Can context diminish the effects of rapid speech recognition in older listeners?

Izumi Furukawa, M.A., Nancy Vaughan, Ph.D., Daniel Storzbach, Ph.D.

VA RR&D National Center for Rehabilitative Auditory Research, Portland, Oregon

Introduction

One of the common complaints among older listeners is that they have difficulty understanding speech especially when it is too fast. However, the cause of this difficulty has not been adequately explained. It is known that hearing loss accompanies aging, but this type of difficulty occurs even among normal hearing listeners. Studies have shown that decline in certain cognitive processes with age, such as speed of processing and working memory, contribute to speech understanding difficulty in older adults. (Brebion, 2001, Wingfield, 2000)

The results represented here are a part of a larger study investigating the effects of age-related cognitive deficits on speech perception among older adults. Time-compressed speech was used to evaluate the effects of age on rapid speech recognition with two different types of speech materials, contextual and non-contextual. The hypothesis was that time-compression increases demand on working memory and on processing speed. The study component reported here analyzed whether context improved performance on time-compressed sentence recognition tasks at various rates of speech.

Methods

Subjects

1. Age: between 50 and 75 years old
2. Two groups based on hearing status

Normal Hearing: Less than or equal to 25 dB HL at all octave frequencies 250 through 4000 Hz.

Mild to Moderate Hearing Loss (HL): Less than or equal to 40 dB through 1000 Hz, to 60 dB HL at 2000 Hz, and to 70 dB at 4000 Hz.

Test materials

1. IEEE sentences (IEEE, 1969). (contextual)
   Example: The bark of the pine tree was shiny and dark.
2. Anomalous sentences (syntactically correct but semantically anomalous)
   Example: Run a bell while feeding a truck to a button.

Time-compression(TC):

All speech materials were time-compressed using the synchronized overlap–add (SOLA) algorithm adapted from Hardam (1990). See poster 4aPP10 by Akiko Kusumoto for more details.

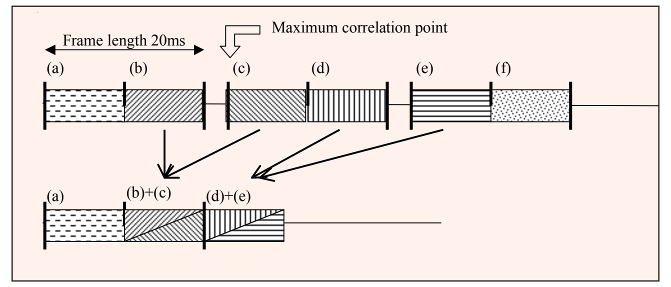

Figure 1a. Schematic of the time-compression algorithm developed at the VA RR&D NCRAR. The software program identified an initial 20 ms frame (a and b) at the beginning of the input signal. The first half (a) of the frame remained unchanged while the latter half (b) of the frame was overlapped onto the next 10 ms of the input signal. The overlapping segment was then moved along to the right sample by sample until a maximum correlation was achieved between the amplitudes of the overlap. At the point of maximum correlation within the 20 ms window, the two overlapping segments were added together with a weighting factor to control for amplitude growth. This process continued for each 20 ms frame.

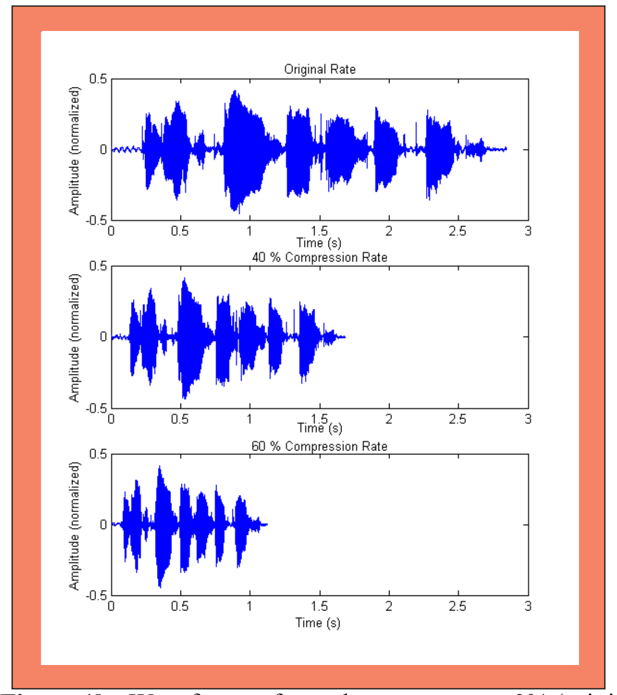

Figure 1b. Waveforms of a spoken sentence at 0% (original rate), and time-compressed at 40%, and 60%.

