

# Perception of alternated speech operates similarly in young and older adults with age-normal hearing

RAJ STEWART, ETHAN YETTON, AND ARTHUR WINGFIELD  
*Brandeis University, Waltham, Massachusetts*

When speech is rapidly alternated between the two ears, intelligibility declines as the rate of alternation approaches 3 to 5 switching cycles per second, and then, paradoxically, returns to a good level beyond that point. We tested intelligibility when shadowing was used as a response measure (Experiment 1), when recall was used as a response measure (Experiment 2), and when time-compression was used to vary the speech rate of the presented materials (Experiment 3). In spite of claims that older adults are generally slower in switching attention, younger and older adults did not differ in the critical alternation rates producing minimal intelligibility. We suggest that the point of minimal intelligibility in alternated speech reflects an interaction between (1) the rate of disruption induced by breaking the speech stream between two sound sources, (2) the amount of contextual information per ear, and (3) the size of the silent gaps separating the speech elements that must be perceptually bridged.

The human auditory system is adept at handling concurrent and overlapping sound sources—such as multiple speakers or speech heard in background clutter—using their spatial, temporal, and spectral features to segregate them into separate auditory streams for further analysis (Bregman, 1990; Sussman, Čeponienė, Shestakova, Näätänen, & Winkler, 2001). The subject of our present inquiry is what happens when the human auditory system is confronted with a single message that is alternated rapidly between two spatial sources.

The first demonstration of the consequences of this effect was by Cherry and Taylor (1954), who used an electronic switch to rapidly alternate speech back and forth between the two earpieces of dichotic earphones. They reported that intelligibility progressively declined with increasing rates of alternation, becoming poorest at approximately 3 to 5 complete switching cycles per second (cps), a rate corresponding to about 167 to 100 msec per ear. As the alternation rate was increased beyond this point, intelligibility paradoxically improved, thus yielding a V-shaped function of intelligibility decline and intelligibility recovery over successively increasing alternation rates. The size of the decrement at this critical rate is smaller when the speech materials are predictable enough to allow missed elements to be inferred from the linguistic context (cf. Speaks & Troien, 1974; Wingfield & Wheale, 1975a) and larger when perceptual processing is made more difficult by mixing masking noise with the speech signal (Rupf, Hughes, & House, 1971). The basic V-shaped intelligibility function remains, however, as interaural alternation rates are progressively increased.

The intelligibility loss at the critical rate of 3 to 5 alternating cps is intriguing because the speech is always present, either in one ear or the other. Cherry and Taylor (1954) attributed this loss to the time it takes a listener to switch attention from one ear to the other in order to follow the shifting signal. If each shift of attention from one ear to the other is accompanied by a perceptual “dead time,” during which no usable information is available from either ear, as alternation rates increase, so would the cumulative loss of information.

Cherry and Taylor (1954) argued that the point of minimal intelligibility in the V-shaped intelligibility function marks the point at which the listener’s switching attention is completely out of phase with the shifting signal: By the time the listener’s attention has shifted from one ear to the other to follow the signal, the speech would already have returned to the original ear. They accounted for the improved intelligibility when alternation rates are increased beyond this point by suggesting that, at these very rapid rates, listeners begin to concentrate on the input from only one ear. The result would be an interrupted signal to the attended ear in which the frequent but brief interruptions would result in a perceptual “picket fence” effect, with the broad speech envelope still recognizable. The faster the rate of alternation, the smaller the silent gaps to be perceptually bridged (e.g., Miller & Licklider, 1950).

An alternative account for the V-shaped function was offered by Huggins (1964), who noted that the 3- to 5-cps region of minimal intelligibility represents an ear dwell-time that approximates half the duration of most syllables. He suggested that, if each ear (i.e., contralateral hemi-

sphere) processed its own information, the critical alternation rate would, on average, represent maximal disruption of syllabic processing units. It might thus be that the point of minimal intelligibility is a function not of the amount of time per ear, as suggested by Cherry and Taylor (1954), but of the amount of speech content per ear. Following this reasoning, Huggins predicted that a shift in the critical alternation rate producing minimal intelligibility would occur if the speech rate were accelerated, as this would reduce the duration of the spoken syllables and thus increase the amount of speech content to a single ear at any given alternation rate. This prediction was confirmed (Huggins, 1964; Wingfield & Wheale, 1975b).

