Canadian Acoustics / Acoustique Canadienne 26(2) 25-28 (1998) ***Research article / Article de recherche***

THE BENEFIT OF CONTEXTUAL CUES FOR THE PERCEPTION OF TEMPORALLY DEGRADED SPEECH

Sharon M. Abel and Robert Najafee

Samuel Lunenfeld Research Institute
Mount Sinai Hospital
Toronto, Ontario, Canada M5G 1X5

ABSTRACT

The effect of contextual cues on the understanding of distorted speech was studied. Two groups of subjects, aged 20-25 and 50-59 years participated. Three lists of 50 sentences that were continuous or interrupted at rates of 8/s and 2/s were presented to each ear. In half the sentences within list, the final word was highly predicable from the context, and in half, poorly predicted. Final word recognition scores decreased with interruption, more so for the 2/s rate. Context was compensatory. Middle-aged subjects did not perform more poorly. Factors possibly accounting for this outcome are discussed.

SOMMAIRE

L'effet d'indices contextuels sur la compréhension de la parole dégradée a été étudié. Deux groupes de sujets, âgés respectivement de 20-25 et 50-59 ans, ont participé à cette étude. Trois listes de 50 phrases ont été présentées, à chacune des oreilles, de façon continue ou interrompue avec un taux de 8/s ou 2/s. Pour la moitié des phrases de chacune des listes, le dernier mot était hautement prévisible sur la base du contexte et, dans la seconde moitié, peu prévisible. La reconnaissance du mot final décroît avec l'interruption, et ce davantage pour le taux de 2/s. Le contexte s'est avéré compensatoire. Les sujets plus âgés ne performent pas moins bien que les plus jeunes. Les facteurs pouvant potentiellement expliqués ces resultats sont discutés.

1.0 INTRODUCTION

Published results from a number of studies support the conclusion that, with aging, there is a slowing of information processing ability (for reviews, see Birren and Schaie, 1977; Fitzgibbons and Gordon-Salant, 1996; Divenyi and Haupt, 1997). This effect is clearly apparent in the perception of degraded speech (Bergman, 1980; Wingfield, 1996). Poorer speech understanding observed in older listeners could be due to a variety of underlying factors, including deficits in the ability to use linguistic cues, memory, selective attention, reaction time, and frequency selectivity (e.g., Working Group on Speech Understanding, 1988; Hutchinson, 1989; Gordon-Salant and Fitzgibbons, 1997). Several experiments have confirmed that the difference limens for duration, frequency and intensity, as well as the interaural time of arrival difference required to perceive laterality increase with aging (Herman et al., 1977; Robin and Royer, 1989; Abel et al., 1990; Schneider et al., 1994; Fitzgibbons and Gordon-Salant, 1996; He et al., 1998).

In a benchmark experiment conducted by Bergman and co-workers in the mid-1970s (Bergman et al., 1976), the understanding of distorted speech was compared in listeners ranging in age from 20-79 years, screened for hearing loss. The effects of reverberation of the listening environment, a competing speaker, and simultaneous presentation of the two parts of two-part words to the right and left ears, as well as low and high pass filtering, time compression, and periodic interruption of the speech materials, were studied in relation to the undistorted condition. Decrements due to aging were observed to some degree under all conditions of distortion. However, the greatest changes in speech understanding occurred when time-processing abilities were stressed, that is, under conditions of reverberant listening, compression and interruption. The understanding of undistorted sentences decreased by about 15% over the age range studied, while periodic interruption resulted in a decrement of about 65% in the number of words correctly recognized. In the case of speech interruption, changes were evident as early as the fourth decade.

A number of studies have demonstrated that the age-related decrement in processing distorted speech is independent of any concurrent presbycusis (Sticht and Gray, 1969; Dubno et al., 1984; Gordon-Salant and Fitzgibbons, 1993). In the study by Sticht and Gray (1969), subjects under and over the age of 65 years, subdivided

into groups with either normal or mild to moderate bilateral sensorineural hearing loss, were tested with time compressed speech. Younger and older subjects in each of the two hearing categories were matched in their ability to understand undistorted speech. Hearing loss, aging and percent compression of the speech materials were all statistically significant factors. Importantly, the effects of age and hearing loss were independent.