Presentation:

- Sentences were presented in the right ear or in the better ear (based on the pure tone average of 1k, 2k, and 4k Hz) through the MedRx Otowizard monaurally via an ER-3 insert earphone.
- The presentation level for normal hearing subjects was 90 dBA.
- When the better ear thresholds were greater than 25 dB HL at any one of test frequencies between 250 and 4k Hz, the master hearing aid (MHA) function on the Otowizard was used with NAL-RP formula. The volume on the Otowizard was kept at the same level but the actual presentation levels were varied depending on the degree of hearing loss of each subject.
- All odd number subjects started with IEEE sentences and even number subjects started with Anomalous sentences.
- Five practice sentences with feedback and five practice sentences without feedback were presented at the beginning of each session for each new time-compression rate.

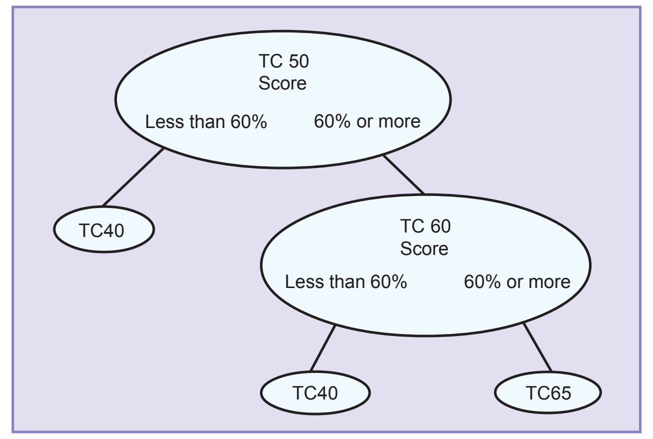

Figure 2. Diagram of adaptive testing method.

Adaptive Method

All subjects were tested at the original rate (TC00) at each material first. Testing was the performed using an adaptive method. Testing was started at 50% time-compression (TC50) and speeded up until the subject scored less than 60% correct key words. If the score was less than 60% correct at TC50, testing was continued at the slower rate (TC40). All the subjects were tested either 2 or 3 different rates in addition to original rate. Ten sentences were presented at each TC rate.

Results

Subjects

A total of 160 subjects (133 male, 27 female) have participated in the study to date. Mean age of the entire group was 60.9 (SD=6.448). Sixty-four out of 160 (40%) had normal hearing in the tested ear. Ninety-six (60%) had mild to moderate hearing loss and used the MHA. The mean age of the normal hearing group was 58.1 years (SD=5.755) and the mean age of the hearing loss group was 62.7 years (SD=6.244). Figure 3 shows the mean thresholds of the tested ear for the normal hearing group, the mild to moderate hearing loss group, and the mean thresholds obtained with the MHA for the mild to moderate hearing loss group.

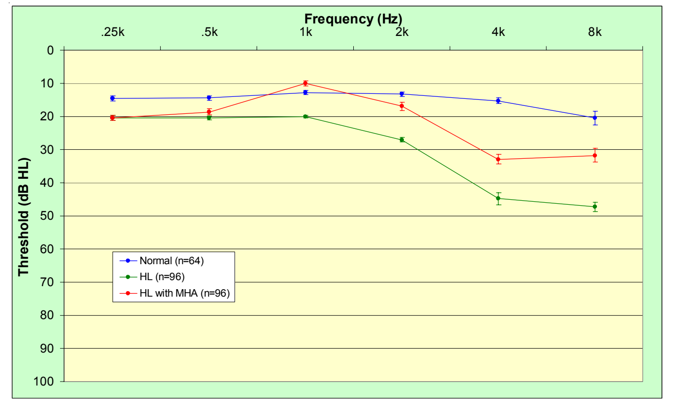

Figure 3. Mean thresholds and standard errors of the tested ear for normal hearing and hearing loss groups and for the hearing loss group with MHA.

Descriptive

Table 1 shows the mean scores and standard deviation for each subject group at each rate for both test materials. The scores decreased as the TC rate increased for both groups and for both materials. The normal hearing group scored better than the hearing loss group at each rate for both test materials. IEEE sentences (Contextual) scores were better than Anomalous sentences (non-contextual) except at the fastest rate (TC65) for both the normal and the hearing loss groups.

Table1. Means and Standard deviations of TC score (percent correct key word) for normal and mild to moderate hearing loss groups for IEEE and Anomalous sentences at each time-compression rate.