Although initially gaining wide acceptance (e.g., Neisser, 1966, pp. 184–187), attempts to confirm Huggins's hypothesis by specifically interrupting speech either between syllables or within syllables failed to support his syllabic interruption view (Samuel, 1991). Although there has yet to be a totally satisfactory account of the V-shaped function of intelligibility on interaural alternation rate, the phenomenon itself is a robust one (Cherry & Taylor, 1954; Huggins, 1964; Samuel, 1991; Schubert & Parker, 1955; Wingfield, 1977; Wingfield & Wheale, 1975a, 1975b).

The phenomenon of alternating speech is particularly interesting in the context of adult aging. In addition to difficulties with binaural fusion (i.e., integrating sounds arriving separately to the two ears; Pichora-Fuller & Souza, 2003), older adults are known to show a general slowing in a variety of perceptual and cognitive operations (Salt-house, 1996). So, for example, if each switch of attention is assumed to take a certain amount of time, during which there is a perceptual "dead time" resulting in information loss, then a slowed perceptual system would not only produce a generally poorer level of intelligibility at all alternation rates, but, more important, one would see the point of minimal intelligibility occur at a slower rate of alternation for older than for younger adults. That is, because of arguments that older adults find it more difficult, and take longer, to switch attention than do younger adults (Panek & Rush, 1981; Verhaeghen & Basak, 2005; Waugh & Barr, 1980), contrasting younger with older adults' performance would be an interesting test of the alternating attention hypothesis as an account for the effects on intelligibility of speech alternating rapidly between the two ears.

There are a number of reasons to expect that older adults would show overall lower intelligibility performance across all alternation rates. Even when an older adult's peripheral hearing acuity remains good in terms of his or her audiometric profile, aging can also be accompanied by central auditory processing deficits that result in a decline in efficiency of temporal and/or frequency resolution that would affect overall speech intelligibility (see B. A. Schneider & Pichora-Fuller, 2000, for a review). Of interest, however, is not the heights of the intelligibility curves in terms of  $y$ -intercepts. Rather, our interest is in whether the point of minimal intelligibility shows a shift of position on the  $x$ -axis of an intelligibility–alternation rate function for older adults in relation to younger adults.

## EXPERIMENT 1

### Method

**Participants.** The older adults were 32 community-dwelling volunteers, 10 men and 22 women, who ranged in age from 66 to 82 years ( $M = 73.3$  years). The group had a mean of 15.9 years of formal education ( $SD = 2.3$ ), and a mean Shipley vocabulary score (Zachary, 1986) of 15.6 ( $SD = 2.4$ ). The younger participants were 32 undergraduate and graduate students, 12 men and 20 women, with ages ranging from 18 to 25 years ( $M = 20.3$  years). The group had a mean of 14.1 years of formal education at time of testing ( $SD = 2.1$ ) and a mean Shipley vocabulary score of 14.4 ( $SD = 2.0$ ). Both groups were thus well-educated and had good vocabularies, with the older group having an average of 1.8 more years of formal education than the younger adults [ $t(62) = 1.69$ , n.s.], and, as is common for older adults (Verhaeghen, 2003), a somewhat higher vocabulary score than the younger adults [ $t(62) = 2.26$ ,  $p < .05$ ]. All participants were native speakers of American English, with no known history of stroke, Parkinson's disease, or other neuropathology that might compromise cognitive function.

Although the older adults in this study had good hearing for their ages (Morrell, Gordon-Salant, Pearson, Brant, & Fozard, 1996), the older group had a higher pure tone average (PTA; taken as the mean thresholds for tones at 500, 1000, and 2000 Hz) across both ears than did the younger adults (older  $M = 15.5$  dB HL,  $SD = 6.7$ ; younger  $M = 6.8$  dB HL,  $SD = 4.3$ ) [ $t(62) = 6.77$ ,  $p < .001$ ]. Although there was a significant difference in PTA for the two age groups, no individual participant in either age group exceeded the cutoff of 25-dB HL that is typically considered to be clinically normal for speech (Hall & Mueller, 1997). Both groups had approximately symmetrical hearing across both ears, with no participant showing an asymmetry greater than 5 dB. Speech reception thresholds (SRTs; the lowest decibel level at which prerecorded, two-syllable words can be correctly identified 50% of the time) were also measured. The older participants also had a higher average SRT (15.9 dB HL,  $SD = 7.7$ ) than did the younger adults (5.6 dB HL,  $SD = 4.8$ ) [ $t(62) = 6.73$ ,  $p < .001$ ], but all participants' SRTs were within the 25-dB upper limit considered to be clinically normal for speech (Hall & Mueller, 1997).

