Published in final edited form as:

J Acoust Soc Am. 2005 August ; 118(2): 1089–1100.

Decline of speech understanding and auditory thresholds in the elderly

Pierre L. Divenyi^a

Speech and Hearing Research Veterans Affairs Medical Center and East Bay Institute for Research and Education Martinez, California 94553

Philip B. Stark

Department of Statistics University of California Berkeley, California 94720-3860

Kara M. Haupt

Speech and Hearing Research Veterans Affairs Medical Center Martinez, California 94553

Abstract

A group of 29 elderly subjects between 60.0 and 83.7 years of age at the beginning of the study, and whose hearing loss was not greater than moderate, was tested twice, an average of 5.27 years apart. The tests measured pure-tone thresholds, word recognition in quiet, and a set of tests of understanding speech with various types of distortion (low-pass filtering, time compression) or interference (single speaker, babble noise, reverberation). Performance declined consistently and significantly between the two testing phases. In addition, the variability of speech understanding measures increased significantly between testing phases, though the variability of audiometric measurements did not. A right-ear superiority was observed but this lateral asymmetry did not increase between testing phases. Using a group of young subjects with normal hearing to establish a baseline, we found that the decline of speech understanding measures accelerated significantly relative to the decline in audiometric measures in the seventh to ninth decades of life. On the assumption that speech understanding depends linearly on age and audiometric variables, there is evidence that this linear relationship changes with age, suggesting that not only the accuracy but also the nature of speech understanding evolves as we mature.

I. INTRODUCTION

From early adulthood on, but especially after age 60, auditory sensitivity and other aspects of hearing gradually deteriorate. More troubling for communication, and thus for quality of life, is the progressive degradation of speech understanding, especially when there is noise such as background speech or reverberation. The decline of auditory communication ability is interesting as basic science and for its clinical implications: understanding this loss would shed light on the relationship between hearing and speech comprehension and could lead to the development of devices to alleviate the deficit.

The nature, degree, and rate of hearing deterioration with age have been studied both diachronically, by testing a group of subjects longitudinally as they age, and synchronically, by comparing subjects of different ages. Longitudinal study is more direct and has fewer inherent ambiguities, but it is costly, difficult, and fraught with unpredictable obstacles, such as dropouts or illnesses affecting auditory functions. Cross-sectional study of subjects of different ages is easier and less expensive; however, it is more vulnerable to confounding.

^aPortions of the findings were reported at the 139th meeting of the Acoustical Society of America, Atlanta, GA, on May 31, 2000. Author's electronic mail: pdivenyi@ebire.org

Interestingly, differences between findings reported in longitudinal and cross-sectional studies with large sample sizes are minor: they present quite consistent comprehensive pictures of auditory aging Brant and Fozard (1990). The gradual presbycusis (i.e., sloping, high-frequency) loss of hearing sensitivity with age has been studied extensively. Pure-tone thresholds tend to increase as individuals age: longitudinal and cross-sectional studies put the rate of decay at about 5.5 to 9 dB/decade for the *better* ear, depending on frequency (Davis, Osti & Parving, 1990, Gates, Cooper, Kannel & Miller, 1990, Osti & Parving, 1991); the worse ear often deteriorates almost 50% faster. Although one longitudinal study of changes in hearing level covering a 23-year span (the Baltimore Longitudinal Study on Aging) estimated the rate of decline to be almost twice that found in other studies (4.5 to 14.7 dB/decade, Pearson, Morrell, Gordon-Salant, Brant, Metter, Klein & Fozard, 1995, Morrell, Gordon-Salant, Pearson, Brant & Fozard, 1996), the discrepancy may be due in part to the different ways the rate of decline was calculated. Several studies also observed that decline in men is faster in the younger advanced years and slower later, whereas in women, the rate of decline is slow at first and accelerates later—with the result that by the eighth or ninth decade of life, almost no gender differences (Gates *et al.*, 1990, Pearson *et al.*, 1995, Morrell *et al.*, 1996) remain.

As mentioned above, auditory thresholds in the two ears are seldom exactly equal. The ear with the higher thresholds is often the left, and its relative disadvantage increases with age (Gates *et al.*, 1990, Gates & Cooper, 1991). There is also typically a left-ear disadvantage in processing some suprathreshold speech presented in interference. In particular, perception of left-ear targets of dichotic sentences deteriorates with age relative to perception of right-ear targets (Jerger, Stach, Johnson, Loisel & Jerger, 1990, Jerger & Jordan, 1992), even in individuals with little or no change in their hearing sensitivity over a period of several years (Stach, Jerger & Fleming, 1985). In contrast, the right-ear advantage for brief, dichotically presented nonsense syllables, digits, or spondees does not appear to change (Gelfand, Hoffman, Waltzman & Piper, 1980, Martini, Bovo, Agnoletto, Da Col, Drusian, Liddeo & Morra, 1988). These reports suggest that a right-ear advantage increases both for peripheral hearing sensitivity and for speech tasks that rely on centrally mediated processes. This increase of asymmetry may be due to a breakdown of the integration of binaural information, possibly caused by gradual demyelination of the callosal pathway necessary for interhemispheric transfer (Jerger, Silman, Lew & Chmiel, 1993, Chmiel & Jerger, 1996, Chmiel, Jerger, Murphy, Pirozzolo & Tooley-Young, 1997).

Decline of speech understanding after age 60 also has been investigated extensively. Generally, understanding of undistorted speech presented in quiet does not change significantly (Blumenfeld, Bergman & Millner, 1969, Gelfand, Piper & Silman, 1985, van Rooij & Plomp, 1990, Tun, 1998), at least until the eighth decade of life (Bergman, Blumenfeld, Cascardo, Dash, Levitt & Margulies, 1976, Pedersen, Rosenhall & Møller, 1991). In contrast, understanding of speech presented in interference deteriorates steadily and sometimes dramatically. For example, understanding of speech with added speech-spectrum random noise or multi-talker babble deteriorates with age, especially from the seventh decade on (Gelfand, Piper & Silman, 1986, Dubno, Lee, Matthews & Mills, 1997); reverberation increasingly affects speech understanding as listeners age (Duquesnoy & Plomp, 1980, Nabelek & Robinson, 1982); thresholds for word, consonant, and even vowel identification increase with age (van Rooij & Plomp, 1990); and age predicts understanding of speech in the presence of reverberation (Humes & Christopherson, 1991, Divenyi & Haupt, 1997a) or an interfering ipsilateral or contralateral sentence (Jerger & Hayes, 1977, Divenyi & Haupt, 1997a). Speech understanding by older listeners is also affected by time compression (Konkle, Beasley & Bess, 1977, Tun, 1998), periodic interruption (Bergman *et al.*, 1976), and low-pass filtering (Cheesman, Hepburn, Armitage & Marshall, 1995)—especially when the filtering is paired with reverberation (Humes & Christopherson, 1991).

Can the decline of speech understanding be attributed to presbycusis threshold shifts? Some studies suggest so for some aspects of speech perception (e.g., Duquesnoy & Plomp, 1980, Humes & Roberts, 1990, Humes, 1991, Dubno *et al.*, 1997), but other studies indicate that deterioration in speech understanding occurs *in addition* to deterioration in hearing sensitivity (Jerger & Hayes, 1977, Bergman, 1980, Humes & Christopherson, 1991, Pedersen, Rosenhall & Møller, 1991, Divenyi & Haupt, 1997a, c, Jerger & Chmiel, 1997). That is, changes in speech understanding in aging, although influenced by hearing sensitivity, also include components separate from the presbycusis process.

Hearing and understanding speech in interference or distortion deteriorate later in life. But how rapidly? Do hearing and speech understanding decline at the same rate? Does the decline in hearing account for the decline in understanding? To answer these questions requires tracking changes in individuals' hearing sensitivity and speech understanding over time. Such changes provide clues to relationships among different auditory functions; in turn, these relationships might hint at the underlying causes.

