Effects of Age and Passage Difficulty on Listening-Rate Preferences for Time-Altered Speech

Arthur Wingfield and Julie L. Ducharme

Department of Psychology and Volen National Center for Complex Systems, Brandeis University, Waltham, Massachusetts.

When younger and older adults were allowed to adjust the speech rate of time-compressed and time-expanded speech passages, older adults tended to select as preferred rates significantly slower speech rates than the younger adults. Both age groups, however, selected slower rates for difficult speech passages (low cloze predictability) than for easy passages (high cloze predictability). Recall performance showed effects of speech rate and passage difficulty, with participants' recall at their selected speech rates comparable to their performance at slower rates. Results suggest that older adults are as effective as the young in their ability to monitor the difficulty of a speech passage as it is being heard and to moderate their listening rate selections accordingly.

TIME compression of natural speech has served as a valuable tool for examining upper limits on processing speed in young (Foulke, 1971), older (Wingfield, 1996), and neurologically impaired (Blumstein, Katz, Goodglass, Shrier, & Dworetsky, 1985) individuals. The most common method of time compression is the sampling method, in which small unnoticed segments are periodically deleted at regular intervals from the speech signal, with the remaining segments then abutted in time. By deleting segments from both word and silent durations, the compressed signal preserves the temporal patterning of the original speech. The method also preserves the original pitch contour and other features of the speech prosody. The degree of time compression is controlled by the frequency with which the segments are deleted. Speech can be expanded in time by reversing the process and iteratively resampling the signal so as to lengthen the duration of the speech elements and silent intervals (Foulke, 1971).

Older adults are known to find time-compressed speech very difficult to comprehend and to recall, a finding that holds even when young and older listeners are equated for pure tone hearing acuity (see Wingfield, 1996, for a review). Older adults also claim that they prefer to listen to slower speech rates (Obler, Fein, Nicholas, & Albert, 1991). This claim of listening preferences, however, has received little direct study. One way to test this claim would be to present listeners with time-compressed speech and allowing them to adjust the speech to a preferred listening rate. This technique has been used in the past with young adults to examine rate preferences after extended exposure to time-compressed speech (Gade & Mills, 1989).

In this experiment we used the rate adjustment technique to examine (1) age effects on selection of preferred speech rates, and (2) the question of whether, for both younger and older participants, preferred speech rates would be moderated by difficulty of the speech materials. An empirically derived measure of passage difficulty was used based on the average predictability of the speech, calculated using a cloze procedure: the percentage of individuals who can correctly guess the identity of words periodically deleted from a prose passage (Miller & Coleman, 1967). Prior use of such passages has shown that average cloze predictability of text correlates highly with individuals' subjective estimates of passage difficulty as well as actual recall scores (Acquino, 1969). Testing for listeners' sensitivity to the content difficulty of the speech materials is an important variable to ensure that rate preferences are a consequence of ease or difficulty of the processing task, rather than, for example, rate preferences being determined by the speech rhythm or other superficial feature of the speech signal.

Once participants' selected rates had been established, we then tested the validity of their selections by examining recall accuracy for passages presented at several speech rates, including the rate selected by that individual.

METHODS

Participants.—The older participants were 16 healthy older adults, 7 men and 9 women, with ages ranging from 66 to 81 years (M = 74.2). The young participants were 16 university undergraduates, 8 men and 8 women, with ages ranging from 17 to 22 years (M = 19.7). The older participants had an average of two more years of formal education than the younger participants [Older: 16 years (SD = 2.5); Younger: 14 years (SD = 1.0), t(30) = 3.20, p < .01] and a higher mean WAIS-R vocabulary score [Older: 71.8 (SD = 6.7); Younger: 66.4 (SD = 6.1), t(30) = 3.84, p < .05]. The two groups did not differ significantly on either forward [Older: M = 7.1 (SD = 1.1); Younger: 7.6 (SD = 1.3), t(30) = 1.33, n.s.] or backward digit spans [Older: 5.6 (SD = 0.9); Younger: 5.6 (SD = 1.3), t(30) < 1].

