Speech Understanding in the Elderly

Larry E. Humes*

Abstract

Three basic hypotheses regarding the speech-understanding difficulties of the elderly are reviewed: the peripheral, central-auditory, and cognitive hypotheses. Evidence obtained in our laboratory and in others is reviewed regarding the viability of each hypothesis. It is concluded that the strongest support exists for the peripheral hypothesis. Specifically, individual variations in the amount of sensorineural hearing loss among the elderly are most responsible for individual variations in speech-understanding performance. (*Speech understanding* is used throughout this article as a general term for the proportion of a speech signal that is accurately perceived by a listener whether in a discrimination, identification, recognition, or comprehension paradigm.) The focus to date, however, has been placed on monaural speech understanding measured in quiet, noise, or reverberation. It is possible that a more complex picture may yet emerge for other forms of temporally distorted speech or for dichotic measures of speech understanding.

Key Words: elderly, presbyacusis, speech recognition, speech understanding

Few would argue with the statement that many people over 60 years of age have difficulty understanding speech and that, as age increases beyond 60 years, the likelihood that such difficulties will be encountered also increases. There is probably more disagreement, however, as to the primary factor or factors underlying the observed speech-recognition difficulties among the elderly.

An excellent review of the hypothesized factors and mechanisms underlying the speech-understanding difficulties of those over 60 years of age was published in 1988 by a Working Group of the Committee on Hearing and Bioacoustics and Biomechanics of the National Research Council (CHABA, 1988). Briefly, the Working Group considered three distinct hypotheses as to the mechanisms accompanying age-related decline in speech understanding. These three basic hypotheses are contrasted in Figure 1 using a highly schematized overview of the structure and function of various components of the auditory system. The simplest of the three hypotheses, illustrated in the top of Figure 1, was the peripheral hypothesis, which maintained that the speech-recognition difficulties of the elderly were primarily attributable to age-related changes in the auditory periphery. This hypothesis could be further subdivided into two versions: (1) simple changes in audibility associated with peripheral changes in structure (outer ear through inner ear and eighth nerve); and (2) other peripheral deficits accompanying cochlear pathology beyond the loss of hearing sensitivity, such as abnormal spectral and temporal resolution. The second hypothesis, shown in the middle panel of Figure 1, maintained that there were structural or functional changes in the auditory pathways of the brain stem or the auditory portions of the cortex. This will be referred to as the central-auditory hypothesis. The final hypothesis, the cognitive hypothesis, is illustrated in the lower portion of Figure 1. Notice that higher centers in the auditory pathway are again involved in this hypothesis, just as in the central-auditory hypothesis. The cognitive hypothesis, however, differs from the central-auditory hypothesis in that a general cognitive deficit is implied, which affects not only the auditory modality, but results in dysfunction for similar processing functions in other sensory modalities. This hypothesis, for example, would suggest that those elderly having difficulty processing auditory stimuli presented dichotically would also have difficulty on similar tasks with visual stimuli (dichoptic vision).

Of course, as the CHABA Working Group was quick to note, it is possible that all of these
hypotheses are valid either for the elderly as a group, with some elderly having pure peripheral explanations of their difficulties and others having pure central-auditory or cognitive explanations, or as combined factors within an elderly individual, with a portion of a given listener’s difficulty being due to peripheral factors, a portion being due to central-auditory dysfunction and the remainder attributable to cognitive decline.

Over the past several years, we have conducted a series of studies examining the contributions of various factors to the speech-recognition deficits of the elderly. For various reasons, we began this series of studies with the objective of seeing how far we could take the simplest of all of the hypotheses, the peripheral hypothesis. Moreover, we began by evaluating the simplest form of the simplest hypothesis, which maintains that the primary peripheral factor underlying the speech-recognition deficit of the elderly is the well-documented peripheral sensorineural hearing loss associated with presbyacusis.