This article presents such a longitudinal study. Although it spans only about 5 years and includes only 29 subjects, according to a recent survey (Divenyi & Simon, 1999) no comparable study has been conducted before. The present investigation focused on how speech understanding under unfavorable conditions declines in elderly individuals whose auditory performance is initially better than average for the general population at their age. Limiting the study to individuals with relatively good initial hearing reduces confounding from presbycusis and makes it easier to separate the changes in the two auditory functions.

II. METHODS

A. Subjects

Twenty-nine subjects, 21 females and eight males with above-average results on a large battery of auditory performance tests (Divenyi & Haupt, 1997a), were selected from an original group of 100 elderly individuals. The subjects, whose ages ranged from 60.0 to 83.7 years (mean: 69.6 years, SD: 6.1 years) at the time of initial testing, were re-tested using the same battery 4.0 to 8.5 years after their first test. Ages ranged from 65.2 to 88.6 years (mean: 74.9 years, SD: 5.8 years) at retesting. The mean speech recognition threshold (*SRT*) of the group was 13.6 dB HL (SD: 8.62 dB HL) at initial testing and 15.0 dB HL (SD: 10.85 dB HL) at retesting. Pure-tone thresholds were measured at 0.5, 1, 2, 4, and 8 kHz (Figure 1). At each frequency, in both testing phases, and for all subjects, the thresholds were symmetrical within 15 dB in 280 of 290 right-left threshold pairs, and always within 25 dB. Figure 1 shows that the hearing loss of our group is presbycusis, typical of persons in their 7th to 9th decades, and does not exceed what is considered mild-to-moderate. Air-conduction thresholds matched bone-conduction thresholds (i.e., any loss was of the sensorineural type); tympanometry was normal; acoustic reflexes were present at 0.5, 1, and 2 kHz; there was no decay at either 0.5 or 1 kHz; and word recognition in quiet was 80 percent or better (at 45 dB sensation level re/ *SRT*). At the time of the initial testing, the subjects' self-assessed hearing was "good" or "excellent." Our inquiry, both at the initial screening and at the time of retesting, revealed no history of unusual noise exposure. None of the subjects had ever complained of tinnitus or vertigo or worn a hearing aid. All were in good general health and were native English speakers – a requirement necessary for the interpretation of speech understanding results.

To estimate the rate of change in the auditory performance of the subjects prior to their first test, we compared their results to those of 11 healthy, normal-hearing individuals with an average age of 22.18 years (SD: 3.15 years) and no prior experience as listeners. Their average auditory performance was used as an estimate of the auditory performance of the elderly subjects 40 to 60 years before the study started.

B. Apparatus and test battery

Audiological status of the subjects was assessed using acoustic immittance, pure-tone, and speech audiometry. Audiometric and psychoacoustic tests were conducted in a sound-attenuated testing booth. Audiometric tests were performed using an Interacoustics AC30 clinical audiometer and a set of Telephonic TDH-39 earphones with MX 41/AR type cushions.

Speech understanding was measured on the word and sentence levels for spectrally and temporally distorted speech as well as for speech presented with various types of interference. Some tests were performed under more than one condition, producing several performance measures; some performance measures were calculated by combining results for several test conditions. The seven tests described below gave a total of 14 measures, each obtained separately for either the right and left ears or the right and left hemifields (for the *SPIN* tests presented in virtual free field). Assessing performance in the right and left ears or hemifields allowed us to estimate changes in lateral asymmetry. The acronyms in bold type are used throughout the rest of the paper.

1. Pure-tone thresholds at 0.5, 1, 2, 4, and 8 kHz, leading to three measures: From the 0.5, 1, 2, and 4-kHz thresholds, we calculated **PTA4**, the average pure-tone threshold of the four lowest frequencies, as well as **PTSLP**, the difference between the 0.5 and 4-kHz thresholds. We used the 8-kHz threshold by itself (**PT8k**).
2. Speech Recognition Threshold (**SRT**) for spondee words.
3. Low-pass filtered speech (see Lynn & Gilroy, 1972, 1975): listeners were asked to attend to and repeat NU-6 words low-pass filtered at 750 Hz (**LP750**), presented at 50 dB SL (re/ the average pure-tone threshold at 0.5, 1, and 2 kHz).
4. Time-Compressed speech (see Beasley, Schwimmer & Rintelmann, 1972): listeners were asked to repeat words from the W-22 list compressed by 60 percent (**TC60**), presented at 50 dB SL (re/ **SRT**).¹
5. Speech understanding in reverberation, using the Modified Rhyme Reverberation Test (see Nabelek & Robinson, 1982) at reverberation times of 0.45, 0.85, and 1.25 seconds or without reverberation, presented at 50 dB SL (re/ **SRT**). Listeners were asked to identify which of six similar-sounding words (differing only with respect to the initial or final consonant) they heard. **RT75** is the estimated reverberation time corresponding to 75 percent correct identification.²
6. Competing Sentence Test (Willeford, 1985): pairs of sentences were presented simultaneously, with one sentence in each pair spoken by a female and the other sentence spoken by a male. In the monaural condition, both sentences in each pair were presented in the same ear at 45 dB SL (re/ **PTA4**) for the test sentences and 50 dB SL for the non-test sentences (**CSTI**), i.e., a -5 dB signal-to-noise ratio (S/N). Listeners were asked to repeat the sentence said by the female speaker. In the binaural condition (**BCST**), one sentence was presented in each ear at 50 dB SL (re/ **PTA4**), i.e., 0 dB S/N. Listeners were asked to repeat both sentences.
7. Speech Perception in Noise Test (*SPIN*, originally developed by Kalikow, Stevens, & Elliott, [1977], revised by Bilger et al. [1984], the “Illinois SPIN test”), presented either monaurally or binaurally. Listeners were asked to repeat the final word of a sentence spoken by a male speaker. The last word was either easy to predict (a “high-probability” item) or hard to predict (a “low-probability” item) from the sentence context. High and low-probability items were scored separately. In the monaural

¹The compression of the words was accomplished using the computer program ProAudio9, by the Cakewalk software company.

²Because performance decreased monotonically with increasing reverberation time, we used linear regression to interpolate or extrapolate to find **RT75**.

condition, the target was presented at 50 dB SL (re/ babble threshold) and the speech-to-babble ratio (S/B) was 4 dB. The monaural results yielded **SPMon**, the average of the high- and low-probability scores, and **SPMonCEf**, the difference between the scores for high-probability and low-probability items, a measure of the effect of context. In the binaural conditions, the sentences were presented in virtual free field at 45° right or left, flanked by two independent four-talker (two female, two male) babble sources on either side of the target speaker, with all speakers positioned on the periphery of a 2-m radius circle. There were three spatial *SPIN* conditions differing in the angular separation between adjacent babble speakers; this separation was 22.5°, 45°, or 72°, resulting in a total azimuthal span of 90°, 180°, and 360°, respectively.³ The presentation level at the nearer ear was 50 dB SL for both the target sentences and the babble. Three measures were derived from the spatial *SPIN* data: (1) The overall performance measure, **SPSpat**, is the average of high- and low-probability item scores across the three angular separations. (2) The context effect, **SPSpatCEf**, is the difference between scores for high-probability and low-probability items averaged across the three conditions. (3) The spatial separation effect, **SPSpatEf**, is the least-squares estimate of the slope of performance as a function of azimuthal separation, weighted by the logarithm of the overall spatial performance. This last measure was included because speech understanding in babble was seen to depend on the azimuthal separation between target and interference for individuals without significant hearing loss, whether young (Shinn-Cunningham, Schickler, Kopco & Litovsky, 2001) or elderly (Gelfand, Ross & Miller, 1988, Divenyi & Haupt, 1997b). The weighting prevents masking of the effect by consistently poor or consistently excellent performance.