All participants received pure-tone audiometric screening, and all participants' hearing was within normal limits for their age group (ANSI, 1989; Morrell, Gordon-Salant, Pearson, Brant, & Fozard, 1996). Sound levels were adjusted to a comfortable listening level for each participant during a pre-test session and then maintained at that level for all conditions throughout the experiment.

Stimulus materials.—The stimuli consisted of 16 150-word passages taken from Miller and Coleman (1967). All of the passages consisted of meaningful prose but they differed in their...
level of average interword predictability. Eight of the passages were high-predictability passages (p = .35 - .42; M = .38) and eight were low-predictability passages (p = .24 - .34; M = .29). The passages were all read by a female speaker of American English in natural intonation at an average speech rate of 174 wpm.

Procedures

Speech rate selection.—For the rate-selection phase of the experiment each participant heard eight passages, four low-predictability and four high-predictability. The passages were presented on a Lexicon Varispeech II compressor/expander that included a rate control knob that could be operated by the participant. Turning the knob in one direction speeded the speech; turning the knob in the other direction slowed it. The stimuli were heard binaurally over headphones. Participants were tested individually in a sound-isolated test room. Two of the passages of each predictability level were begun at twice the originally recorded speech rate and two were begun at half the originally recorded speech rate. The participants' instructions were to use the rate-setting control on the Varispeech apparatus to adjust the rate to the point where he or she felt that the passage could be understood and accurately recalled. Participants were placed under no time pressure in completing their adjustments although all participants were able to select their desired rates approximately half way through the passage. (The particular passages heard with faster or slower initial rates were counterbalanced across participants such that, by the end of the experiment, each passage had been heard an equal number of times with slow or fast initial starting rates. Order of presentation of passages and conditions were also counterbalanced across participants.)

Recall performance.—Following the rate-selection session, each participant listened to an additional eight passages not used in the previous session. Two passages, one high-predictability and one low-predictability, were presented at each of three speech rates: 260 wpm (i.e., reproduction in 67% of the original playing time), 330 wpm (reproduction in 53% of original playing time), and 140 wpm (reproduction in 124% of the original playing time). Participants also heard one high-predictability passage presented at the mean speech rate they had selected for the high-predictability passages in the rate-selection session and one low-predictability passage presented at the mean speech rate they had selected for the low-predictability passages. As each passage ended participants were asked to recall the passage as accurately, and with as much detail, as possible. Responses were given aloud and recorded for later scoring.

Results

Both young and older participants responded to the rate adjustment task in a similar way, either slowing or speeding the speech depending on the speech starting rate. Final rate selection was typically preceded by a brief period of fine-tuning adjustments as the desired rate selection was completed. The left panel of Figure 1 shows the mean speech rates selected by the younger and older adults for the high- and low-predictability passages. For each case we show separately the speech-rate selections for slow (S; half original rate) and fast (F; twice original rate) initial starting speeds. The broken horizontal line indicates the 174 wpm speech rate of the original recordings. As can be seen, participants selected faster speech rates for high-predictability passages than for low-predictability passages, F(1,30) = 166.02, MSE = 110.46, p < .0001, the older adults selected slower speech rates than the younger adults, F(1,30) = 4.18, MSE = 826.93, p < .05, and participants generally selected faster speech rates when their initial starting rate was faster than the original rate than when the initial starting rate was slower than the original rate, F(1,30) = 20.10, MSE = 209.77, p < .0001. None of the interactions was significant (using a p < .05 criterion), consistent with the appearance in Figure 1 (left panel) that neither initial starting rate nor passage predictability affected either age group differentially. Paired comparison tests showed that regardless of age, passage predictability, or initial starting direction, selected speech rates were always significantly faster than the original recorded rate (p < .05 or better in all cases).

The right panel shows the mean percentage of propositions (“idea units”) correctly recalled for the passages heard at 140, 260, and 330 wpm, with lines of best fit plotted through these three points for each condition. Recall scores for passages presented at the participants' own selected rates taken from the first session are plotted in the shaded area. (A vertical arrow is plotted on the abscissa to indicate the relative position of the original 174 wpm recording rate.)