	Normal		Hearing Loss	
	IEEE	Anomalous	IEEE	Anomalous
TC00	99.38 ±1.28	97.56 ±3.71	99.04 ±1.67	95.52 ±5.30
TC40	96.14 ±4.19	92.88 ±5.56	92.90 ±7.38	89.13 ±10.47
TC50	92.00 ±6.45	88.41 ±7.40	82.06 ±17.23	78.63 ±15.02
TC60	80.78 ±13.32	76.44 ±13.04	68.68 ±16.99	64.64 ±17.76
TC65	50.28 ±16.34	51.46 ±13.26	39.31 ±15.18	42.68 ±12.69

ANCOVA

Table 2 shows the result of the Analysis of covariance (ANCOVA). With age as a significant covariate, there were significant main effects of Material, Group, and Rate. Only the two-way interactions between Group and Rate, and between Rate and Material were significant.

Table2. The results of Analysis of covariance (ANCOVA) comparing Material, Group, and TC Rate with age as a covariate.

Source	DF	F-ratio	p-value
Age (Covariate)	1	65.18	0.000000*
Material	1	9.99	0.001608*
Group	1	45.24	0.000000*
Rate	4	726.44	0.000000*
Material x Group	1	0.00	0.974254
Group x Rate	4	10.56	0.000000*
Rate x Material	4	2.94	0.019725*
Material x Group x Rate	4	0.21	0.933597

Main Effect of Material:
The IEEE sentence (contextual) scores were significantly better than the Anomalous sentence (non-contextual) scores across rate and group.

Main Effect of Group:
The normal hearing group performed significantly better than the mild to moderate hearing loss group across rate and material.

Main Effect of Rate:
The TC scores significantly decreased as the TC rate increased (faster) across material and group.

Group x Rate Interaction:
The post hoc Newman-Keuls Multiple-Comparison Test showed that the normal hearing group performed significantly better than the mild to moderate hearing loss group at TC50, TC60, and TC65 (faster rates). However there were no significant differences between two groups at TC0 and TC40 (slower rates).

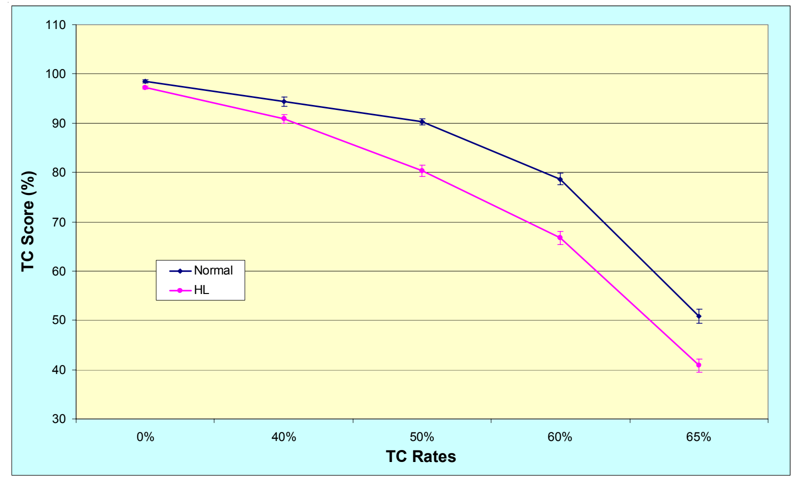

Figure 4. The means and standard errors of TC scores at each TC rate by group.

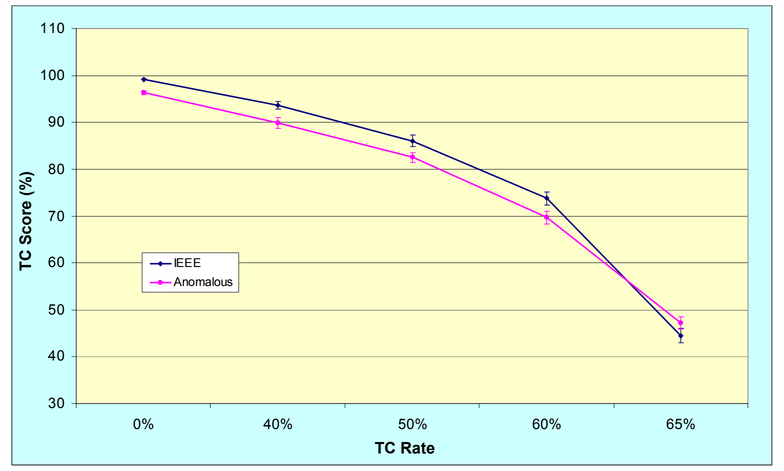

Figure 5. The means and standard errors of TC scores at each TC rate by material.

Material x Rate Interaction:
The post hoc Newman-Keuls Multiple-Comparison Test showed that the TC scores of the IEEE sentences were significantly better than the Anomalous sentence scores at TC50 and TC60. However, there were no significant differences between IEEE and Anomalous sentence scores at TC0, TC40 (slower rates), and TC65 (fastest rate).