The battery thus produced four audiometric measures (**PTA4**, **PTSLP**, **PT8k**, and **SRT**) and ten measures of speech understanding: two of temporally or spectrally distorted speech (**TC60** and **LP750**) and eight of speech in interference (**RT15**, **SPMon**, **SPMonCEf**, **SPSpat**, **SPSpatCEf**, **SPSpatEf**, **BCST**, and **CSTI**). Testing procedures and presentation levels were those recommended by the developers of each test except for the *SPIN* test.⁴ For tests presented to both ears or in both hemifields, the first ear tested was counterbalanced across subjects and the order of test presentation was pseudo-random. Procedures for the initial and second tests were identical. In each phase, subjects were tested for a total of eight to ten hours over an average span of three weeks. Each testing session lasted approximately 90 minutes. Breaks were provided as needed. Units of the performance measures differed by task, but larger values correspond to better performance—thus, a positive difference between the first and second phases shows deterioration in performance. Percent correct measures were logit-transformed to emphasize differences at the high and low ends of the range. *SPIN* test results were logit-transformed before any further data manipulation, such as averaging and/or differencing. The remaining measures represent threshold values, expressed in dB or in seconds (for **RT75**).

C. Statistical methods

The results were analyzed using nonparametric methods. In some cases, parametric tests were also performed for comparison. We eschew common parametric methods, such as ANOVA and the t-test, because their assumptions are difficult to justify in the present study and, arguably, in audiological studies in general. Significance levels for common parametric tests are computed assuming that the data are independent samples from normally distributed populations, which is questionable for several reasons: (1) Test and retest data for the same measure for the same individual may be dependent, as may data for a given subject's two ears.

³The target sentences and babble were recorded digitally using an artificial head in anechoic space with five identical loudspeakers, one for the speech signal and four for

⁴The *SPIN* test was presented at a speech-to-babble ratio lower than the recommended 8 dB S/B, to increase the inter-subject performance variability.

(2) There is little reason to believe that test values or measurement errors have normal distributions. (3) Measures can be difficult to compare when they have different scales or are bounded by different limits. (4) Intrasubject variability is large, and differs across individuals (Marshall, 1981). The significance levels of parametric tests on audiological data can be misleading, causing one to reject the null hypothesis not because the alternative is a better explanation of the data, but because the null hypothesis is unrealistic. Conversely, violations of the statistical assumptions can obscure real effects. Nonparametric methods obviate the need for some of these assumptions and provide a robust alternative to parametric tests (Lehmann and D'Abrera, 1988). In the last two decades there has been an explosion in methods based on resampling the data, which also require fewer assumptions about the distribution of the data than parametric methods (e.g., Efron, 1982, Davison & Hinkley, 1997). The present study uses nonparametric methods and resampling to test hypotheses about the changes in audiometric and speech understanding measures as subjects age.

III. RESULTS

Analysis focused on three points: (1) Comparing performance between the two testing phases, (2) determining the rate of change in audiometric and speech understanding measures, and (3) assessing the relationship between speech understanding performance and both age and auditory thresholds. The first point was addressed by analyzing the data using univariate methods; the second, by a comparison between the results for elderly subjects and reference data for young, normal-hearing subjects; and the third, by bootstrap tests of a simple linear model that represents speech understanding as a function of age and hearing sensitivity.

A. Univariate statistics and hypothesis tests

Means and standard deviations of the 14 measures of performance over the two testing phases averaged across the right and left ears of the elderly subjects are illustrated in Figure 2. In all analyses in this section, each subject's measurements in the second testing phase were replaced with linearly interpolated or extrapolated estimates of the subject's performance 5.27 years (the mean lapse) after the first testing phase, in order to normalize the elapsed time between the two testing phases. (The largest extrapolation interval was 1.27 years; the largest interpolation interval was 3.23 years.) Figure 2 shows that essentially all performance measures deteriorated between Phase 1 and Phase 2. The average pure-tone threshold (**PTA4**) increased by 3.33 dB (SD: 2.47 dB), corresponding to 6.32 dB/decade (SD: 4.69 dB), a result very similar to threshold changes reported by other investigators (Davis, Ostri & Parving, 1990, Gates *et al.*, 1990, Ostri & Parving, 1991) and comparable to the lower limit of decay estimated by Pearson *et al.* (1995). Figure 1 shows that the audiometric decline is dominated by the threshold increases at 4 kHz (7.67 dB, SD 9.23 dB, or 14.61 dB/decade) and 8 kHz (5.97 dB, SD 9.09 dB, or 11.32 dB/decade), which explains the relatively large average change of the audiogram slope (**PTSLP**, 5.78 dB, SD 6.16 dB, or 10.98 dB/decade). Because increased thresholds at higher frequencies impair the perception of consonantal information in speech (Dubno, Dirks & Morgan, 1984, Frisina & Frisina, 1997), it is likely that threshold increases at 4 and 8 kHz contributed to the decline of speech understanding in interference. However, the presentation level was threshold-dependent (SL or re/SRT), so the speech signal was always audible even in the 4-kHz range and, because threshold increases at high frequencies also impede perception of the interfering speech, loss of speech understanding is unlikely to be solely the result of simultaneous masking of the signal⁵. Moreover, although we included the 8-kHz threshold as an indicator of the subject's hearing status, no speech material presented had information beyond 5 kHz—the developers of the tests were careful to minimize any direct effect of high-frequency loss. The relatively large standard deviation of the threshold changes at the high

⁵This refers to energetic masking. Informational masking is a different issue; if the definitions advocated by several authors (e.g., Durlach, Mason, Shinn-Cunningham,

frequencies indicates that presbycusis hearing loss increased substantially between testing phases, at least for some subjects.

Differences in the 14 measures by testing phase and laterality were examined using nonparametric methods based on randomization. Observations were paired by subject and by ear across the two phases, which gave 812 differences (14 measures $\times$ 29 subjects $\times$ 2 ears). We sought to determine whether the observed differences between phases were larger than chance would account for if the measurements in each pair had been labeled randomly as arising from phase 1 or from phase 2, as if by tossing a fair coin 812 times. To that end, each difference was coded as -1 , 0 or 1 , depending on whether the subject's measure for that ear at Phase 1 was larger than (-1), equal to (0), or smaller than ($+1$) at Phase 2. We then compared the observed mean of these coded data with the distribution of means that would arise from random assignment by coin tossing.⁶ This allowed us to assign a P -value to the hypothesis that there was no change in each measure between testing phases. We made the analogous test for all the measures pooled together, and to test whether the measures differed by laterality. (These tests involving the mean of the coded responses are equivalent to tests involving the medians of the original responses). The null hypothesis for these tests is that the collection of scores is fixed in advance, and that the measurement process is equivalent to a random assignment of one of each pair of responses to each of the two phases or two ears. In this paper, P -values not exceeding 0.05 are considered to be statistically significant.

Table I shows P -values⁷ for the test of differences in individual measures by testing phase and laterality. For comparison, corresponding P -values obtained in one-tailed t -tests are also shown; when they differ, we trust the nonparametric test for reasons mentioned in the Methods section; see p. 11.8.⁸ Table I also shows P -values for differences pooled across measures. The observed differences are statistically significant at level 0.0001 : the evidence that the median of the measurements is different for the two epochs is quite strong. For all but two individual measures (**LP750** and **RT75**), the difference between Phase 1 and Phase 2 is statistically significant. There is an overall right-ear superiority, but it is statistically significant for fewer than half of the individual measures. Phase-laterality interaction was found to be statistically insignificant (P -value 0.35): laterality did not change significantly between the two testing phases.