As can be seen in the right panel, recall was higher for high-predictability passages than for low-predictability passages, F(1,30) = 144.15, MSE = 126.14, p < .0001, the younger adults recalled more than the older adults, F(1,30) = 27.85, MSE = 629.38, p < .0001, and there was also a significant main effect of speech rate, F(2,60) = 55.06, MSE = 108.57, p < .0001. None of the interactions was significant, although a potential age by speech rate interaction may have been constrained by a floor effect for the older participants for the low-predictability passages at the fastest speech rate (330 wpm). Although participants' selected rates were faster than the original recording rate of 174 wpm, we can see that recall performance was close to extrapolated levels for this rate (indicated by the short vertical lines plotted on the speech rate functions). Paired comparison tests showed that participants' recall performance at their selected rates and at the slowed (140 wpm) rate did not differ significantly (using a p < .05 criterion).

It is possible that factors such as hearing acuity and span scores might affect rate preferences and/or recall performance in a broad population of young and older adults. In the present case, however, care was taken to select young and older adults with good digit spans and levels of hearing acuity, such that an analysis of our data failed to show an effect of either of these factors on speech rate preferences or recall performance.

Discussion

The present results demonstrate first, that older adults select slower speech rates than young adults when listening to connected speech, and second, that both age groups reliably mod-
LISTENING-RATE PREFERENCES

Figure 1. Left panel shows selected speech rates in words per minute (wpm) for younger and older adults listening to high and low predictability speech passages. Bar heights show selected rates when the speech was initially presented faster (F) or slower (S) than the original recording rate. Right panel shows recall performance for the younger and older participants when high and low predictability passages were presented at 140, 260, and 330 wpm and at participants' own selected rates (plotted in the shaded area). Error bars represent one standard error.

The latter finding shows that the older adults' speech rate selections were responsive to the listening task, rather than, for example, reflecting an immutable cognitive style. It is interesting to note that both participant groups tended to select speech rates that were significantly faster than the speech rate of the original recording. In view of the fact that the older adults showed poorer recall than the young, however, it could be argued that the older adults' preferred rates were inappropriately farther from their optimal rates than the rates selected by the young. Older adults are generally known to be at a disadvantage when information loads are high, whether the information load is manipulated by increasing speech rates or by varying the informational complexity of the speech (Riggs, Wingfield, & Tun, 1993). That is, at any given rate the information load will be proportionally higher for the older adults with consequent effects on comprehension and memory.

Examination of both the young and older participants’ recall performance relative to a slower rate (140 wpm) than their average selections, however, suggests that their rate selections were reasonable ones for the task. This latter finding, however, should be interpreted with caution. Unlike slowing speech using uniform time-expansion, when speakers are asked to speak slower they do so more by adding pauses at strategic syntactic points (e.g., clause and sentence boundaries), than by lengthening individual speech sounds (Lane & Grosjean, 1973; Speer, Wayland, Kjelgaard, & Wingfield, 1994). Indeed, it can be reliably shown that slowing speech by lengthening pauses at syntactic boundaries is more beneficial to listener comprehension and recall than expanding all parts of the speech signal (cf., Blumstein et al., 1985; Schnitt, 1983). Studies with older adults confirm that adding processing time at salient linguistic boundaries is effective to a degree that adding time elsewhere in the speech signal is not (Wingfield, Tun, Koh, & Rosen, 1999). The present results should thus not imply that listeners could not have shown better recall had slowing been accomplished by mirroring the way speech is slowed in natural discourse.

Our results do show, however, that older adults can be as adept as the young in their ability to monitor passage difficulty as it is being heard, and to moderate their selection of preferred listening rates accordingly.

ACKNOWLEDGMENTS

This work was supported by NIH grant R37 AG04517 from the National Institute on Aging. We also gratefully acknowledge support from the W.M. Keck Foundation.

Address correspondence to Dr. Arthur Wingfield, Volen National Center for Complex Systems (MS 013), Brandeis University, Waltham, MA 02454-9110. E-mail: wingfield@volen.brandeis.edu

REFERENCES


Received July 20, 1998
Accepted December 18, 1998