A common control procedure used in several of these studies was to introduce a spectrally shaped masking noise into the ear of normal-hearing young adults and then test these control subjects in a manner identical to that used with the elderly. The spectrally shaped masking noise was designed so that the young normal-hearing adults would have thresholds shifted by the noise to levels identical to those of the average audiogram obtained in quiet from the elderly subjects. Previous work on modeling the perceptual effects of sensorineural hearing loss for younger hearing-impaired listeners, including loudness recruitment and level-dependent changes in frequency resolution, had indicated that this would also provide a good approximation of these effects in young normal-hearing subjects (Humes et al., 1988; Humes and Jesteadt, 1991).

The first study in this series (Humes and Roberts, 1990) compared the monaural and binaural speech-identification performance of young normal-hearing listeners, elderly hearing-impaired subjects, and young listeners with hearing loss simulated via noise masking (simulation for monaural conditions only). Conditions evaluated included reverberation and background noise, both separately and in combination, with the materials consisting of a closed-set nonsense-syllable identification task. As would be expected solely from the presence of sensorineural hearing loss, the hearing-impaired elderly performed worse than young normal-hearing subjects on all speech-identification tasks. The monaural performance of the elderly, however, was virtually identical to that observed in the noise-masked simulated hearing loss. Moreover, among the elderly, there were strong correlations ($r = -0.7$ to $-0.9$) between the amount of pure-tone hearing loss (average at 1000, 2000, and 4000 Hz) and speech-identification scores, regardless of the listening condition. In general, these results were consistent with the peripheral hypothesis, which ascribes individual differences in performance to individual differences in peripheral processing accompanying sensorineural hearing loss.

Next, we further studied monaural speech-identification performance in the elderly by examining the contributions of hearing loss and aging to the identification of nonsense syllables

---

**Figure 1** Highly schematic overview of the auditory system from periphery to cortex used to contrast three hypotheses regarding the mechanisms underlying age-related changes in speech understanding: peripheral hypothesis (top), central-auditory hypothesis (middle), and cognitive hypothesis (bottom). The dashed-dotted box highlights the region of presumed dysfunction in each of the hypotheses.
that were degraded both spectrally and temporally (Humes and Christopherson, 1991). Temporal degradation was accomplished via reverberation (\( T_{60} = 0.8 \) sec), as in the previous study, and spectral distortion was introduced by bandpass filtering the materials from 500 to 2000 Hz. The addition of spectral distortion via filtering was introduced for two primary reasons. First, it had been suggested that the effects of aging on speech recognition may be subtle and would require use of more difficult speech materials to observe age-related effects (Bergman, 1980). In addition, we were attempting to correlate speech-identification performance to a battery of auditory-discrimination tests, referred to as the Test of Basic Auditory Capabilities (TBAC), (Watson et al., 1982; Christopherson and Humes, 1992). This battery of tests makes use of stimuli that range in frequency from about 500 to 2000 Hz, and it has been demonstrated previously that correlations between psychoacoustic measures and measures of speech understanding are optimized when the frequency regions examined with both tests are similar (Thibodeau and Van Tasell, 1987).