The analyses above do not distinguish among audiometric and speech-related measures. Since the median of most of the measures—both audiometric and speech-related—seems to have changed between phases, identical processes could drive their deterioration. However, if the variability changed differently for the two groups of measures, that could indicate that more than hearing loss is responsible for the decline in speech understanding. To investigate this possibility, we tested the hypotheses that the variability of the speech understanding measures is the same in both phases and that the variability of the audiometric measures is the same in both phases against the alternative hypothesis that the variability is higher in the second phase.⁹ Results indicate that the variability changed significantly only for the speech-related variables: the test does not reject the null hypothesis for the audiometric variables (P -value approximately 0.46), but rejects it for the speech-related variables (P -value approximately 0.03). This suggests that changes in central factors may play a role in the degradation of speech understanding—as proposed by other investigators (e.g., Frisina and Frisina, 1997; Gordon-Salant and Fitzgibbons, 1997).

⁶This test is analogous to the sign test for the median (see, e.g., Lehmann & D'Abrera, 1988).

⁷The P -value is the probability that the test statistic would have a value as extreme as or more extreme than its observed value if the null hypothesis is true. Small P -values

⁸ P -values for the t -test were computed using Student's t distribution. P -values for the permutation test were computed using the binomial distribution (for dichotomous d

⁹We tested the null hypothesis that the variability did not change across testing phases using a permutation test. To calculate the test statistic, we subtracted from the mea

B. Rate of decline of hearing sensitivity and of speech understanding in interference

A primary objective of the present research was to explore whether a decline of central function with age—beyond the decline of peripheral function—plays a role in the accelerating decline with age of the ability to understand speech. Comparing the rate of decline before the first testing phase with the rate of decline between the first and second testing phases can help address this question. A young group was used to estimate the elderly subjects' performance at a young age, so the analysis in this section has both cross-sectional and longitudinal aspects. For each elderly subject i and measure j , a “decline difference sign” $d_{i,j}$ was computed for each of the four audiometric variables and each of the eight measures of speech understanding in interference, **RT75**, **SPMon**, **SPMonCEf**, **SPSpat**, **SPSpatCEf**, **SPSpatEf**, **BCST**, and **CSTI**:

$$d_{i,j} = \text{Sign} \left[\frac{(m_{i,j,1} - m_{i,0})}{(a_{i,1} - a_0)} - \frac{(m_{i,j,2} - m_{i,j,1})}{(a_{i,2} - a_{i,1})} \right] \quad (1)$$

where $m_{i,0}$ is the reference value of measure j in the young control group, $m_{i,j,1}$ is measure j for subject i at phase 1, $m_{i,j,2}$ is the corresponding datum at Phase 2, a_0 is the reference (young) age, $a_{i,1}$ is the age of subject i at Phase 1 and $a_{i,2}$ is at the subject's age at Phase 2.¹⁰ The decline difference sign compares the average slope between the reference age (22.18 years) and age at Phase 1 to the average slope between age at Phase 1 and age at Phase 2. If the deterioration accelerates—i.e., if the measure declines more rapidly on average between Phase 1 and Phase 2 than between the reference age and Phase 1—then $d_{i,j}$ is -1 ; if the two rates are the same, $d_{i,j}$ is 0; and if the deterioration decelerates, $d_{i,j}$ is $+1$. Using the signs of the differences of rates puts all the measurements on an equal footing, making it possible to pool across measures to increase the power of the test, and reducing the influence of outliers.

A permutation test was used to test the null hypothesis that the change in rate of decline of a measure does not depend on whether the measure is audiometric or speech-related against the alternative hypothesis that the change in rate of decline for the eight speech-related variables tends to be larger than that of the four audiometric variables. The test statistic was the sum of the decline difference signs for the speech-related measures. The P -value for the null hypothesis was estimated by comparing the observed value of this statistic with the distribution of values of the test statistic obtained by randomly labeling $4 \times 29 = 116$ slope ratios as audiometric, and $8 \times 29 = 232$ slope ratios as speech-related, 10,000 times. The estimated P -value was 0.032: the decline with age of speech-related measurements accelerates faster than the decline with age of audiometric measurements by an amount that is statistically significant. Figure 3 shows the performance decline per year as a function of age at the testing phase separately for the set of four audiometric variables and the set of eight measures of speech understanding in interference. Performance change at the reference age was, by definition, zero. The larger acceleration of decline in the speech-related measures compared to the acceleration of decline in the audiometric measures is clearly visible.

C. Differences in the relationship between audiometric and speech variables across the two testing phases

The analyses above indicate that both audiometric and speech understanding measures declined between the first and second testing phases, and that the variability of speech understanding increased between testing phases. The acceleration analysis in the previous section shows that speech understanding declined more rapidly later in life than audiometric measures did, relative to their decline between about age 20 and about age 60. These findings, however, do not explore the relationships among the variables. We have seen that measures of speech understanding

¹⁰We did not interpolate measures to standardize the time lapse, as we did in the previous section.

are associated with both peripheral hearing and age, but is the association the same at the two testing phases? In this section, we test the hypothesis that the relationship between audiometric variables, age, and speech understanding is constant with time.

Testing this hypothesis requires much stronger assumptions than those made so far: (1) individual cases are independent and identically distributed, as if the subjects were drawn at random from a population,¹¹ (2) speech-related performance variables at the two phases have a linear relationship to the explanatory variables—age and audiometric variables—and (3) the speech performance variables have additive measurement errors that have the same distribution, have zero expected value, and are independent for different individuals, but that need not be independent across testing phases nor across measurements for a given individual.

To reduce the dimensionality of the problem, the largest principal components of the four audiometric and the ten speech-related variables were used instead of all 14 raw measures. Enough principal components were retained to account for at least 90 percent of the total variance of the raw measures: the top two audiometric principal components accounted for 88.9 percent of the variance of the audiometric measures, and the top five speech principal components accounted for 94.0 percent of the variance of the speech understanding measures. The loading matrices (rotated for easier interpretability) are shown in Table II. The first principal component of speech, **Speech₁**, is related mainly to *SPIN* sentence context effect, spatial separation of speech target and babble sources, and temporal distortion; **Speech₂**, to monaural processing of speech in babble and reverberation; **Speech₃**, to sentence-based interference; **Speech₄**, to spectral distortion; and **Speech₅**, to spatially distributed speech processing in babble. **Audio₁** summarizes overall hearing level and **Audio₂**, high-frequency hearing. The model in the null hypothesis is

$$\text{Speech}_{ij} = a_0 + a_1 \text{Age}_{ij} + a_2 \text{Audio}_{1ij} + a_3 \text{Audio}_{2ij} + \text{error}_{ij}, \quad i = 1, 2, \dots, 5; j = 1, 2, \dots, 29. \quad (2)$$

We wish to test the null hypothesis that, for all $i=1, 2, \dots, 5$, each of the principal components of the speech measures, the coefficients a_0 , a_1 , a_2 and a_3 are the same at both phases. The test statistic compares the fit of a model that requires each coefficient a_{ij} to be the same at the two phases (the *restricted* model) with the fit of a model that allows the coefficients to differ (the *unrestricted* or *full* model). The full model will always fit at least as well as the restricted model, because it contains the restricted model as a special case. The measure of fit we use is related to the Chow statistic (Kennedy, 1998), pp. 229-230):

$$\left[(RSS_R - RSS_U) / k \right] / \left[RSS_U / (N - k) \right], \quad (3)$$

where RSS_R is the sum of squared from the restricted model, RSS_U is the sum of squared residuals from the unrestricted model, k is the number of variables in each regression ($k = 5$ here), and N is the number of data in the regression (twice the number of subjects, because there are two testing phases).¹² To test the overall hypothesis that any of the relationships changed, we need to combine the Chow statistics for the five speech variables. Using the maximum of the Chow statistics would be particularly sensitive to a change in only one of the relationships; using the sum of the Chow statistics is more sensitive to small changes in several of the relationships.¹³ We report the result of tests based on the sum of the Chow statistics.¹⁴

To determine whether the observed improvement in this summary measure of fit would occur frequently by chance if the model were correct and the coefficients were constant across phases,

¹¹The assumption that the subjects are like a random sample is implausible, even though it is common in audiological research. The assumption that the relationship between

¹²The ratio (3) would have an F distribution under the null hypothesis if all the data were drawn independently from normal distributions. This assumption can be avoided

¹³It also would be reasonable to use the ratio of the sum of normalized numerators to the sum of normalized denominators in equation (3).