2.0 RATIONALE

The present experiment was conducted to explore whether contextual cues might be used to counteract the effect of speech degradation. Miller and colleagues (Miller and Licklider, 1950; Miller et al., 1951) showed that, in young listeners, word recognition scores improved with familiarity of the test items. Scores decreased as the number of possible alternatives increased. More recently, Hutchinson (1989) compared subjects in their 20s and 60s on their ability to understand words in sentences with either good or poor contextual cues, presented against a babble noise background. Both groups were aided, almost equally, by the availability of contextual cues. Older subjects were more severely affected by the absence of context, particularly as the speech to noise ratio decreased.

3.0 EXPERIMENTAL DESIGN

In the present study, speech degradation was accomplished by means of periodic interruption. This method was chosen because of the demonstration that it provides the earliest indicator of the age effect (Bergman, 1983). Electronic interruption of speech has previously been investigated in detail by Miller and Licklider (1950) for young normal listeners. In their study, monosyllabic words were interrupted at regularly spaced intervals ranging from 0.1/s to 10,000/s, with an on-off fraction (duty cycle) of 0.5. Word recognition improved from 50% at 0.1/s to near 100% at 100/s. Given a constant interruption rate, performance improved as the speech on time increased.

The study by Bergman et al. (1976) utilized an interruption rate of 8/s. In later work, Gordon-Salant and Fitzibbons (1993) were unable to show an aging effect with an interruption of 12.5/s. In the present experiment, two values were chosen, 8/s and 2/s, to maximize the likelihood that time-processing abilities would be stressed. Performance under these two conditions was compared with continuous undistorted speech. Speech materials were presented in quiet to preclude possible confounding by age-related differences in the perception of acoustic cues with masking (Hutchinson, 1989). Right and left ears were tested independently. The nontest ear was fitted with a sound attenuating ear plug. Previous research by Johnson et al. (1979) showed an age-related decline in dichotic memory for spoken digits presented to the left ear but not the right ear. This difference was attributed to hemispheric differences in the encoding of language.

4.0 METHODS AND MATERIALS

4.1 Subjects

Two groups of eight subjects, aged 20-25 years (1 male and 7 females) and 50-59 years (2 males and 6 females) participated. All had normal hearing thresholds of less than 10 dB HL in both ears at 2 kHz. This frequency has been shown to be a good predictor of speech perception in cases of mild hearing loss (Abel, 1993).

4.2 Apparatus

The apparatus has been described previously (Abel et al., 1990). Each subject was tested in a sound proof IAC booth with ambient noise levels less than the maximum recommended for headphone testing (ANSI, 1991). The SPIN test (Bilger et al., 1984) used for the experiment was commercially available on audio cassette. Lists of sentences pre-recorded by a male American speaker in quiet, were presented to the subject monaurally over a TDH-39 matched headset by means of a Nakamichi Bx-125 Cassette Deck. A sustained 1-kHz tone at the start of the tape, recorded at a level equivalent to the rms value of the speech, allowed calibration to a comfortable listening level of 75 dB SPL. The interuption rate of the speech (2/s or 8/s) was controlled by means of a Hewlett-Packard 3325A Synthesizer/Function Generator. The duty cycle was fixed at 0.5. The rise/decay time of speech segments was 10 ms.

4.3 Procedure

Eight alternative lists were available. Each comprised 50 sentences. In half the sentences, the final word was highly predictable from the preceding context (e.g., THE SLEEPY CHILD TOOK A NAP), and in half, the final word was poorly predictable (e.g., THE CLASS SHOULD CONSIDER THE FLOOD). Each subject was presented six lists, three to the right ear and three to the left, with either no interruption (i.e., continuous speech) or with interruption rates of 2/s and 8/s, respectively. The continuous condition was always given first to maximize practice. The subject's task was to write down the final word in each sentence. Guessing was encouraged, in case of uncertainty. In half the subjects within each group, the right ear was tested first, and in half, the left. The order of the two interruption rates and choice of list for each of the six listening conditions was counter-balanced across subjects in the group. No list was presented twice within subject to control familiarity.