Once again, the results of this study indicated that the primary factor determining performance on the nonsense-syllable identification task was the sensorineural hearing loss of the listener. The greater the hearing loss, the lower the speech-identification score. However, group differences were also observed between young normal-hearing listeners and elderly hearing-impaired listeners on a number of the auditory-discrimination tasks. By comparing the performance of a third experimental group, noise-masked normal-hearing young adults, to that of the other two groups on the same task, however, it was possible to discern which effects were likely due to the presence of hearing loss and which were due to age or the interaction of age and hearing loss. Doing so, it was clear that there were several auditory-discrimination tasks on which the elderly hearing-impaired subjects performed worse than young normal-hearing subjects, whether the latter group was tested in quiet or in the presence of a spectrally shaped noise simulating the average hearing loss of the elderly. We also observed that, when we divided the elderly subjects into two groups according to their age, young-old (63-74 years) and old-old (75-83 years), there appeared to be further age-related decline in performance among the elderly on these same tasks. In the study by Humes and Christopherson (1991), we did not have cognitive measures for the subjects, however, and it has since been shown that performance on the TBAC battery among young adults is affected by cognitive ability (Watson, 1991). We decided to conduct a follow-up study in which we controlled for general cognitive differences between the young-old and old-old (Lee and Humes, 1992). In the follow-up study, we matched two groups of elderly subjects for hearing loss and general cognitive function (IQ and digit-span memory), with age being the only between-group difference (young-old vs old-old). As can be seen in Figure 2, the more recent study by Lee and Humes (1992, vertical bars) failed to find an age-related difference in performance between the old-old and young-old groups (filled and unfilled vertical bars are identical). When these data are compared to the results of the previous study obtained from young normal-hearing subjects with and without simulated hearing loss (filled and unfilled circles in Fig. 2, respectively), both elderly groups performed considerably poorer than either group of young normal-hearing subjects on four of the TBAC tests. In summary, there do seem to be some age-related differences in auditory-discrimination performance using a test that has proven to be reliable when applied to the elderly (Christopherson and Humes, 1992). In particular, the elderly demonstrate observed deficits in

![Figure 2](image-url)
auditory-discrimination tasks that involve pure-tone frequency discrimination ("delta F"), temporal-order discrimination for sequences of tones ("temp ord") or syllables ("syll seq"), and intensity discrimination involving complex tonal patterns ("10-tone").

Are these differences in auditory-discrimination performance between young and elderly subjects related in any way to differences in speech identification? In both the studies by Humes and Christopherson (1991) and Lee and Humes (1992), multiple-regression analyses revealed that the vast majority of the differences across subjects could be explained by individual variations in the degree of sensorineural hearing loss. In both studies, however, significant percentages of additional variance could be accounted for by inclusion of some measures of auditory-discrimination performance. Across studies and conditions, the auditory-discrimination test emerging most frequently as an important contributor was simple tonal frequency discrimination for a 1000-Hz tone. The percentage of variance accounted for by this variable ranged from about 3 to 14 percent, however, whereas that accounted for by the pure-tone hearing loss ranged from 55 to 90 percent.

Perhaps our inability to find correlations between speech-understanding performance and auditory measures other than hearing loss resulted from the limited data collected in the studies described above. It could be argued, for instance, that use of nonsense syllables at a moderate level (70 dB SPL) in each of the previous studies could increase the likelihood that performance would be related to peripheral factors, such as the hearing loss of the listener. Few cognitive resources are required, for instance, to identify consonant-vowel or vowel-consonant syllables in a closed set. Further, use of a conversational level could enhance the importance of the hearing loss to an extent that might not be observed at higher levels, especially in a noise background. In addition, it was possible that nonauditory factors, such as cognition, memory, and attention, could affect speech-understanding performance for more complex speech materials in a way not captured by the simple auditory-discrimination measures used previously.

To address these issues, a larger scale study was conducted in which 50 elderly persons, between the ages of 63 and 84 years, served as subjects (Humes et al, 1994). A wide range of speech-understanding measures was included that spanned from closed-set identification of nonsense syllables to open-set recognition of the final word in meaningful sentences (five types of speech materials all together). In addition, two speech levels (70 and 90 dB SPL) and two background conditions (quiet and +7 dB signal-to-noise ratio) were included. In addition to these 20 measures of speech understanding, all subjects were again administered the TBAC battery of auditory-discrimination tasks. In addition, measures of cognitive function were obtained with the revised Wechsler Adult Intelligence Scale (WAIS-R; Wechsler, 1981) and the revised Wechsler Memory Scale (WMS-R; Wechsler, 1987).