¹⁴The other approaches to combining the Chow statistics gave similar results.

a bootstrap approach was used.¹⁵ Starting with the best-fitting model that restricts the coefficients to be equal across phases, we calculate 29 vectors, each consisting of five pairs of residuals for each of the 29 subjects at the two phases (one pair for each of the five speech principal components at the two phases). Then, vectors of residuals are drawn at random with replacement from the population of 29 vectors of residuals. We add these residuals to the predictions of the best-fitting model, and fit the restricted and unrestricted models to these synthetic data. By repeating this process 1000 times, the sampling distribution of the sum of the Chow statistics was estimated under the assumption that the null hypothesis is true. That allowed us to assign an approximate *P*-value to the null hypothesis.

Differences in the coefficients across phases were statistically significant ($P = 0.017$): On the assumption that individual cases are independent and identically distributed, we would reject the hypothesis that the first five principal components of the speech-related variables have the same linear relationship to age and the first two principal components of the audiometric variables at the two phases. Figure 4 shows the five pairs of coefficients of the three predictors, age and the first two audiometric principal components. The coefficient vectors differ between phases, that is, the five principal components of the speech measures are associated with age and hearing sensitivity differently at Phase 1 and at Phase 2. With increasing age, the monaural perception of speech in babble and in reverberation changes, as does comprehension of sentences in the presence of another simultaneous sentence. As overall hearing threshold level changes, so does the perception of spectrally distorted and spatially distributed speech. As high-frequency loss changes, so does the perception of temporally distorted speech and the utility of sentence context.

IV. DISCUSSION

The most compelling observation of this study is that auditory performance in advanced age declines significantly over a period as short as five years in subjects whose initial hearing loss is only mild to moderate. This decline manifests in pure-tone thresholds and in almost all speech understanding performance measures. Although auditory thresholds might not characterize the status of the peripheral hearing system completely, understanding speech in interference clearly depends on auditory thresholds. In every reasonable model of speech perception, the peripheral auditory system extracts spectral-temporal features of the stimulus, features that comprise the input to higher centers. Using one such simple model, in which speech perception measures depend linearly on age and pure-tone thresholds (Figure 4), we found that the relationships *between* measures of speech perception on the one hand, and age and auditory thresholds on the other, changes with age. This suggests that the nature of speech perception changes with age. While auditory thresholds gauge the role of the peripheral auditory system in speech perception, the effect of age includes the degradation of central structures responsible for low-rate temporal processing, among other things. These two effects can be regarded as complementary and, as Figure 3 shows, they progress at different rates.

We find that audiometric thresholds increase by roughly 6.32 dB/decade, which is consistent with threshold changes reported by others (Davis, Ostri & Parving, 1990, Gates *et al.*, 1990, Ostri & Parving, 1991) and with the lower limit of decay estimated by Pearson *et al.* (1995). The decline in speech understanding measures, illustrated in Figure 2 and Table I, agrees with studies showing an increasing presbycusis loss accompanied by deteriorating speech perception performance (e.g. Bergman *et al.*, 1976, Pedersen, Rosenhall & Møller, 1991, Frisina & Frisina, 1997).

¹⁵The bootstrap approximation of the distribution of an estimator is the distribution of the estimator when it is applied to random samples with replacement from the data.

We also find that the variability of speech-related measures increased significantly between phases, but the variability of audiometric measures did not. Since speech understanding depends on the combination of auditory and cognitive processes, it is not surprising that the variability of this combination might change with age even if the variability of audibility alone does not. The increase of intersubject variability of speech understanding with age may be a facet of the large individual differences observed across a variety of cognitive functions in aged populations (Salthouse, 1991). Speech perception relies in a complex way on many cues, which in turn depend, most likely in a nonlinear fashion, on the acoustic signal impinging on listeners' ears. Thresholding effects that cause some cues to become inaudible are another form of nonlinearity. These nonlinearities could also account for the increased variability of speech understanding relative to audibility. So could individual differences in how audible cues are used. The increased variability differences of measures of speech understanding compared with measures of audibility could also arise from differences between the processes involved in their declines. This hypothesis is also supported by Figure 3, which shows an increase in the rate of decline of speech understanding in interference—i.e., an acceleration. Taken together, the comparative increase in variability of speech understanding compared with audiometric measures, the comparative acceleration of the decline in speech understanding compared with audiometric measures, and the apparent change with time in the relationship between speech understanding and age and audiometric variables¹⁶ suggest that the decline in auditory performance during the later decades of life has two major components: (1) presbycusis due to deterioration of the peripheral auditory structures, and (2) loss of speech understanding involving a non-peripheral (i.e., central) component.

In fact, a centrally originating component of auditory decline has been proposed by many other investigators (Bergman, 1980, Jerger, Jerger, Oliver & Pirozzolo, 1989, Pedersen, Rosenhall & Møller, 1991, Jerger, Chmiel, Allen & Wilson, 1994, Pichora-Fuller, Schneider & Daneman, 1995, Gordon-Salant & Fitzgibbons, 1997, Pichora-Fuller, 1997, Schneider, Speranza & Pichora-Fuller, 1998, Divenyi & Simon, 1999, Gordon-Salant & Fitzgibbons, 1999). What is the functional significance of this central component? Previous results (Bergman *et al.*, 1976, Gelfand, Ross & Miller, 1988, Gordon-Salant & Fitzgibbons, 1993, Divenyi & Haupt, 1997a) suggest, the decline affects most aspects of the temporal processing of simple and complex sounds (Gordon-Salant & Fitzgibbons, 1997)—and thus is quite likely to impede the ability to separate simultaneous speech sounds. Most temporal processing deficits affect the time range most useful for speech processing, i.e., amplitude modulation frequencies of 3 to 6 Hz corresponding to the syllabic rate (e.g., Takahashi & Bacon, 1992, Humes, Watson, Christensen, Cokely, Halling & Lee, 1994, Fitzgibbons & Gordon-Salant, 1995). In our study, the **CSTI** and **BCST** measures are affected by the speech of even a single interfering talker, which can degrade the envelope information at these low AM frequencies. However, speech also can be degraded at higher modulation rates (20 to 50 Hz), where fluctuations important for segmental processing occur. One example is reverberation, which reduces intelligibility by degrading speech envelopes at modulation rates up to 50 Hz (Avendano & Hermansky, 1996). The decline in understanding of reverberant speech, also reported by others (Duquesnoy & Plomp, 1980, Nabelek & Robinson, 1982, Nabelek & Dagenais, 1986, Helfer & Huntley, 1991, Helfer, 1992, Gordon-Salant & Fitzgibbons, 1995, 1999), indicates that temporal information processing in both the syllabic and segmental ranges is impaired in the elderly. Deterioration of temporal processing in the 3 to 50-Hz envelope frequencies thus could cause the decline of speech understanding in interference observed here and elsewhere (Bergman *et al.*, 1976). Babble, reverberation and other modulated non-target sounds have syllabic-rate fluctuations difficult to distinguish from those of the target. If the ability to discern those fluctuations declines, understanding of reverberant speech or speech in babble noise

¹⁶This last assertion depends on the assumption that the relationship is linear in the principal components, as described more fully in section III C of the Results.

presumably will be affected. Our laboratory has found that the ability of elderly listeners to segregate non-speech streams using only envelope information is highly correlated with their ability to understand speech in babble and reverberation (Divenyi, 2004a). Therefore, studying the connection between speech understanding in interference and the decline of auditory temporal processing at low modulation rates might give insight into auditory losses beyond presbycusis. Unfortunately, neither the precise characteristics of the age-related decline in temporal processing nor its precise role in speech perception impairments is yet clear, although research in the area is progressing (Fitzgibbons & Gordon-Salant, 1996, Gordon-Salant & Fitzgibbons, 1999, Snell & Frisina, 2000). It seems that the decline in speech understanding found here is likely to reflect a general deficit of auditory temporal processing in the syllabic and sub-syllabic range (Fitzgibbons & Gordon-Salant, 2004, Gordon-Salant & Fitzgibbons, 2004), i.e., a time scale significantly longer than changes in the auditory periphery would account for.