**Stimulus materials.** The stimuli consisted of sixteen 125-word English passages representing a variety of general interest topics and a range of common syntactic forms. Proper names and difficult technical terms were removed. All passages were spoken by a female speaker of American English at an average speech rate of 183 words per minute (wpm).

**Procedure.** Participants were tested individually in a sound-isolated testing room with the speech stimuli presented using E-A-RTONE 3A (E-A-R Auditory Systems, Acaro Company, Indianapolis, IN) insert earphones. The stimuli were prepared, amplitude calibrated, and stored as 44.1-kHz WAV format computer sound files routed for presentation through a Grason-Stadler GSI 61 audiometer (Grason-Stadler, Madison, WI). Prior to the experiment, the intensity of the stimuli was adjusted individually for each ear for each participant, with all participants hearing the stimuli at 40 dB above their individual SRT for that ear.

Each participant heard two passages under each of eight conditions: passages presented in a normal, nonalternating manner binaurally to both ears, or alternated at 1, 2, 3, 4, 5, 8, or 16 alternating cps, corresponding to 500, 250, 167, 125, 100, 63, or 31 msec per ear. The particular passages heard in each alternation condition were counterbalanced across participants so that, by the end of the experiment, each passage had been heard an equal number of times at each alternation rate or in the baseline, nonalternating condition. The order of presentation of alternation conditions was also varied between participants.

Participants were told they would be hearing speech passages presented either simultaneously to both ears or alternated between the two ears. In either case, participants were instructed to listen carefully and "shadow" the speech, saying aloud what was being heard on a word-by-word basis as it was being heard. This followed Cherry

and Taylor's (1954) original procedure, designed to minimize effects of memory on report accuracy.

Participants were told that, if they should miss an occasional word while shadowing, they were not to stop, but to pick up where they could and continue shadowing. Participants were told that their responses were being audio recorded for transcription and accuracy scoring. The main experiment was preceded by a brief practice session with four passages to familiarize the participants with the shadowing task and the sound of alternated speech. None of these passages was used in the main experiment.

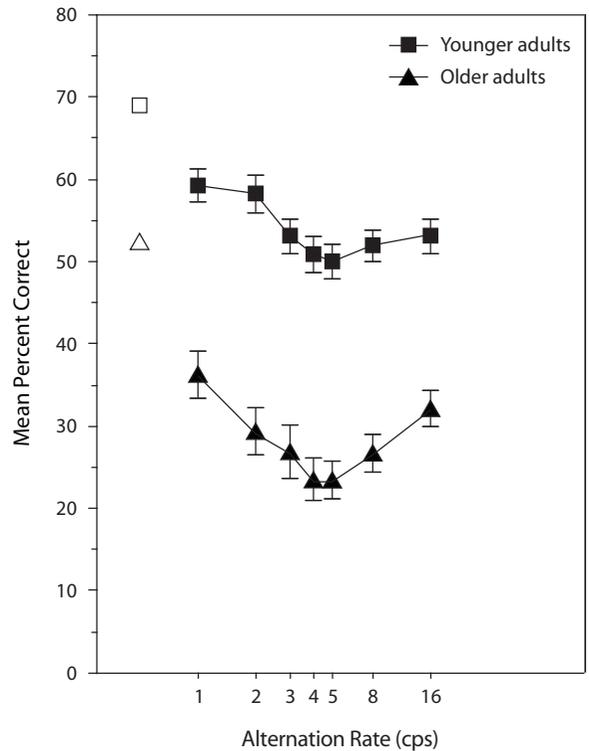
## Results

Intelligibility was taken as the number of words correctly reported from the middle 100 words of each passage, with the first 15 and last 10 words of the passage not scored, in order to eliminate end effects. Credit was given for each word reported correctly, regardless of order. No partial credit was given for incomplete words. Figure 1 shows the mean intelligibility scores for uninterrupted speech (open symbols on the left side) and for each of the seven alternation rates (closed symbols). Data are shown separately for the younger and the older participants. The abscissa is shown on a logarithmic scale.