The two primary findings that emerged from this study were that (1) the 20 measures of speech understanding were highly correlated, such that subjects who performed poorly on one task tended to perform poorly on all of the speech-understanding tasks; and (2) the single strongest factor that accounted for individual differences in speech-understanding performance was the degree of sensorineural hearing loss of the listener. A very strong association (canonical correlation of 0.89) was observed between the set of 20 speech-recognition measures and the set of auditory and cognitive measures with the primary, if not sole, factor underlying that association being due to the effects of a single predictor variable, the average hearing loss at 1000, 2000, and 4000 Hz.

We began this series of studies fully expecting to observe a complex relationship between individual differences in speech understanding and a host of other predictive measures. Instead, a strong association was observed repeatedly between speech-understanding ability and the average high-frequency hearing loss. That being the case, we decided to investigate a couple of the areas that led us to believe that the situation would be much more complicated than observed. In prior auditory perceptual research conducted with the elderly over the past 20 years, two of the most influential findings suggesting that more than loss of hearing sensitivity was involved in the elderly were a study on age-related changes in the auditory filter (Patterson et al, 1982) and the frequent identification of a "distortion" component in a widely used model of speech-reception threshold (SRT) developed by Plomp (1978).

Briefly, the age-related widening of the auditory filter was demonstrated to be due to associated declines in hearing threshold accompanying the sensorineural impairment (Sommers...
and Humes, 1993). In that study, elderly listeners with normal hearing had auditory filters essentially identical to those of young normal-hearing subjects, whereas those obtained from elderly hearing-impaired subjects, although broader than in young normal-hearing subjects, were identical to those observed in young normal-hearing subjects having thresholds elevated by a second masking noise. Several other recent studies have also found that the observed broadening of auditory filters in the elderly is primarily due to the associated sensorineural hearing loss (Peters and Moore, 1992a, b).

Regarding the Plomp (1978) SRT model and the repeated appearance of a “distortion” factor in this model when applied to the elderly, a recent study by Lee and Humes (1993) demonstrated that the observed “distortion” may actually be a consequence of the loss of audibility in the high frequencies. The primary observation that results in a significant “distortion” term in Plomp’s SRT model is that the elderly frequently require a higher than normal signal-to-noise ratio to achieve SRT (50% correct) at moderately high noise levels. Lee and Humes (1993) demonstrated, however, that, when care is taken to ensure that every subject in the study experiences masking from 250 to 4000 Hz for at least the highest noise levels used, then the SRT measured for the elderly hearing-impaired subjects is not significantly different from that of the young normal-hearing listeners at these same high noise levels. That is, once the background noise is sufficiently intense to be the factor limiting the audibility of the speech signal across the 250- to 4000-Hz frequency range, both groups perform equivalently. At lower noise levels, the hearing loss of the elderly subjects in the high frequencies reduces the audible bandwidth of both the speech and the noise. To compensate for the reduced bandwidth while still keeping the performance level constant (SRT or 50% correct), the signal-to-noise ratio must be increased. Thus, it is the loss of audibility in the high frequencies associated with aging that underlies the observed emergence of a significant distortion factor among the elderly in this paradigm. As demonstrated by Plomp (1986), comparable audibility losses produced in young normal-hearing listeners via low-pass filtering produce significant “distortion” factors within Plomp’s SRT model. Thus, the “distortion” term may capture spectral distortion associated with frequency-specific changes in audibility, not unlike that associated with filtering of speech and noise, and not some audibility-independent form of distortion.

The foregoing evidence argues strongly in favor of the peripheral hypothesis as the major mechanism underlying age-related changes in speech understanding. Several other recent studies have also found the degree of sensorineural hearing loss to be either the primary or sole factor accounting for individual differences in speech recognition among the elderly (van Rooij et al, 1989; van Rooij and Plomp, 1990, 1992; Helfer and Wilber, 1990; Jerger et al, 1991; Helfer, 1992; Souza and Turner, 1994).