Right-ear superiority in elderly subjects has been reported by several investigators (Divenyi & Haupt, 1997a, Jerger & Chmiel, 1997, Greenwald & Jerger, 2001). The present study confirms this (see Table I). Because one of our subject selection criteria was symmetrical audiograms, it is not surprising that we saw no laterality effect for audibility measures. The right-ear advantage we find in the present study therefore strongly suggests a central imbalance. This conclusion is consistent with physiological-anatomical observations that attribute much of the right-ear advantage (or, rather, the left-ear disadvantage) in aging to cortical impairment—a deterioration of the fibers in the corpus callosum (Jerger, Alford, Lew, Rivera & Chmiel, 1995, Chmiel *et al.*, 1997). Although we might have predicted an increase in the right-ear advantage from testing phase 1 to testing phase 2, no significant increase was observed, suggesting that the development of central imbalance occurs over a longer time than the duration of our study.

We find that the effect of sentence context can be measured with the *SPIN* test, although the context effect interacts nonmonotonically with the overall *SPIN* score: both poor and excellent performances reduce the apparent utility of sentence context, but for opposite reasons. Subjects who cannot understand speech in noise at all cannot pick out context to use it, and for subjects who can understand every word, context is redundant. To be a useful indicator of the higher-order cortical function of sentence processing, the context effect needs to be weighted by performance level.

It is perhaps counterintuitive that the effect of spatial separation of speech and babble sources was statistically significant but small, possibly because several subjects had difficulty understanding the targets at any separation. There is, however, another likely explanation: auditory localization deteriorates with age (Herman, Warren & Wagener, 1977, Divenyi & Haupt, 1997a), so poor localization could have reduced or obliterated the benefits of increasing the spatial separation between sources. Because auditory localization in the elderly has been measured only with simple signals (clicks or tone pips), the last explanation remains a possibility until tested directly, for example, by replicating on an elderly population the (virtual) free-field speech masking level difference experiments of Shinn-Cunningham and her colleagues (Shinn-Cunningham *et al.*, 2001).

We want to reiterate the importance and utility of nonparametric methods for analyzing auditory performance data. When the assumptions underlying parametric tests are satisfied, nonparametric tests tend to be more conservative, possibly increasing the risk of missing a real effect. But because their assumptions are less restrictive, nonparametric tests, such as the permutation-based tests used in this study, tend to behave well in many situations in which the parametric tests do not, and we believe that the assumptions of usual parametric tests are often not satisfied in studies on hearing. Moreover, many nonparametric tests are in fact simpler and

easier to understand than the corresponding parametric tests. We recommend these nonparametric methods, especially when the sample size is small, when subjects are not chosen at random, and when it is not reasonable to assume that the variables of interest have Gaussian distributions.

Can conclusions of the present study be generalized to the entire aging population? Because the criteria for enrollment in this study excluded people with hearing loss worse than moderate or with significant medical problems, our subject sample does not represent the elderly population as a whole. Had the criteria for inclusion been relaxed, inter-subject variability due to illness and auditory status might well have interfered with the effects we wanted to examine. We believe this study shows important effects of normal aging on speech understanding and that it can serve as a baseline against which to compare results on subject samples having auditory or other pathologies.

To conclude, our results suggest that the decline of speech understanding in interference in the seventh to ninth decades of life is similar to the presbycusis deterioration of hearing with important differences: deterioration of speech understanding accelerates and its variability increases compared with the rate of deterioration and variability of absolute hearing thresholds, suggesting an accelerating decline of central auditory processing. This might have been discovered sooner if central auditory decline with age had received as much attention as peripheral. The central component of these deficits may also explain the limited success elderly hearing aid users have understanding conversations in noisy situations, especially in speech interference. Providing robust speech understanding in adverse listening conditions might eventually be possible using intelligent, automatic speech-separation devices still in development.

ACKNOWLEDGMENT

We are grateful to Brian Gygi, Ken Grant and two anonymous reviewers for comments on an earlier version of the paper, and to David Freedman for numerous helpful conversations. The research has been supported by Grant R01-AG07998 from the National Institute on Aging and by the Veterans Affairs Biomedical and Laboratory Research and Development.

REFERENCES

- Avendano C, Hermansky H. Study on the dereverberation of speech based on temporal envelope filtering. *Proc ICSLP'96* 1996;889–892.
- Beasley DS, Schwimmer S, Rintelmann WF. Intelligibility of time-compressed CNC monosyllables. *J Sp Hear Res* 1972;15:340–350.
- Bergman, M. *Aging and the Perception of Speech*. University Park Press; Baltimore, MD: 1980.
- Bergman M, Blumfield VG, Cascardo D, Dash B, Levitt H, Margulies MK. Age-related decrement in hearing for speech. Sampling and longitudinal studies. *J Gerontol* 1976;31:533–538. [PubMed: 950447]
- Bilger RC, Nuetzel JM, Rabinowitz WM, Rzeczkowski C. Standardization of a test of speech perception in noise. *J Sp Hear Res* 1984;27:32–48.
- Blumenfield VG, Bergman M, Millner E. Speech discrimination in an aging population. *J Sp Hear Res* 1969;12:210–217.
- Brant LJ, Fozard JL. Age changes in pure-tone hearing thresholds in a longitudinal study of normal human aging. *J Acoust Soc Am* 1990;88:813–820. [PubMed: 2212307]
- Cheesman MF, Hepburn D, Armitage JC, Marshall K. Comparison of growth of masking functions and speech discrimination abilities in younger and older adults. *Audiol* 1995;34:321–333.

¹⁷For a review, see Divenyi (2004b)