As might be expected, the older adults showed an overall lower accuracy level than did the younger adults, to include the uninterrupted speech condition [ $t(62) = 4.84$ ,  $p < .001$ ]. Of special interest, however, is that both the younger and older adult participants showed a pattern of decline in accuracy scores as the rate of interaural alternation was progressively increased from 1 alternating cps to a minimal point falling within Cherry and Taylor's (1954) 3 to 5 alternating cps, followed by a recovery of intelligibility when alternation rates were further increased. Both the younger and older adults' intelligibility curves yielded a significant quadratic trend component (Keppel, 1991), confirming a canonical V-shaped intelligibility function for both age groups [younger adults,  $t(31) = 4.11$ ,  $p < .001$ ; older adults,  $t(31) = 7.97$ ,  $p < .001$ ]. A 7 (alternation rate: 1, 2, 3, 4, 5, 8, 16 cps)  $\times$  2 (age: younger, older) mixed design ANOVA was conducted on these data, with alternation rate as a within-participants variable and age as a between-participants variable. The ANOVA confirmed a significant main effect of alternation rate on report accuracy [ $F(6,434) = 5.33$ ,  $p < .001$ ] and a significant main effect of age [ $F(1,62) = 390.67$ ,  $p < .001$ ]. Consistent with the appearance of similar points of minimal intelligibility for both age groups, the ANOVA failed to show a significant alternation rate  $\times$  age interaction [ $F(6,434) < 1$ ].

## Discussion

The results of Experiment 1 reproduce the previously described V-shaped intelligibility function for alternated speech and show a point of minimal intelligibility to occur in the 3 to 5 alternating cps region, as observed in previous reports (Cherry & Taylor, 1954; Huggins, 1964; Samuel, 1991; Wingfield & Wheale, 1975a, 1975b). The overall lower intelligibility levels observed for the older adults is not surprising for the several reasons alluded to previously. For example, although the older adults were tested with speech at the same intensities in relation to their individual



**Figure 1.** Percentage of words correctly repeated for speech alternated between the two ears by younger and older adults as a function of the rate of alternation. The abscissa is shown on a logarithmic scale. Open symbols on the left side of the figure show younger and older adults' mean accuracy levels for binaurally presented, nonalternated speech. Error bars represent  $\pm 1$  SE.

SRTs as were the younger adults, amplification alone may not fully compensate for age-related, higher level auditory processing deficits, such as the lowered efficiency with temporal and spectral resolution that might be expected to reduce the efficiency of speech perception (Pichora-Fuller & Souza, 2003; B. A. Schneider & Pichora-Fuller, 2000). It is thus especially notable that the shapes of the intelligibility functions are similar for both age groups.

Cherry and Taylor's (1954) switch-time theory would lead one to expect older adults (with their generally slowed systems) to show a steeper intelligibility decline with increasing alternation rates, reflecting a greater cumulative effect of information loss at each point of attention switching as the speech ceases in one ear and begins in the other. Also, to the extent that the point of minimal intelligibility reflects an attentional system out of phase with the alternating input, a slower switching rate would predict a point of minimal intelligibility at a different (slower) position on the abscissa for the older than for the younger adults. As can be seen, such a shift in the point of minimal intelligibility was not evident in these data.

In Experiment 1, we used the traditional shadowing method to estimate the intelligibility of the alternated speech (e.g., Cherry & Taylor, 1954; Huggins, 1964). The advantage of shadowing is that the report of the speech is immediate and does not rely heavily on participants'

memory. It might be argued, however, that the shadowing requirement of speaking and listening at the same time puts older adults at a disadvantage. First, shadowing requires the participant to engage in the two simultaneous activities of speaking and listening, which might differentially challenge older adults' susceptibility to dual-task interference (Anderson, Craik, & Naveh-Benjamin, 1998; Tun, Wingfield, Stine, & Meccas, 1992). Second, at the perceptual level, older adults might be more susceptible to the acoustic masking of the target speech by their own voice (Bergman, 1980). Finally, the requirement in shadowing that the speakers' responses keep up with the input might put older adults at a disadvantage because of generally slower articulation rates (Benjamin, 1997; Smith, Wasowicz, & Preston, 1987). Any or all of these factors could account for the generally lower performance level of the older adults at all alternation rates.