Regression Analysis:
In order to determine the amount of variance of the sentence recognition scores that is accounted for by age and hearing loss, stepwise regression analyses were performed for each combination of sentence type and time-compression rate (Figure 6). Hearing loss was defined as the pure tone average of 0.5, 1, 2, and 4 kHz of the tested ear (PTA4) for purposes of these analyses.

At the TC00 and TC40, mean sentence recognition scores were quite high (greater than 90% correct). Age did not account for any of the variance in the TC scores at these slower rates. At TC50, TC60, and TC65, on the other hand, age accounted for between 3% and 10% of the variance. Age was more important at TC60 for both speech materials than at any other TC rate. As expected, hearing loss accounted for most of the variance in this analysis. However, age and hearing loss combined explained less than 29% of the total variance.

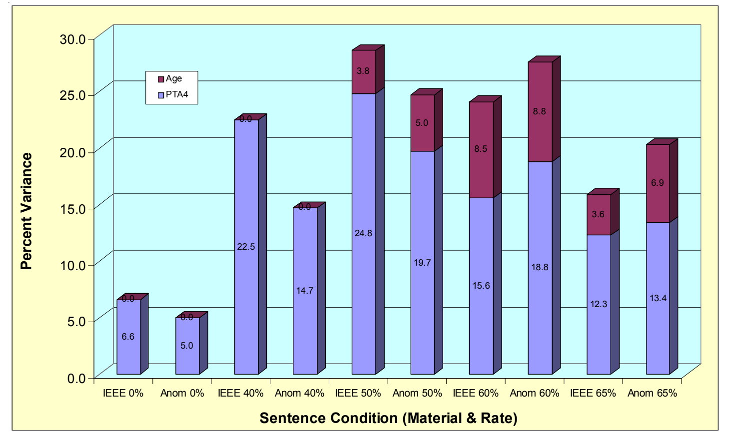

Figure 6. The results of regression analysis using hearing loss (PTA4) and age as the predictor variables for each sentence condition represented by the combination of test materials and TC rates.

Discussion

Performance on the IEEE sentences (contextual) was significantly better than the Anomalous (non-contextual) sentences statistically, but the difference was not clinically substantial, at the same time compression rates (TC50 and TC60). This finding suggests that older listeners benefit marginally from contextual cues with faster rates of speech. The average Anomalous sentence score was better than the average IEEE sentence score at the most difficult rate (TC65) though the difference was not statistically significant. This may be a result of subjects' willingness to guess on the Anomalous sentences more than on the IEEE sentences since the Anomalous sentences did not need to be meaningful. The difference between IEEE and Anomalous sentences scores were also not statistically significant at easier rates (TC00 and TC40). The scores at these rates were high for both materials probably reflecting ceiling effects.

The normal hearing group scores were significantly better than those of the mild to moderate hearing loss group at faster speech rates (TC50, TC60, and TC65). The scores were not significantly different between the two groups at slower speech rates (TC00 and TC40). These slower rates were easy for both the hearing impaired and normal hearing listeners.

As expected, a stepwise regression analysis using age and hearing loss as the predictor variables revealed that hearing loss accounted for most of the variance of the sentence scores. Age accounted for 3 to 9% of the variances across the different conditions. As speech rate increased, the amount of variance accounted for by age increased also. These findings suggest that age plays a larger role as sentence difficulty increases due to rate. These data are part of a larger on-going study investigating the effects of age-related cognitive changes on speech perception in older listeners. It is expected that further analysis of cognitive variables will provide an explanation for some of the remaining variance.

References

Brebion, G. (2001). Language processing, slowing, and speed/accuracy trade-off in the elderly. *Exp Aging Res, 27*(2), 137-150.

Hardam, E. (1990). *High quality time scale modification of speech signals using fast synchronized-overlap-add algorithms*. Paper presented at the International Conference on Acoustics, Speech, and Signal Processing.

IEEE. (1969). *IEEE Recommended Practice for Speech Quality Measurements*. New York: Institute Electrical and Electronic Engineers, Inc.

Wingfield, A. (2000). Speech perception and the comprehension of spoken language in adult aging. In D. C. Park & N. Schwarz (Eds.), *Cognitive Aging* (Vol. 13, pp. 175-195). Philadelphia, PA: Psychology Press/Taylor & Francis.

Acknowledgements

This work was supported by the Department of Veterans Affairs Office of Research and Development Grant C2832R and the VA RR&D National Center for Rehabilitative Auditory Research. The authors wish to acknowledge Craig M. Dennis, Computer Specialist, for his significant contribution to this presentation.