THE PERCEPTION OF SPOKEN LANGUAGE BY ELDERLY LISTENERS: CONTRIBUTION OF AUDITORY TEMPORAL PROCESSES

Natalie Haubert and Kathleen Pichora-Fuller
University of British Columbia, 5804 Fairview Avenue, Vancouver, BC, V6T 1Z3

I INTRODUCTION

Older adults, even those with normal pure-tone hearing thresholds, commonly report difficulties understanding spoken language, particularly as the listening task becomes more challenging (CHABA, 1988). Fast speaking rate and background noise both make listening more challenging.

Numerous studies have established that older normal-hearing adults show decrements in the perception of spoken language presented in a background of competing noise (e.g. Gordon-Salant & Fitzgibbons, 1995; Stuart & Phillips, 1996). Age-related decrements in comprehension for fast speech rates have also been found in numerous studies (e.g. Letowski & Poch. 1996; Schmitt & Carroll, 1985; Wingfield, 1996). At the level of a single word, Price & Simon (1984) showed that rate can influence an elderly listener's ability to identify words where the distinguishing feature is a silent interval (e.g. 'rabid' vs. 'rapid').

A number of possible factors have been proposed to account for older adults' difficulties understanding spoken language. Some researchers believe that small changes in pure-tone high frequency sensitivity account for the poorer speech perception abilities of older adults (e.g., Humes, 1996). Others propose that changes in cognitive ability are responsible for these language processing deficits (e.g., Wingfield, 1996). Hypotheses based on auditory processing deficits, other than threshold hearing loss, constitute another plausible position.

One particular auditory processing ability of interest is temporal resolution. The ability of the auditory system to accurately analyze events that occur over time is critical to the understanding of spoken language. In fact, a large portion of sentence information is carried solely by temporal structure (e.g., Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995).

Gap detection thresholds provided one means of quantifying temporal resolution ability. In regards to older adults, many studies do indicate that they have reduced temporal resolution ability when examined in relation to young listeners and that this impaired gap detection ability is unrelated to audiometric thresholds (e.g., Strouse, Ashmead, Ohde, & Grantham, 1998; Schneider, Pichora-Fuller, Kowalchuk, & Lamb, 1994). Furthermore, an important relationship has been demonstrated between the duration of the markers surrounding a gap and gap detection thresholds, with age effects being significant when the marker duration is short (e.g. 2.5 ms) but not when it is long (e.g. 500 ms) (Hamstra & Schneider, 1999).

Decreased ability to detect gaps seems likely to interfere with some aspects of speech perception, especially at fast speaking rates. To date, however, no clear relationship has been found between gap detection ability and speech perception in old age (e.g. Strouse, Ashmead, Ohde, & Grantham, 1998). It may be that some relationship does exist, but that the particular stimuli used in these studies did not allow for this potential connection to be observed.

The purpose of the present study is to examine gap detection thresholds and speech perception abilities for stimuli selected to be representative of the challenging listening conditions in question. We assume that there is a correspondence between the marker-gap-marker sequence in a gap detection experiment and the VCV sequence in speech where C is a stop consonant. Therefore, a clearer relationship between measures of gap detection ability and speech perception should be observed in challenging conditions when the marker duration is short and the speech rate is fast compared to less challenging conditions when the marker duration is long and the speech rate is slow.

II METHOD

A. SUBJECTS

Eight young listeners (20-30 years old) and eight older listeners (68-74 years old) were tested. All listeners had pure-tone audiometric thresholds equal to or better than 25 dB HL from .25 to 3 kHz. All listeners were native speakers of English.

B. PROCEDURES

Each listener completed three tasks in the following order: word identification in quiet, gap detection, and word identification in noise. All testing took place in a double-walled, sound-attenuating IAC booth with stimuli presented monaurally over TDH-39 headphones.

Word Identification in Quiet

Four word pair continuums consisting of 14 tokens each were tested in quiet: 'cash'-'catch', 'dish'-'ditch', 'slit'-'split', and 'soon'-'spoon'. The slow speech rate continuums were based on naturally-spoken tokens of 'cash', 'dish', 'slit', and 'soon', respectively. To create the other members of each continuum a silent interval was inserted into the words and systematically adjusted in length to progress from short (5-10 ms) to long (75-120 ms). As the length of the inserted silent interval increased, the other member of the stimulus continuum, e.g. 'catch', was approximated. A fast speech rate continuum was also created for each word pair based on the originally spoken token. The original tokens were compressed to about two thirds of their original duration by deleting portions from the steady-state portions of the signal. All deletions were made at the zero-crossings in the signal to avoid any waveform discontinuities. Using these edited original tokens, silent intervals were inserted into the words and the length of the interval adjusted to construct the remaining 13 members of each continuum.

