PERCEPTUAL AND COGNITIVE FACTORS AFFECTING SPEECH UNDERSTANDING

Bruce Schneider
Department of Psychology
University of Toronto

As Figure 1 illustrates, to understand speech in noisy situations, listeners must not only "extract" the speech signal from the background noise, but also determine the meaning of the words and phrases in the extracted signal, integrate this information with past knowledge, and store it for future use. Psychoacousticians, audiologists, and auditory physiologists typically have directed their efforts towards understanding the perceptual processes that are involved in signal extraction. Linguists and cognitive scientists, on the other hand, have focussed their efforts on how the words and phrases in the speech signal are processed linguistically, and how this processed information is integrated with past knowledge and stored in memory for future use. The endeavours of both groups have enhanced our understanding of how listeners process speech signals. However, for the most part, both groups have ignored the complex and often subtle interactions (indicated by the presence of a feedback loop in Figure 1) that occur between these two levels of processing.

To illustrate why such interactions cannot be ignored, consider the difficulties that the elderly often experience when attempting to understand speech in noisy conversational settings. These difficulties cannot be attributed solely to changes in auditory sensitivity because clinical tests of the hearing acuity of many of these individuals indicate that their hearing status is in the normal range. It could be that these clinical tests are insensitive to the perceptual deficits that are responsible for these comprehension difficulties, or that these difficulties are a consequence of age-related changes in more central or cognitive processes. What is more likely, however, is that age-related declines in bothperceptual and cognitive factors interact to produce comprehension difficulties.

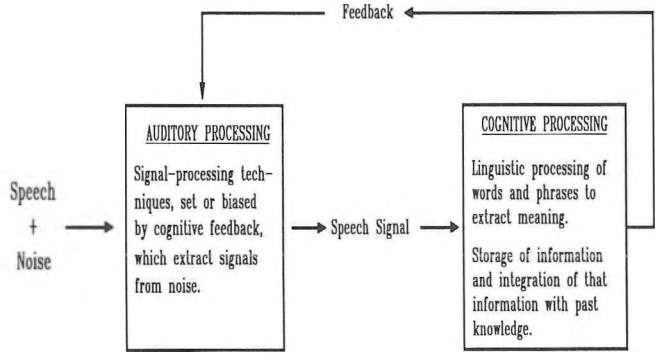

Figure 1. Simplified model of speech processing.

Age-related changes in perceptual factors such as the ability of listeners to use binaural cues to unmask signals in noise (Pichora-Fuller & Schneider, 1991, 1992), or the ability to temporally resolve two sounds (Schneider, Pichora-Fuller, Kowalchuk, & Lamb, 1994; Schneider, Speranza, & Pichora-Fuller, 1994) could contribute to difficulties in understanding speech. For example, old subjects with good hearing appear to have poorer temporal resolution than young subjects. Young and old subjects were asked to discriminate between two tone pips and a continuous tone of the same total energy and duration (Schneider, Pichora-Fuller, Kowalchuk, & Lamb, 1994, Schneider, Speranza, & Pichora-Fuller, 1994). In general, the size of the temporal gap required for discrimination was almost twice as large in older subjects. Moreover, as Figure 2 indicates, these thresholds are independent of the subject's audiometric threshold. Thus there are significant age-related declines in temporal acuity that are unrelated to the degree of sensorineural hearing loss. Because speech understanding requires listeners to process the rapid fluctuations in amplitude that are characteristic of speech, declines in temporal acuity may contribute to the speech understanding difficulties of the elderly by effectively degrading the speech signal and producing errors in speech perception.

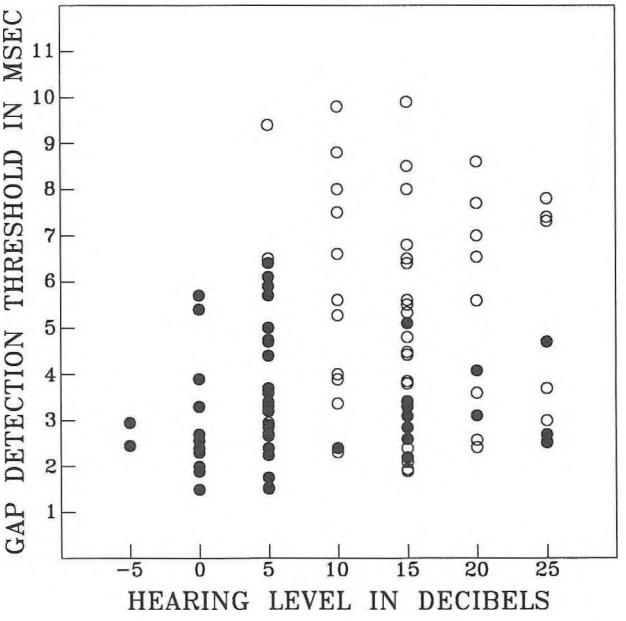

Figure 2. Gap detection thresholds for young (filled circles) and old (unfilled circles) subjects as a function of their hearing level.