What then of the central-auditory and cognitive hypotheses? We have recently reviewed much of the evidence regarding the existence of central-auditory processing disorders among the elderly population and their relation to measured speech-understanding deficits (Humes et al, 1992). Estimates of the prevalence of central-auditory processing disorders among the elderly range from about 10 to 20 percent among a stratified random sample of the US population (Cooper and Gates, 1991) to 80 to 90 percent in some studies of clinical populations (Stach et al, 1990). As noted in our previous review, however, the high estimates of prevalence can be criticized on several grounds. First, the primary test approach used to detect the presence of central auditory disorder has been to employ a parallel test battery with loose criterion. Thus, if three tests of central-auditory processing are administered to a group of elderly subjects, failure on any one of the tests constitutes presence of the disorder. As noted by Humes et al (1992), this tends to bias prevalence estimates toward higher values. Second, the tests used have infrequently been evaluated in terms of reliability and those that have been evaluated in this population to date have proven to be unreliable (Dubno and Dirks, 1983; Cokely and Humes, 1992). Finally, it is difficult, if not impossible, to validate the findings against a “gold standard.” The observation of a test result that is consistent with observations made previously in subjects with known central lesions, especially on just one of several tests in a battery, does not in and of itself imply that the same type of lesion is present. By way of analogy, measurement of a low-frequency air-bone gap in all subjects with fluid-filled middle ears does not mean that all subsequent observations of low-frequency air-bone gap are due to fluid in the middle ear. Moreover, it has seldom been the case that observations of significant central-auditory processing disorders among the elderly have definitively eliminated the possibility that the observed deficit is cognitive in nature. This could best be accomplished, as noted
previously, by confirming that the observed deficit in the auditory modality did not also manifest itself for an identical task in another modality. Thus, arguments in favor of a modality-specific central auditory deficit could be enhanced by demonstrating that comparable deficits are not evident in other modalities. Subjects with deficits in dichotic processing of digits presented acoustically, for example, could also be examined for the ability to process visual digits dichotically. Similarly, recognizing speech in noise auditorily could also be examined visually to confirm that the problem is unique to the auditory modality and not a general cognitive (figure-ground) deficit. Although a few recent studies have attempted to isolate cognitive from central-auditory deficits in evaluating speech understanding in the elderly (Jerger et al., 1991), these studies have not taken the approach of using identical test procedures in the different modalities.

Perhaps the most definitive studies examining the role of auditory and cognitive factors in speech recognition among the elderly are those conducted by van Rooij and colleagues (van Rooij et al., 1989; van Rooij and Plomp, 1990, 1992). A large battery of auditory and cognitive tests, with parallel auditory and visual versions of the latter, were administered to various groups of elderly, along with measures of speech understanding in quiet and noise. The primary findings from this series of studies were that (1) hearing loss was the major factor underlying individual differences in speech understanding among the elderly; (2) cognitive measures could account for only a very small portion of additional variance, especially for a subject group selected randomly from the general public; and (3) cognitive processing deficits observed auditorily were also manifested visually on identical tasks.

The focus in the work described above has been placed on understanding the speech-recognition difficulties of the elderly for speech in quiet, in noise, or in reverberation. Repeatedly, the sensorineural hearing loss of the listener emerged as either the sole or primary explanatory variable for these measures of speech understanding. It remains possible, however, that other factors will account for individual variations in speech-understanding performance degraded in other ways. In particular, there is increasing evidence that speech unnaturally distorted temporally by means other than reverberation, such as time compression or interruption, may be particularly difficult for the elderly and may not be explained by hearing loss alone. Much of this work is reviewed later in this issue in the article by Fitzgibbons and Gordon-Salant.

In addition, most of the work described above has involved attempts to understand monaural speech recognition in the elderly. It is possible that attempts to account for individual differences in speech understanding among the elderly will not be so straightforward for dichotic listening situations. Jerger et al. (1991), although finding the average hearing loss to again be the most significant factor associated with both monotic and dichotic speech understanding, also noted that this factor accounted for much less of the variance for the dichotic conditions. With less accounted for by the sensorineural hearing loss, the door is open for other variables to make significant contributions as explanatory factors.

REFERENCES