- Chmiel R, Jerger J. Hearing aid use, central auditory disorder, and hearing handicap in elderly persons. *J Am Acad Audiol* 1996;190–202. [PubMed: 8780992]
- Chmiel R, Jerger J, Murphy E, Pirozzolo F, Tooley-Young C. Unsuccessful use of binaural amplification by an elderly person. *J Am Acad Audiol* 1997;8:1–16. [PubMed: 9046064]
- Davis AC, Ostri B, Parving A. Longitudinal study of hearing. *Acta Otolaryng Suppl* 1990;476:12–22.
- Davison, AC.; Hinkley, DV. *Bootstrap Methods and Their Application*. Cambridge University Press; Cambridge, U.K.: 1997.
- Divenyi, P. Masking the feature-information in multi-stream speech-analogue displays. In: Divenyi, P., editor. *Speech Separation by Humans and Machines*. Kluwer Academic Publishers; New York: 2004a. p. 269–281.
- Divenyi, PL., editor. *Speech Separation by Humans and Machines*. Kluwer Academic Publishers; New York, NY: 2004b.
- Divenyi PL, Haupt KM. Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. I. Age and laterality effects. *Ear Hear* 1997a;18:42–61. [PubMed: 9058037]
- Divenyi PL, Haupt KM. Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. II. Correlation analysis. *Ear Hear* 1997b;18:100–113. [PubMed: 9099559]
- Divenyi PL, Haupt KM. Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. III. Factor representation. *Ear Hear* 1997c;18:189–201. [PubMed: 9201454]
- Divenyi PL, Simon HJ. Hearing in aging: Issues old and new. *Curr Opin Otolaryng* 1999;7:282–289.
- Dubno JR, Dirks DD, Morgan DE. Effects of age and mild hearing loss on speech recognition in noise. *J Acoust Soc Am* 1984;76:87–96. [PubMed: 6747116]
- Dubno JR, Lee F, Matthews LJ, Mills JH. Age-related and gender-related changes in monaural speech recognition. *J Speech Lang Hear Res* 1997;40:444–452. [PubMed: 9130212]
- Duquesnoy AJ, Plomp R. Effect of reverberation and noise on the intelligibility of sentences in cases of presbycusis. *J Acoust Soc Am* 1980;68:537–544. [PubMed: 7419810]
- Durlach NI, Mason CR, Shinn-Cunningham BG, Arbogast TL, Colburn HS, Kidd G Jr. Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *J Acoust Soc Am* 2003;114:368–379. [PubMed: 12880048]
- Efron, B. *The jackknife, the bootstrap, and other resampling plans*. Society of Industrial and Applied Mathematicians (SIAM); Philadelphia: 1982.
- Fitzgibbons PJ, Gordon-Salant S. Age effects on duration discrimination with simple and complex stimuli. *J Acoust Soc Am* 1995;98:3140–3148. [PubMed: 8550939]
- Fitzgibbons PJ, Gordon-Salant S. Auditory temporal processing in elderly listeners. *J Am Acad Audiol* 1996;7:183–189. [PubMed: 8780991]
- Fitzgibbons PJ, Gordon-Salant S. Age effects on discrimination of timing in auditory sequences. *J Acoust Soc Am* 2004;116:1126–1134. [PubMed: 15376678]
- Frisina DR, Frisina RD. Speech perception and presbycusis: relations to possible neural mechanisms. *Hear Res* 1997;106:95–104. [PubMed: 9112109]
- Gates GA, Cooper JC Jr. Incidence of hearing decline in the elderly. 1991
- Gates GA, Cooper JC, Kannel WB, Miller NJ. Hearing in the Elderly: The Framingham Cohort, 1983–1985. Part I. Basic Audiometric Test Results. *Ear Hear* 1990;11:247–256. [PubMed: 2210098]
- Gelfand SA, Hoffman S, Waltzman SB, Piper N. Dichotic CV recognition at various interaural temporal onset asynchronies: effect of age. *J Acoust Soc Am* 1980;68:1258–1261. [PubMed: 7440848]
- Gelfand SA, Piper N, Silman S. Consonant recognition in quiet as a function of aging among normal hearing subjects. *J Acoust Soc Am* 1985;78:1198–1206. [PubMed: 4056214]
- Gelfand SA, Piper N, Silman S. Consonant recognition in quiet and in noise with aging among normal hearing subjects. *J Acoust Soc Am* 1986;80:1589–1598. [PubMed: 3794064]
- Gelfand SA, Ross L, Miller S. Sentence reception in noise from one versus two sources: Effect of aging and hearing loss. *J Acoust Soc Am* 1988;83:248–256. [PubMed: 3343444]

- Gordon-Salant S, Fitzgibbons PJ. Temporal factors and speech recognition performance in young and elderly listeners. *J Speech Hear Res* 1993;36:1276–1285. [PubMed: 8114494]
- Gordon-Salant S, Fitzgibbons PJ. Recognition of multiply degraded speech by young and elderly listeners. *J Speech Hear Res* 1995;38:1150–1156. [PubMed: 8558883]
- Gordon-Salant S, Fitzgibbons PJ. Selected cognitive factors and speech recognition performance among young and elderly listeners. *J Speech Lang Hear Res* 1997;40:423–431. [PubMed: 9130210]
- Gordon-Salant S, Fitzgibbons PJ. Profile of auditory temporal processing in older listeners. *J Speech Lang Hear Res* 1999;42:300–311. [PubMed: 10229448]
- Gordon-Salant S, Fitzgibbons PJ. Effects of stimulus and noise rate variability on speech perception by younger and older adults. *J Acoust Soc Am* 2004;115:1808–1817. [PubMed: 15101658]
- Greenwald RR, Jerger J. Aging affects hemispheric asymmetry on a competing speech task. *J Am Acad Audiol* 2001;12:167–173. [PubMed: 11332516]
- Helfer KS. Aging and the binaural advantage in reverberation and noise. *J Sp Hear Res* 1992;35:1394–1401.
- Helfer KS, Huntley RA. Aging and consonant errors in reverberation and noise. *J Acoust Soc Am* 1991;90:1786–1796. [PubMed: 1960274]
- Herman GE, Warren LR, Wagener JW. Auditory lateralization: Age differences in sensitivity to dichotic time and amplitude cues. *J Gerontol* 1977;32:187–191.
- Humes LE. Understanding the speech-understanding problems of the hearing impaired. *J Am Acad Audiol* 1991;2:59–69. [PubMed: 1768875]
- Humes LE, Christopherson L. Speech identification difficulties of hearing-impaired elderly persons: The contributions of auditory processing deficits. *J Sp Hear Res* 1991;34:686–693.
- Humes LE, Roberts L. Speech recognition difficulties of the hearing-impaired elderly: The contributions of audibility. *J Sp Hear Res* 1990;33:726–735.
- Humes LE, Watson BU, Christensen LA, Cokely CG, Halling DC, Lee L. Factors associated with individual differences in clinical measures of speech recognition among the elderly. *J Sp Hear Res* 1994;37:465–474.
- Jerger J, Alford B, Lew HL, Rivera V, Chmiel R. Dichotic listening, event-related potentials, and interhemispheric transfer in the elderly. *Ear Hear* 1995;16:482–498. [PubMed: 8654903]
- Jerger J, Chmiel R. Factor analytic structure of auditory impairment in elderly persons. *J Am Acad Audiol* 1997;8:269–276. [PubMed: 9272749]
- Jerger J, Chmiel R, Allen J, Wilson A. Effects of age and gender on dichotic sentence identification. *Ear Hear* 1994;15:274–286. [PubMed: 7958527]
- Jerger J, Silman S, Lew HL, Chmiel R. Case studies in binaural interference: Converging evidence from behavioral and electrophysiologic measures. *J Am Acad Audiol* 1993;4:122–131. [PubMed: 8471784]
- Jerger, J.; Stach, BA.; Johnson, K.; Loiselle, LH.; Jerger, S. Patterns of abnormality in dichotic listening in the elderly. In: Jensen, JH., editor. *Presbycusis and other age related aspects - 14th Danavox Symposium*. Danavox; Copenhagen: 1990. p. 143-150.
- Jerger JF, Hayes D. Diagnostic speech audiometry. *Arch Otolaryng* 1977;103:216–222. [PubMed: 849199]
- Jerger JF, Jerger S, Oliver T, Pirozzolo F. Speech understanding in the elderly. *Ear Hear* 1989;10:79–89. [PubMed: 2707505]
- Jerger JF, Jordan C. Age-related asymmetry on a cued-listening task. *Ear Hear* 1992;13:272–277. [PubMed: 1397770]
- Kalikow DN, Stevens KN, Elliott LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *J Acoust Soc Am* 1977;61:1337–1351. [PubMed: 881487]
- Kennedy, P. *A guide to econometrics*. MIT Press; Cambridge, MA: 1998.
- Konkle DF, Beasley DS, Bess FM. Intelligibility of time-altered speech in relation to chronological aging. *J Sp Hear Res* 1977;20:108–115.
- Lehmann, EL.; D'Abrera, HJM. *Nonparametrics: Statistical Methods Based on Ranks*. 2nd. McGraw Hill; New York: 1988.