Although age-related declines in perceptual processing may lead to errors in speech perception, part or all of the information lost in the early stages of auditory processing may be partially or completely recovered by higher-order processes. Listeners who "mishear" a word in a sentence can often recover the word from the context provided by the sentence. If the elderly, because of age-related declines in perceptual processing, are "mishearing" more of the words and phrases in everyday conversational settings, then they are forced to rely more often on context than their younger counterparts. Indeed, the elderly require a higher signal-to-noise ratio than the young in order to recognize words that cannot easily be recovered by the sentence context (Pichora-Fuller, Schneider, & Daneman, submitted). Therefore, in

many conversational settings, the elderly have to depend on context whereas the young do not. As a result, speech processing is more effortful for them, and there is some evidence that they are more effective than young adults in utilizing the contextual information (Pichora-Fuller, Schneider, & Daneman, submitted), perhaps because they are more often forced to rely on it.

When the elderly have to use context to recover information lost or distorted by external or internal noise, it is likely that they do so at a cost. Although cognitive psychologists have seldom considered the toll placed on cognitive resources when listening occurs under degraded conditions, it is conceivable that when the signal is degraded, more resources will be required for listening, which might deplete the cognitive resources available for linguistic and cognitive processing. For example, speech understanding draws heavily on working memory, a system responsible for both the processing and the temporary storage of information during the performance of complex cognitive tasks (Baddeley & Hitch, 1974; Daneman & Carpenter, 1980, 1983; Craik et al., 1990). In order to integrate successively heard words, phrases, and sentences into a coherent representation, listeners must have access to the results of earlier processes. In addition to having access to previously stored information they must also be able to simultaneously manipulate the stored material during ongoing processing. If, in addition, to these tasks, working-memory resources are required to remove the ambiguity and recover the information in the signal that has been lost during perceptual processing, comprehension would suffer. For example, diverting cognitive resources to the task of recovery, may make them less efficient at storing information in memory (Pichora-Fuller, Schneider, and Daneman, submitted). Thus, perceptual deficits can have serious consequences for the linguistic and cognitive processing of speech.

If, in addition to perceptual deficits, the elderly also experienced difficulties when forced to divide their processing resources between the task of recovery of lost information, and comprehension of the message, they would be particularly disadvantaged. In such a case, perceptual and cognitive deficits would interact in such a way as to make it very difficult for some of the elderly to understand speech in everyday listening situations.

If, in order to function well in ordinary conversational settings, the elderly must divert cognitive resources to the recovery of information lost during perceptual processing, they will have fewer resources available for higher-order cognitive functions. Thus, they might not be as efficient or as fast at integrating this incoming information with past knowledge, or at storing it for future use, and therefore appear to have a comprehension deficit. At the very least, listening is likely to be more effortful for elderly listeners which might explain why some of them often observe that it is not so much that they cannot understand what is being said in these settings, but that they find it very fatiguing to do so.

Thus, in order to effectively study the speech understanding difficulties of the elderly, we are forced to consider the speech processing system as a whole because of the subtle interactions that can occur between perceptual and cognitive levels of processing when one is listening to speech..

References

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G.A. Bower (Ed.) *The Psychology of Learning and Motivation*, Academic: NY.

Craik, F.I.M., Morris, R. G., & Gick, M.L. (1990). Adult age differences in working memory. In G. Vallar and T. Shallice *Neuropsychological impairments of short-term memory*. Cambridge University Press, Cambridge.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior. 19*, 450-466.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. *Journal of Experimental Psychology: Learning , Memory, and Cognition, 25*, 1-18.

Pichora-Fuller, M. K., & Schneider, B. A. (1991). Masking-level differences in the elderly: A comparison of antiphasic and time-delay methods with burst and with continuous masking noise. *Journal of Speech and Hearing Research, 34*, 1410-1422.

Pichora-Fuller, M. K. & Schneider, B. A. (1992). The effect of interaural delay of the masker on masking-level differences in young and elderly listeners. *Journal of the Acoustical Society of America, 91*, 2129-2135.

Pichora-Fuller, M. K., Schneider, B. A., & Daneman, M. (submitted). How young and old adults listen to and remember speech in noise.

Schneider, B. A., Pichora-Fuller, M.K., Kowalchuk, D., & Lamb, M. (1994). Gap detection and the precedence effect in young and old adults. *Journal of the Acoustical Society of America, 95*, 980-991.

Schneider, B. A., Speranza, F, & Pichora-Fuller, M. K. (1994). Age-related changes in temporal resolution: Envelope and intensity effects. Paper presented at the XXII International Congress of Audiology, Halifax, Nova Scotia.