor variable was hearing sensitivity (Humes & Christopherson, 1991; van Rooij & Plomp, 1990). The full set of criterion variables was the 13 scores from all speech recognition conditions (undistorted, time-compressed, interrupted, and reverberant speech). The raw scores were transformed to z-scores, and a multiple correlation matrix was derived. The correlation matrix showed that there was high multicollinearity ( $r > .70$ ) between numerous pairs of variables within a set (criterial or predictor variables). One strategy for reducing high multicollinearity is to delete one or more of the highly redundant variables (Tucker & Chase, 1980). Applying this strategy, individual variables were omitted, one at a time, until the correlation matrix indicated that each variable was reasonably unique.

The final set of predictor variables submitted to the CANCERR procedure was PTA2 (average of thresholds at 1000 Hz, 2000 Hz, and 4000 Hz), age, duration discrimination at 500 Hz and 4000 Hz (DD5 and DD4), gap duration discrimination at 500 Hz and 4000 Hz (GD5 and GD4), gap detection for same tones at 4000 Hz (SAME4), and gap detection for different tones at 4000 Hz (DIFF4). The criterion variables in the final set included undistorted SPIN scores (SPIN), 40% time-compressed SPIN scores (TC40), 12.5% interrupted SPIN scores (INT12), and reverberant SPIN scores at 0.2 sec and 0.6 sec (REV2 and REV6).

A summary of the overall significance tests for the canonical correlation analysis is shown in Table 3. The first canonical correlation is .907, which represents the correlation between the linear combination of the dependent variables and the linear combination of the independent variables. This value is substantially larger than any simple correlation between a predictor variable and a criterion variable. The probability level for the null hypothesis that the first canonical correlation is 0 in the population is less than .0001. Thus, the null hypothesis can be rejected and the conclusion drawn that the correlation is significant. A second canonical correlation is derived such that the linear combination of predictor variables and the linear combination of dependent variables are each uncorrelated with those derived for the first canonical correlation. In addition, the new linear combinations themselves are maximally correlated. The second canonical correlation is necessarily less than the first canonical correlation. Table 3 shows that the second canonical correlation is .807, which also is significant ( $F = 2.3992$ ,  $p < .0008$ ). Thus, the first two canonical correlations were significant and warrant further interpretation.

Historically, researchers have examined the size and configuration of the weights in each set of canonical variables. However, Tucker and Chase (1980) argue that examining the canonical component loadings (the correlations of the original data variables with the derived canonical vari-

TABLE 3. Summary of canonical correlation analysis.

	Eigenvalue	Canonical $r$	Approx. $F$	$df$	$p$
Variate					
1	4.64	.907	3.7875	40	<.0001
2	1.87	.807	2.3992	40	<.0008

ates) is preferred. Table 4 shows this canonical structure of the correlations between the predictor and criterial variables and the first two canonical variables. The succeeding interpretation of these correlations follows the general rule that a correlation  $> .5$  is significant.

Inspection of Table 4 indicates that the first canonical variable (CV1) is highly correlated with the predictor variable PTA2 ( $-.92$ ). For the second canonical variable (CV2), high correlations are shown with the predictor variables GD5 (.72), GD4 (.65), and age (.63). The correlations shown in Table 4 also indicate significant relationships between the first canonical variable and the criterial variables TC40 (.91), SPIN (.83), and REV2 (.71). The second canonical variable shows significant negative correlations with REV6 scores ( $-.6648$ ).

Two conclusions can be drawn from this canonical correlation analysis. First, the presence of high-frequency hearing loss is associated with decreased scores on undistorted and nearly all temporally distorted speech tests. This effect is most prominent for time-compressed speech (40% TC ratio). The second finding is that age and gap duration discrimination (at both 500 Hz and 4000 Hz) contribute significantly to recognition of reverberant speech. Specifically, advanced age and increased gap discrimination thresholds (reflecting poorer performance) are related to reduced reverberant speech scores.

The CANCERR procedure also calculates a canonical redundancy matrix, which assesses the degree to which the sets of criterial and predictor variables overlap. Thus, it assesses the amount of variance in the dependent variables

TABLE 4. Summary of the canonical structure analysis between the set of predictor variables (age, hearing loss, and temporal processing measures) and the set of criterial variables (speech measures).