Subjects were presented with these tokens in a single-interval two-alternative forced-choice paradigm. The listeners first heard the ordered continuum presented twice. This served to familiarize the listeners to the materials. The tokens were then presented ten times each in random order. This sequence of events was then repeated allowing for the collection of 20 responses to each member of each stimulus continuum by each subject. The order of presentation of the stimulus continuums was counterbalanced across listeners with each listener first listening to the slow speaking rate continuums and then the fast speaking rate continuums. The stimuli were all presented at 40 dB SL referenced to the listener's SRT.

Gap Detection

Following Schneider and Hamstra (1999), the materials consisted of 2-kHz tonal gap and non-gap stimuli created by summing a number of temporally distributed Gaussian envelopes. Five gap marker durations were assessed: 0.83, 5, 10, 80, and 400 ms. The rise and fall times of the stimuli was held constant with stimulus duration specified from the peak of the first Gaussian envelope to the peak of the final Gaussian envelope. In order to prevent off-frequency cues resulting from spectral splatter, the gap detection stimuli were all presented in an amplitude-modulated broadband noise constructed to simulate a 6-person babble (ICRA, 1997). The gap detection stimuli were presented monaurally at 90 dB SPL and the noise was presented at 85 dB SPL to the same ear.

The gap detection stimuli were presented in a two-interval forced-choice paradigm. The gap size was either increased or decreased according to a 3 down 1 up rule to determine the 79.7% point on the psychometric function (Levitt, 1971). The duration of the gap markers changed after each run in the same irregular order for each participant. This order was cycled through four times, with the gap detection threshold being established as the mean of the final three runs.

Word Identification in Noise

The testing of word identification in noise was based on each listener's word identification performance in quiet. Only two stimulus tokens from each stimulus continuum were selected for further testing in noise, one for which one member of the pair was identified correctly at least 80% of the time and one for which the other member of the pair was identified correctly at least 80% of the time. In the few cases where a listener did not reach this criterion level of performance in quiet, testing was not completed in noise for that particular stimulus continuum. Testing was conducted at three signal to noise levels: 10, -5, and –15 dB, presented in order from most to least favorable and with the signal level held constant at 40 dB SL referenced to the listener's SRT. The same noise that was used in the gap detection task was used for the word identification test in noise. Following a period of familiarization at each noise level, the selected tokens were presented 20 times each. The selected tokens were presented in the same order as the corresponding continuums had been presented in quiet. As in quiet, word identification performance was assessed using a single-interval two-alternative forced-choice design.

III RESULTS

A. WORD IDENTIFICATION IN QUIET

All young subjects were able to distinguish the word pairs in quiet at both rates; however, some old subjects were unable to distinguish the 'slit'-'split', 'soon'-'spoon', and 'dish'-'ditch' contrasts even in quiet, and especially for the fast rate. For both slow and fast speech rates, the word boundary for the old listeners occurred at longer gap durations than for the younger listeners. For both age groups the word boundary occurred at longer gap durations for the slow rate than for the fast rate. Significant age by gap duration effects on word identification were found for 'cash'-'catch' at the slow rate (Figure 1) and for all of the other pairs at the fast rate. These findings were confirmed by analyses of variance (ANOVA) conducted for each word pair at each rate.

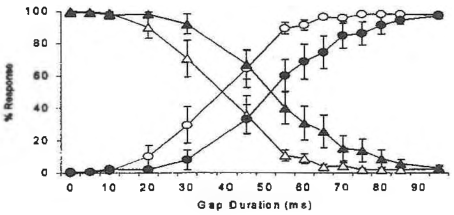

Figure 1: Word identification scores and standard error bars as a function of gap duration for 'cash'-'catch', slow rate. Unfilled triangles: young 'cash'; filled triangles: old 'cash'; unfilled circles: young 'catch'; filled circles: old 'catch'.

B. GAP DETECTION

As shown in Figure 2, the results of this study replicate prior findings by Schneider and his colleagues (Schneider & Hamstra, 1999; Schneider et al., 1994). At the two shortest marker durations (0.83 and 5 ms), the mean gap detection thresholds of the older listeners (mean = 7.399 ms and 6.073 ms, respectively) differed significantly from those of the younger listeners (mean = 2.938 ms and 2.227 ms respectively), as confirmed by an ANOVA [F(1,4)= 2.987, $p<.05$] and a Student-Newman-Keuls test ($p<0.01$).

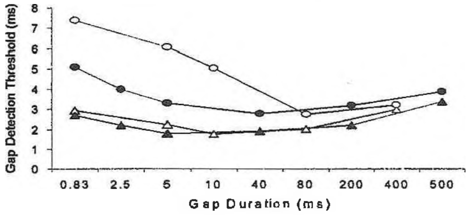

Figure 3: Gap detection thresholds as a function of marker duration for younger and older adults. Unfilled triangles: young adults, present study; filled triangles: young adults, Schneider and Hamstra (1999); unfilled circles: young adults, present study; filled circles: Schneider and Hamstra (1999).