5.0 RESULTS

The results of the experiment are presented in Table 1. Mean correct high and low context final word recognition scores, each based on 25 sentences, are shown for each of the continuous and interrupted conditions by group and

ear. A nested analysis of variance with repeated measures on ear, context, and interruption rate indicated that context, interruption rate and their interaction were statistically significant ($p<0.001$). There was no effect of age or ear. Post hoc pairwise comparisons using Fisher's LSD method (Daniel, 1983) confirmed that for the high context sentences, there was no difference between speech presented continuously or with an interruption rate of 8/s. Both conditions gave significantly higher scores than the 2/s condition. In the case of the low context sentences, the three levels of interruption were significantly different from each other. For both the 8/s and 2/s rates, scores were significantly higher with high compared with poor contextual cues ($p<0.05$).

6.0 DISCUSSION

The results of the experiment indicated that interrupted speech understanding in subjects aged 50-59 yrs was no different than that of subjects aged 20-25 yrs. This is in contrast to Bergman's (1976) finding that the effect of degrading speech in the time domain was evident by the fourth decade. It is possible that the negative outcome was due to the difference in speech materials and task in the two studies, or the small sample size and possible lack of sensitivity of the test in the present study. In contrast, the significance of rate of interruption and contextual cues were clearly apparent. For the two groups combined, the availability of context significantly improved the outcome by 26% in the case of the 8/s interruption rate and by 19% for the the 2/s interruption rate. Collapsed across the two levels of context, mean percent correct scores, 78% and 38% for 8/s and 2/s respectively, were comparable to the results reported by Miller and Licklider (1950).

Findings of Wingfield and coworkers (see Wingfield, 1996) may shed light on the nature of the context effect. Wingfield et al. (1985) presented subjects, aged 18-22 yrs and 65-73 yrs, with normal sentences, syntactic word strings (i.e., word strings without meaning but with syntactic form), and random word strings. The percentage of words correctly recognized decreased, as the presentation rate increased from 275 to 425 words/min. Both groups achieved at least 80% correct, if given normal or syntactic strings but performance deteriorated, particularly for the older group, when syntax was eliminated. Thus, when time processing ability was stressed, both groups performed well, as long as syntactic cues were available. In the present study, language structure was always preserved, regardless of the interruption rate. Both groups gained from the linguistic cues provided by context.

ACKNOWLEDGMENTS

Supported in part by a Scientist award (SMA) from the Saul A. Silverman Family Foundation. The authors thank an anonymous reviewer for comments on an earlier draft.

REFERENCES

Abel, S.M. (1993). "The development of speech communication capability tests." Contract report, DCIEM, North York, Ont., Canada.

Abel, S.M., Krever, E.M. and Alberti, P.W. (1990). "Auditory detection, discrimination and speech processing in ageing, noise-sensitive and hearing-impaired listeners." Scand. Audiol. 19, 43-54.

ANSI S3.1-1991. (1991). "Maximum permissible ambient noise levels for audiometric test rooms." American National Standards Institute, New York, NY.

Bergman, M. (1980). Aging and the Perception of Speech. Baltimore: University Park Press.

Bergman, M., Blumenfeld, V.G., Cascardo, D., Dash, B., Levitt, H. and Margulies, M.K. (1976). "Age-related decrement in hearing for speech." J. Gerontol. 31, 533-538.

Bilger, R.C., Nuetzel, J.M., Rabinowitz, W.M. and Rzeczkowski, C. (1984). "Standardization of a test of speech perception in noise." J. Sp. Hear. Res. 27, 32-48.

Birren, J.E. and Schaie, K.W. (eds). Handbook of the Psychology of Aging. New York: van Nostrand Reinhold.

Daniel, W.W. (1983). Biostatistics: A Foundation for Analysis in the Health Sciences, 3rd ed. New York: Wiley.