In Experiment 2, a different group of younger and older adults similar in characteristics to those in Experiment 1 heard the same passages, but this time the passages were halted at linguistically salient intervals to allow the participants time to recall the speech segment they had just heard. This procedure obviated the need for the participants to talk and listen simultaneously. Although requiring recall of what had been heard in each segment could put the older adults at a memory disadvantage, it relieved them of having to provide verbal output while simultaneously receiving input and of masking of the speech input by the sound of their own voices.

To the extent that the shadowing requirement created differentially greater difficulty for the older participants, one would expect overall levels of performance to be higher, with a reduction in the magnitude of age difference at all alternation rates in comparison with that observed in Experiment 1. Of primary interest, however, is to insure that the older adults' minimal intelligibility region, which matched that of the younger adults' in Experiment 1, would not have been different had shadowing not been required.

## EXPERIMENT 2

### Method

**Participants.** The older adults were 16 community-dwelling volunteers who had not taken part in Experiment 1—6 men and 10 women, ranging in age from 66 to 80 years ( $M = 74.3$  years). The group had a mean of 16.8 years of formal education ( $SD = 2.7$ ) and a mean Shipley vocabulary score of 15.2 ( $SD = 1.4$ ). The younger participants were 16 undergraduate and graduate students who had not taken part in Experiment 1—5 men and 11 women, with ages ranging from 19 to 23 years ( $M = 20.9$  years). The group had a mean of 13.5 years of formal education at time of testing ( $SD = 1.9$ ) and a mean Shipley vocabulary score of 14.3 ( $SD = 1.9$ ). Both groups were thus well educated and had good vocabularies, with the older group having an average of 3.3 more years of formal education than the younger adults [ $t(30) = 4.00, p < .001$ ] and a somewhat higher—but not significantly so—vocabulary score than the younger group [ $t(30) = 1.53, n.s.$ ]. All participants were native speakers of American English, with no known history of stroke, Parkinson's disease, or other neuropathology that might compromise cognitive function.

The older adults in this study had good hearing for their ages (Morrell et al., 1996). As in Experiment 1, the older group had a

higher PTA (500, 1000, and 2000 Hz) across both ears than did the younger adults (older  $M = 16.2$  dB HL,  $SD = 8.7$ ; younger  $M = 7.1$  dB HL,  $SD = 4.9$ ) [ $t(30) = 3.65, p < .001$ ]. Both groups had approximately symmetrical hearing across both ears, with no individual participant showing an asymmetry greater than 5 dB. SRTs were higher for the older adults ( $M = 16.8$  dB HL,  $SD = 7.4$ ) than for the younger adults ( $M = 5.7$  dB HL,  $SD = 4.2$ ) [ $t(30) = 5.21, p < .001$ ], although all participants were within the range considered to be clinically normal for speech (Hall & Mueller, 1997).

**Stimuli and Procedure.** The stimuli consisted of the same 16 paragraph-length passages as used in Experiment 1, presented without alternation and at the same 7 alternation rates (1, 2, 3, 4, 5, 8, and 16 alternating cps). As before, the speech was presented over insert earphones at 40 dB above each individual's SRT for each ear.

To obviate the need for speaking and listening simultaneously, while not requiring participants to attempt to recall an entire 125-word passage, the speech passages were interrupted after clause and sentence boundaries in the passages, resulting in an average segment size of 8.5 words. When the speech stopped, the participant was to recall the words heard in that segment. Participants were told that if they remembered a word that they had failed to give in their recall of a prior segment, they would be given credit for that word if they reported it along with the segment currently being recalled. Fifteen seconds was allowed for recall, at which point the next segment of the passage (a full sentence or clause) was presented for recall. In this way, the participant was allowed to recall the passage on a segment-by-segment basis. Accuracy scoring was conducted following the same procedures as were used in Experiment 1.