C. WORD IDENTIFICATION IN NOISE

For all word pair contrasts, the addition of noise had a more adverse effect on word identification in the fast speech conditions than in the slow speech conditions. Specifically, word identification declined more for the member of the pair with the stop consonant. Age effects were not apparent for the 'cash'-'catch' and 'dish'-'ditch' pairs, but the old subjects performed more poorly than young subjects on the 'slit'-'split' and 'soon'-'spoon' pairs, especially at the fast rate (Figure 3).

D. CORRELATIONS

Overall, there was no clear pattern of correlations between audiometric thresholds and gap thresholds, consistent with the results of Schneider and his colleagues (Schneider & Hamstra, 1999; Schneider et al., 1994). Neither was there any clear pattern of correlations between the audiometric thresholds or the gap thresholds and word identification measures. The absence of clear patterns of significant correlations is not suprising given the small number of subjects in each age group.

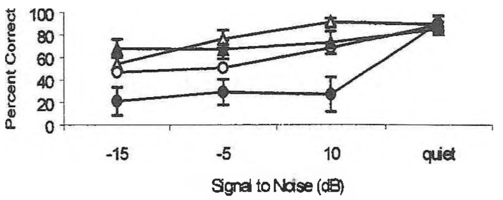

Figure 3: Word identification as a function of signal to noise ratio for 'slit' and 'split' fast speech rate. Unfilled triangles: 'slit' young adults; filled triangles 'slit' old adults; unfilled circles: 'split' young adults; filled circles: 'split' old adults.

IV DISCUSSION

The data from this study indicate that older adults are poorer than younger adults at detecting gaps embedded in non-speech stimuli (2 kHz markers) when the surrounding material is short. Correspondingly, older adults were also shown to have greater difficulty identifying speech when a gap served to differentiate two words, particularly for fast speech. Common to both these cases is the poorer performance evidenced by old listeners in the context of fast rates/ short duration stimuli. This pattern of findings is consistent with a hypothesis that proposes that recovery from neural adaptation occurs more slowly in the old compared to the young auditory system (Schneider & Hamstra, 1999). As the duration of the material surrounding the gap is shortened, less time is available for neural recovery and therefore fewer neurons are available to mark the onset of the material trailing the gap. With the response to the trailing material weakening, detection of the gap will become more difficult. If recovery from neural adaptation occurs more slowly in older adults, then it follows that they will be particularly disadvantaged when required to detect a gap in short duration non-speech stimuli. They would also face greater challenges with fast speech where the detection of a silent interval can influence stop gap detection and subsequent word identification.

V REFERENCES

CHABA. (1988). Speech understanding and aging. J of the Acoustical society of America, 83, 859-895.

Gordon-Salant, S., & Fitzgibbons, P.J. (1995). Recognition of multiply degraded speech by young and elderly listeners. J of Speech and Hearing Research, 38, 1150-1156.

Humes, L.E. (1996). Speech Understanding in the Elderly. J of the American Academy of Audiology, 7, 161-167.

ICRA. (1997). Noise Signals, Version 0.3. International Collegium of Rehabilitative Audiology, Hearing Aid Clinical Test Environment Standardization Work Group.

Letowski, T., & Poch, N. (1996). Comprehension of Time-Compressed Speech: Effects of Age and Speech Complexity. J of the American Academy of Audiology, 7, 447-457.

Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. J of the Acoustical Society of America, 49(2), 467-476.

Schmitt, J.F., & Carroll, M.R. (1985). Older listeners' ability to comprehend speaker-generated rate alteration of passages. J of Speech and Hearing Research,28, 309-312

Schneider, B.A., & Hamstra, S.J. (1999). Gap detection as a function of tonal marker duartion for younger and older listeners. J of the Acoustical Society of America, 106 (1), 371-380.

Schneider, B.A., Pichora-Fuller, M.K., Kowalchuk, D., & Lamb, M. (1994). Gap detection and the precedence effect in young and old adults. J of the Acoustical Society of America, 95(2), 980-991.

Shannon, R.V., Zeng, F-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. Science, 270, 303-304.

Strouse, A., Ashmead, D.H., Ohde, R.N., & Grantham, D.W. (1998). Temporal processing in the aging auditory system. J of the Acoustical Society of America, 104(4), 2385-2399).

Stuart, A., & Phillips, D.P. (1996). Word Recogntion in Continuous and Interrupted Broadband Noise by Young Normal-Hearing, Older Normal-Hearing, and Presbyacusic Listeners. Ear and Hearing, 17(6), 478-489.

Wingfield, A. (1996). Cognitive Factors in Auditory Performance: Context, Speed of Processing, and Constraints of Memory. J of the American Acedemy of Audiology, 7, 175-182).

Funded by NSERC, MRC, BCMSF.