- Lynn GE, Gilroy J. Neuro-audiological abnormalities in patients with temporal lobe tumors. *J Neurol Sci* 1972;17:167–184. [PubMed: 5053923]
- Lynn, GE.; Gilroy, J. Effects of brain lesions on the perception of monotonic and dichotic speech stimuli. In: Sullivan, MD., editor. *Central auditory processing disorders*. University of Nebraska Medical Center; Omaha: 1975. p. 47-83.
- Marshall L. Auditory processing in aging listeners. *J Sp Hear Disord* 1981;46:226–240.
- Martini A, Bovo R, Agnoletto M, Da Col M, Drusian A, Liddeo M, Morra B. Dichotic performance in elderly Italians with Italian stop consonant- vowel stimuli. *Audiol* 1988;27:1–7.
- Morrell CH, Gordon-Salant S, Pearson JD, Brant LJ, Fozard JL. Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. *J Acoust Soc Am* 1996;100:1949–1967. [PubMed: 8865630]
- Nabelek AK, Dagenais P. Vowel errors in noise and in reverberation by hearing-impaired listeners. *J Acoust Soc Am* 1986;80:741–748. [PubMed: 3760327]
- Nabelek AK, Robinson PK. Monaural and binaural speech perception in reverberation for listeners of various ages. *J Acoust Soc Am* 1982;71:1242–1248. [PubMed: 7085996]
- Ostri B, Parving A. A longitudinal study of hearing impairment in male subjects--an 8-year follow-up. *Br J Audiol* 1991;25:41–48. [PubMed: 2012903]
- Pearson JD, Morrell CH, Gordon-Salant S, Brant LJ, Metter EJ, Klein LL, Fozard JL. Gender differences in a longitudinal study of age-associated hearing loss. *J Acoust Soc Am* 1995;97:1196–1205. [PubMed: 7876442]
- Pedersen KE, Rosenhall U, Møller MB. Longitudinal study of changes in speech perception between 70 and 81 years of age. *Audiol* 1991;30:201–211.
- Pichora-Fuller MK. Language comprehension in older listeners. *J Sp Lang Pathol Audiol* 1997;21
- Pichora-Fuller MK, Schneider BA, Daneman M. How young and old adults listen and remember speech in noise. *J Acoust Soc Am* 1995;97:593–608. [PubMed: 7860836]
- Schneider B, Speranza F, Pichora-Fuller MK. Age-related changes in temporal resolution: Envelope and intensity effects. *Can J Exp Psychol* 1998;52:184–191. [PubMed: 10095852]
- Shinn-Cunningham BG, Schickler J, Kopco N, Litovsky R. Spatial unmasking of nearby speech sources in a simulated anechoic environment. *J Acoust Soc Am* 2001;110:1118–1129. [PubMed: 11519579]
- Snell KB, Frisina DR. Relationships among age-related differences in gap detection and word recognition. *J Acoust Soc Am* 2000;107:1615–1626. [PubMed: 10738815]
- Stach BA, Jerger JF, Fleming KA. Central presbycusis: a longitudinal case study. *Ear Hear* 1985;6:304–306. [PubMed: 4076554]
- Takahashi GA, Bacon SP. Modulation detection, modulation masking, and speech understanding in noise in the elderly. *J Sp Hear Res* 1992;35:1413–1429.
- Tun PA. Fast noisy speech: age differences in processing rapid speech with background noise. *Psychol Aging* 1998;13:424–434. [PubMed: 9793118]
- van Rooij JCGM, Plomp R. Auditive and cognitive factors in speech perception by elderly listeners. II: Multivariate analyses. *J Acoust Soc Am* 1990;88:2611–2624. [PubMed: 2283434]
- Willeford, JA. Sentence tests of central auditory dysfunction. In: Katz, J., editor. *Handbook of clinical audiology*. 3rd. Williams & Wilkins; Baltimore: 1985. p. 404-420.

Figure 1. Average of right and left-ear audiograms of the 29 subjects measured in the first phase (round symbols/solid lines) and second phase (square symbols/broken lines). Error bars are one standard deviation.

Figure 2.

Performance means and standard deviations of the 29 subjects in the first phase (solid bars) and second phase (open bars) for four audiometric tests (**PTA4**, **PTSLP**, **PT8K**, and **SRT**) and ten tests of speech understanding in interference, averaged across the right and left ears. Data for individual subjects in the second testing phase were interpolated to correspond to a 5.27-year lapse between phases. The *P*-values above the data bars indicate the significance level for a permutation test of the difference between phases for the raw data. The measure acronyms are **PTA4**: average of four pure-tone thresholds; **PTSLP**: slope of the audiogram (4 kHz – 0.5 kHz thresholds); **PT8K**: threshold at 8 KHz; **SRT**: speech recognition threshold; **LP750**: recognition of speech low-pass filtered at 750 Hz; **TC60**: recognition of speech

compressed 60%; **CSTI**: sentence intelligibility in the presence of a monaurally mixed sentence by another speaker; **BCST**: sentence intelligibility in the presence of mixed with a contralaterally presented sentence by another speaker; **SPMon**: average monaural SPIN test score; **SPMonCEf**: monaural *SPIN* sentence context effect; **SPSpat**: average scores for spatially distributed speech target and four babble sources; **SPSpatCEf**: sentence context effect for spatially distributed *SPIN* test; **RT75**: reverberation time estimated to correspond to 75% word recognition. Statistical significance of the difference between results obtained in the two testing phases (see Table I) is indicated for each measure (***: $P < 0.001$, **: $P < 0.005$, +: $P < 0.05$).

Figure 3. Average normalized decline of three audiometric measures (solid line) and eight measures of speech understanding in interference (broken line), as a function of the subjects' average age.

Figure 4. Estimated coefficients in the linear model of equation (3) for the five top principal components of speech understanding measures as a function of age and the two top principal components of audiometric measures, in testing Phase 1 (solid symbols) and 2 (open symbols).

Table 1
P-values of Differences under Sign-of-Data Permutation Test and T-Test

Measure	Phase 1 - Phase 2 Difference		Right - Left Laterality Difference	
	Permutation Test	T-Test	Permutation Test	T-Test
PTA4	0.0000	0.0000	0.2051	0.2908
PTSLP	0.0000	0.0000	0.0063	0.0043
P8kT	0.0002	0.0002	0.9703	0.9972
SRT	0.0214	0.0350	0.9186	0.9718
SPMon	0.0000	0.0000	0.0435	0.0073
SPMon	0.0267	0.0001	0.7441	0.5838
SPSpat	0.0001	0.0000	0.3470	0.4055
SPSpat	0.0000	0.0000	0.2559	0.0496
SPSpat	0.0027	0.0008	0.0004	0.0000
RT75	0.8209	0.2887	0.9760	0.9334
LP750	0.1144	0.0135	0.2483	0.2402
TC60	0.0001	0.0000	0.0220	0.0142
BCST	0.0006	0.0006	0.0099	0.0017
CSTI	0.0000	0.0000	0.0012	0.0004
Overall	0.0000	0.0000	0.0001	0.4235

Table 2
Varimax Rotated Loading Matrices from Principal Component Analysis

Audiometric Measures			Speech Measures					
Principal Components Percent Variance Accounted for	1	2	Principal Components Percent Variance Accounted for	1	2	3	4	5
SRT	47.55	41.36	SPSpCEf	29.75	17.21	16.93	15.86	14.34
PTA4	0.989	0.075	SPMonCEf	−0.943	−0.163	−0.186	−0.067	0.130
	0.904	0.378	SPSpatEf	−0.878	−0.270	−0.095	−0.117	−0.259
PTSLP			TC60	0.736	0.180	0.277	0.151	0.479
PT8K	0.076	0.923		0.610	0.268	0.396	0.504	0.260
	0.320	0.809	SPMon					
			RT75	0.222	0.800	0.136	0.285	0.379
			BCST	0.470	0.753	0.297	0.161	0.196
			CSTI	0.241	0.128	0.919	0.201	0.014
			LP750	0.208	0.385	0.633	0.560	0.018
			SPSpat	0.106	0.201	0.254	0.872	0.293
				0.138	0.323	−0.025	0.262	0.886