Variables	Structure coefficients	
	CV1	CV2
Predictor		
AGE	-.016	0.63
PTA2	-.92	0.22
DD5	0.10	0.21
DD4	-.07	0.14
GD5	0.21	0.72
GD4	0.41	0.65
SAME4	0.14	0.00
DIFF4	-.19	0.19
Criterial		
SPIN	0.83	0.06
TC40	0.91	-.027
INT12	0.48	0.10
REV2	0.71	-.49
REV6	0.42	-.66

that are predictable from a knowledge of the independent variables. Based on the first two canonical variables, 67% of the proportion of variance of the dependent variables can be explained by their own canonical variables and 52% of the proportion of this variance can be explained by the opposite canonical variables. Thus, approximately 36% of the variance in these speech recognition scores was unexplained by the present analysis.

## Discussion

Analyses of speech recognition performance indicated that there were significant effects of age, independent of the effects of hearing impairment, on the reverberant, time-compressed, and interrupted speech measures. Thus, elderly subjects performed more poorly than their younger counterparts on these three temporally distorted speech tasks, even when possible performance differences in undistorted speech recognition scores were covaried out. These results suggest that age-related factors other than peripheral sensitivity loss contribute to the speech recognition deficits of elderly listeners.

Few previous studies using temporally distorted speech signals and elderly listeners employed careful controls of peripheral hearing sensitivity in young and elderly subjects, rendering comparisons between studies difficult. However, one previous study (Humes & Christopherson, 1991) used a filtered + reverberant nonsense syllable speech task; they found no differences in recognition scores between young noise-masked listeners and elderly listeners with hearing loss (aged 65–75 years), although an older group of elderly listeners (76–86 years) performed more poorly than the other two groups. Methodological differences between the current study and that of Humes and Christopherson (1991) could account for the divergent findings, including differences in speech materials, reverberation times, presentation levels, and number of distortions imposed on the speech signal.

Recognition of undistorted LP-SPIN sentences was affected by hearing loss but not by age. This finding is consistent with numerous previous reports showing that age, *per se*, is not associated with diminished speech recognition performance in undistorted, quiet listening conditions (e.g., Gordon-Salant, 1987). Listeners with hearing loss exhibited recognition deficits for LP-SPIN stimuli, regardless of age, even at high presentation levels. This indicates that the LP-SPIN sentences, even in nonstandard quiet conditions, are sensitive for revealing speech recognition difficulties of listeners with hearing loss.

Similarly, hearing loss was a significant variable affecting recognition of the three forms of temporally distorted speech, regardless of the degree of distortion. This finding is in agreement with earlier reports of recognition performance of listeners with hearing loss for monosyllabic words distorted by time-compression (Grimes et al., 1984), reverberation (Nabelek & Mason, 1981), and interruption (Korsan-Bengsten, 1973). The current results extend the effect of hearing impairment to difficulties with temporally distorted sentence-length materials at high presentation levels. The implication is that efforts to maximize recognition performance of listen-

ers with hearing loss by increasing the stimulus level are not successful with temporally distorted speech.

Assessment of relations between speech and nonspeech measures showed that high frequency hearing sensitivity was the most important variable contributing to undistorted LP-SPIN scores, time-compressed scores at 40% time-compression, and reverberant speech scores at 0.2 sec RT. The strong correlation with hearing loss validates the ANCOVA results, and also is in agreement with previous studies that examined relations between nonspeech and speech measures (Humes & Christopherson, 1991; van Rooij & Plomp, 1990).

The canonical correlation analysis also revealed a second set of canonical variables, age and gap duration discrimination, that was highly (and negatively) correlated with the reverberant speech measure at 0.6 sec RT. Increased gap duration discrimination thresholds were associated with decreased reverberant speech scores. Similarly, older age was associated with reduced reverberant speech scores. The gap duration discrimination threshold reflects a listener's ability to detect changes in the duration of silent intervals between stimuli. This task may be analogous to detecting changes in silent intervals between words in a sentence. Because reverberation effectively alters the modulation characteristics of the speech waveform, pauses between words in the sentence as well as individual phonemes in the sentence may become difficult to resolve. The present findings suggest that listeners who have difficulty detecting changes in the duration of silent intervals will also have difficulty resolving the reduced temporal fluctuations in the reverberant speech waveform.

The second canonical correlation failed to reveal a correlation between age and time-compressed speech, and between age and interrupted speech, as might have been expected on the basis of the ANCOVA results. This outcome could be attributed in part to the specific time-compression ratio and interruption rate selected for the analysis, because the age effect appears to be relatively minor at the 40% time-compression ratio (see Figure 2) and 12.5/sec interruption frequency (see Figure 4). Across the entire range of time-compression ratios and interruption frequencies evaluated, however, the main ANCOVA showed that age was a significant factor.