Divenyi, P.L. and Haupt, K.M. (1997). "Audiological correlates of speech understanding deficits in elderly listeners with mild-to-moderate hearing loss. I. Age and lateral asymmetry effects." Ear & Hear. 18, 42-61.

Dubno, J.R., Dirks, D.D. and Morgan, D.E. (1984). "Effects of age and mild hearing loss on speech recognition in noise." J. Acoust. Soc. Am. 76, 87-96.

Fitzgibbons, P.J. and Gordon-Salant, S. (1996). "Auditory temporal processing in elderly listeners." J. Am Acad. Audiol. 7, 183-189.

Gordon-Salant, S. and Fitzgibbons, P.J. (1993). "Temporal factors and speech recognition performance in young and elderly listeners." J. Sp. Hear. Res. 36, 1276-1285.

Gordon-Salant, S. and Fitzgibbons, P.J. (1997). "Selected cognitive factors and speech recognition performance among young and elderly listeners." J. Sp. Lang. Hear. Res. 40, 423-431.

He, N., Dubno, J.R. and Mills, J.H. (1998). "Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects." J. Acoust. Soc. Am. 103, 553-565.

Herman, G.F., Warren, L.R. and Wagener, J.W. (1977). "Auditory lateralization: Age differences in sensitivity to dichotic time and amplitude cues." J. Gerontol. 32, 187-191.

Hutchinson, K.M. (1989). "Influence of sentence context on speech perception in young and older adults." J. Gerontol.: Psycholog. Sc. 44, P36-P44.

Johnson, R.C., Coles, R.E., Bowers, J.K., Foiles, S.V., Nikaido, A.M., Patrick, J.W., and Woliver, R.E. (1979). Hemisphere efficiency in middle and later adulthood." Cortex 15, 109-119.

Miller, G.A., Heise, G.A. and Lichten, W. (1951). "The intelligibility of speech as function of the context of the test materials." J. Exp. Psychol. 41, 329-335.

Miller, G.A. and Licklider, J.C.R. (1950). "The intelligibility of interrupted speech." J. Acoust. Soc. Am. 22, 167-173.

Robin, D.A. and Royer, F.L. (1989). "Age-related changes in auditory temporal processing." Psychol. and Aging 4, 144-149.

Schneider, B.A., Pichora-Fuller, M.K., Kowalchuk, D. and Lamb, M. (1994). "Gap detection and the precedence effect in young and old adults." J. Acoust. Soc. Am. 95, 980-991.

Sticht, T.G. and Gray, B.B. (1969). "The intelligibility of time compressed words as a function of age and hearing loss." J. Sp. Hear. Res. 12, 443-448.

Wingfield, A. (1996). "Cognitive factors in auditory performance: context, speed of processing, and constraints of memory." J. Am. Acad. Audiol. 7, 175-182.

Wingfield, A., Poon, L.W., Lombardi, L. and Lowe, D. (1985). "Speed of processing in normal aging: effects of speech rate, linguistic structure, and processing time." J. Gerontol. 40, 579-585.

Working Group on Speech Understanding and Aging (1988). "Speech understanding and aging." J. Acous. Soc. Am. 83, 859-893.

Table 1
Effects of age, interruption rate, ear and contextual cues on word recognition.

Age (yrs)	Interruption Rate	Context/Ear: High Left	High Right	Low Left	Low Right
20 - 25	Cont.	24.9 (0.4)	25.0 (0.0)	23.9 (0.8)	24.4 (1.2)
	8/s	23.5 (1.5)	23.6 (2.1)	16.3 (3.0)	17.0 (2.1)
	2/s	11.8 (5.1)	13.9 (5.2)	7.6 (2.5)	7.5 (3.0)
50 - 59	Cont.	24.9 (0.4)	25.0 (0.0)	23.8 (1.2)	23.1 (1.4)
	8/s	21.1 (4.2)	22.6 (2.6)	15.3 (3.5)	16.4 (2.1)
	2/s	10.9 (4.4)	10.9 (3.4)	5.5 (2.9)	6.9 (2.6)

Mean raw score (1 S.D.), 25 items