The significant relationship between age and reverberant speech scores confirms that there is something associated specifically with aging that reduces the ability to process temporally distorted speech signals. The current results also indicate that this aging effect is independent of changes in hearing sensitivity and gap duration discrimination (because each canonical variable is orthogonal to the other canonical variables). Thus, there is a distinct aging factor that appears to affect performance beyond the hearing impairment and peripheral temporal processing factors. The source of the aging factor is still unresolved, although various researchers have implicated age-related changes in central auditory processing (Jerger, Jerger, Oliver, & Pirozzolo, 1989) or cognitive decline (van Rooij & Plomp, 1990). Further research aimed at elucidating the role of higher-order temporal processing factors in relation to age-related speech recognition performance decline is warranted.

Overall, the comparisons of mean performance and the correlational analyses converge on two main points: Hearing loss and age contribute independently to recognition deficits for reverberant speech and time-compressed speech. However, it also appears that at least some suprathreshold temporal processing measures (gap duration discrimination) contribute to the ability to recognize reverberant speech.

Performance by all subject groups deteriorated with increasing degrees of distortion on the time-compressed and reverberant speech tasks. Combining the three main effects (age, hearing impairment, degree of distortion) suggests that elderly listeners with hearing loss may be at a severe disadvantage when they are listening in everyday reverberant environments or to rapid speech. If methods to alleviate the negative impact of hearing loss through amplification, rehabilitation, and environmental design are to be successful, they must take into consideration these particular contributing factors to speech recognition deficits in elderly people.

## Acknowledgments

This research was supported by a grant from the National Institute on Aging (R01-AG09191). The authors are grateful to Timothy Bunnell of the Alfred I. Dupont Institute, Jont Allen of Bell Laboratories, and James Whitney (formerly of the University of Maryland) for providing the software for the time-compression, reverberation, and interruption simulations, respectively. We also thank Robert Bilger for supplying the original, unmodified R-SPIN tapes used in the experiments. Finally, we express our appreciation to Linda Carr-Kraft, Susan Phillips, and Sarah Hargus for their assistance in the stimulus preparation, data collection, and data analysis stages of the experiment.

## References

- Allen, J., & Berkley, D. (1979). Method for simulating small-room acoustics. *Journal of the Acoustical Society of America*, 65, 43-50.
- American National Standards Institute (1989). *Specification for audiometers* (ANSI S3.6-1989). New York: ANSI.
- Bergman, M. (1971). Hearing and aging. *Audiology*, 10, 164-171.
- Bergman, M., Blumenfeld, V., Cascardo, D., Dash, B., Levitt, H., & Margules, M. (1976). Age-related decrement in hearing for speech: Sampling and longitudinal studies. *Journal of Gerontology*, 31, 533-538.
- Bilger, R.C., Nuetzel, J.M., Rabinowitz, W.M., & Rzezchowski, 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.
- Blumenfeld, V.G., Bergman, M., & Millner, E. (1969). Speech discrimination in an aging population. *Journal of Speech and Hearing Research*, 12, 210-217.
- Dreschler, W.A., & Plomp, R. (1985). Relation between psychophysical data and speech perception for hearing-impaired subjects. II. *Journal of the Acoustical Society of America*, 78, 1261-1270.
- Dubno, J.R., Dirks, D.D., & Morgan, D.E. (1984). Effects of age and mild hearing loss on speech recognition in noise. *Journal of the Acoustical Society of America*, 76, 87-96.
- Findlay, R.C., & Denenberg, L.J. (1977). Effects of subtle mid-frequency auditory dysfunction upon speech discrimination in noise. *Audiology*, 16, 252-259.
- Fitzgibbons, P.J., & Gordon-Salant, S. (1992). *Age effects on measures of auditory temporal sensitivity*. Unpublished manuscript.
- Goetzinger, C., Proud, G., Dirks, D., & Embrey, J. (1961). A study of hearing in advanced age. *Archives of Otolaryngology*, 73, 662-674.
- Gordon-Salant, S. (1987). Age-related differences in speech recognition performance as a function of test format and paradigm. *Ear and Hearing*, 8, 277-282.
- Grimes, A., Mueller, G., & Williams, D. (1984). Clinical considerations in the use of time-compressed speech. *Ear and Hearing*, 5, 114-117.
- Harris, J.D., Haines, H.L., & Myers, C.K. (1963). The importance of hearing at 3kc for understanding speeded speech. *Laryngoscope*, 70, 131-146.
- Hotelling, H. (1935). The most predictable criterion. *Journal of Educational Psychology*, 26, 139-142.
- Hotelling, H. (1936). Relations between two sets of variables. *Biometrika*, 28, 321-377.
- Humes, L.E., & Christopherson, L. (1991). Speech identification difficulties of hearing-impaired elderly persons: The contributions of auditory processing deficits. *Journal of Speech and Hearing Research*, 34, 686-693.
- Irwin, R.J., & McCauley, S.F. (1987). Relations among temporal acuity, hearing loss, and the perception of speech distorted by noise and reverberation. *Journal of the Acoustical Society of America*, 81, 1557-1565.
- Jerger, J., Jerger, S., Oliver, T., & Pirozzolo, F. (1989). Speech understanding in the elderly. *Ear and Hearing*, 10, 79-89.
- Jerger, J., Jerger, S., & Pirozzolo, F. (1991). Correlational analysis of speech audiometric scores, hearing loss, age, and cognitive abilities in the elderly. *Ear and Hearing*, 12, 103-109.
- Jokinen, K. (1973). Presbycusis VI. Masking of speech. *Acta Otolaryngologica*, 76, 426-430.
- Kirk, R.E. (1968). *Experimental design: Procedures for the behavioral sciences*. Belmont: Brooks/Cole Publishing Co.
- Konkle, D., Beasley, D., & Bess, F. (1977). Intelligibility of time-altered speech in relation to chronological aging. *Journal of Speech and Hearing Research*, 20, 108-115.
- Korsan-Bengsten, M. (1973). Distorted speech audiometry: A methodological and clinical study. *Acta Otolaryngologica Supplement*, 310, 7-75.
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49, 467-477.
- Miller, G.A., & Licklider, J.C.R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America*, 22, 167-175.
- Moore, B., Peters, R., & Glasberg, B. (1992). Detection of temporal gaps in sinusoids by elderly subjects with and without hearing loss. *Journal of the Acoustical Society of America*, 92, 1923-1932.
- Nabelek, A., & Mason, D. (1981). Effect of noise and reverberation on binaural and monaural word identification by subjects with various audiograms. *Journal of Speech and Hearing Research*, 24, 375-383.
- Nabelek, A., & Robinson, P.K. (1982). Monaural and binaural speech perception in reverberation for listeners of various ages. *Journal of the Acoustical Society of America*, 71, 1242-1248.
- 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.
- Salthouse, T.A. (1982). *Adult cognition*. New York: Springer-Verlag.
- Salthouse, T.A. (1985). *A theory of cognitive aging*. New York: North-Holland.
- SAS Institute, Inc. (1989). *SAS/STAT user's guide, Version 6, 4th Ed., Vol. 1*. Cary, NC: SAS Institute Inc.
- Tillman, T., & Carhart, R. (1966). *An expanded test for speech discrimination utilizing CNC monosyllabic words* (Northwestern University Auditory Test No. 6; USAF School of Aerospace Medicine Technical Report). Brooks Air Force Base, TX.
- Townsend, T.H., & Bess, F.H. (1980). Effects of age and sensorineural hearing loss on word recognition. *Scandinavian Audiology*, 9, 245-248.

- Tucker, R.K., & Chase, L.J. (1980). Canonical correlation. In P.R. Monge, & J.N. Cappella, *Multivariate techniques in human communication research*. New York: Academic Press.
- Tyler, R., Summerfield, Q., Wood, E.J., & Fernandes, M.A. (1982). Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 72, 740-752.
- van Rooij, J.C.G.M., & Plomp, R. (1990). Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses. *Journal of the Acoustical Society of America*, 88, 2611-2624.
- Wingfield, A., Poon, L., Lombardi, L., & Lowe, D. (1985). Speech processing in normal aging: Effects of speech rate, linguistic structure, and processing time. *Journal of Gerontology*, 40, 579-585.
- Yanz, J.L., & Anderson, S.M. (1984). Comparison of speech perception skills in young and old listeners. *Ear and Hearing*, 5, 134-137.

---

Received January 7, 1993

Accepted May 19, 1993

Contact author: Sandra Gordon-Salant, University of Maryland at College Park, Department of Hearing and Speech Sciences, College Park, MD 20742.