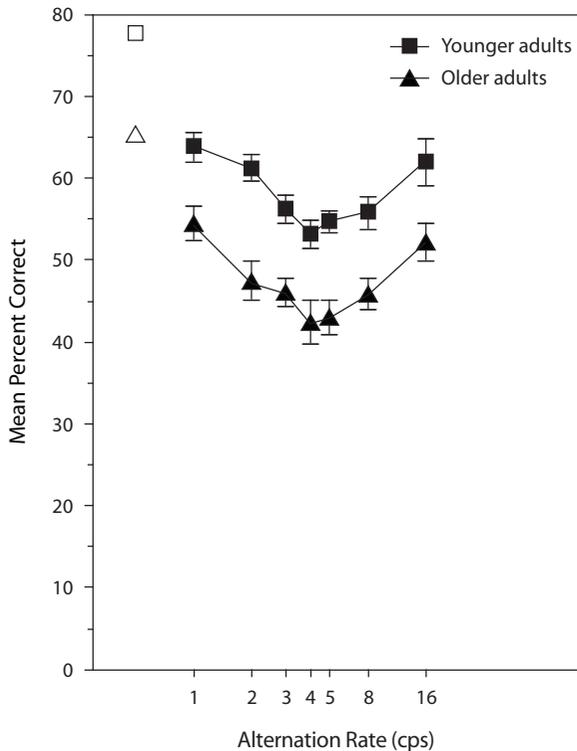
## Results and Discussion

The open symbols on the left side of Figure 2 show the mean percentage of words correctly recalled by the younger and older participants when speech was presented continuously to both ears. The age difference is diminished but is still significant [ $t(30) = 9.40, p < .001$ ]. Although both the younger and older adults show higher recall scores when the shadowing requirement was removed, this procedure gave a differential advantage to the older adults, reducing the overall size of the age difference. One sees in Figure 2, however, that both age groups continue to show a V-shaped intelligibility function (confirmed by a quadratic contrast), with the region of minimal intelligibility remaining the same for both the younger and older adults. A 7 (alternation rate)  $\times$  2 (age) ANOVA showed significant main effects of alternation rate [ $F(6,222) = 60.44, p < .001$ ] and of age [ $F(1,30) = 19.93, p < .001$ ], but no alternation rate  $\times$  age interaction [ $F(6,222) < 1$ ]. Experiment 2 also shows a more clearly defined point of minimal intelligibility at 4 alternating cps, or 125 msec per ear—a point at the midrange between Cherry and Taylor's (1954) 3- to 5-cps estimated region of reduced intelligibility.

We thus conclude that the similarity in the region of minimal intelligibility for younger and older adults as observed in Experiment 1 was not simply confined to the use of shadowing as a response measure.

## EXPERIMENT 3

We previously cited the early argument that the critical feature defining the V-shaped intelligibility function over alternation rates is not due to the amount of time per ear, reflective of a minimal perceptual switching time with attendant information loss, but to the amount of speech



**Figure 2.** Percentage of words correctly reported by younger and older adults for speech alternated between the two ears when speech was interrupted periodically to allow participants time to recall what had been heard. The abscissa is shown on a logarithmic scale. Open symbols on the left side of the figure show mean accuracy levels for younger and older adults for binaurally presented, nonalternated speech. Error bars represent  $\pm 1$  SE.

content per ear (Huggins, 1964). The two features can be independently varied by increasing the speech rate of the stimuli so that a greater amount of uninterrupted speech will be presented to each ear at any given alternation rate. Although Huggins's specific syllable-interruption account has not found clear support (cf. Huggins, 1967; Rupf et al., 1971; Samuel, 1991), it has been shown that the point of minimal intelligibility can be shifted to a faster alternation rate when the speech rate (i.e., speech content per ear) is increased (Huggins, 1964; Wingfield & Wheale, 1975b).

The similarity in results for younger and older adults in Experiments 1 and 2 in terms of the shapes of the functions and points of minimal intelligibility suggests the possibility of a similar response to an increase in speech rate for older and younger adults. It is well known that older adults are differentially affected by time compression of speech in comparison with younger adults (Gordon-Salant & Fitzgibbons, 1993; Wingfield, Tun, Koh, & Rosen, 1999). As such, one would expect to see the overall height of older listeners' intelligibility curves to be especially affected by rapid speech. However, if the mechanisms of response to alternating speech are affected by the amount of speech content per ear, and if these mechanisms are similar for younger and older adults, one would expect

both age groups to show an equivalent shift in the point of minimal intelligibility to a faster alternation rate (i.e., shorter dwell time per ear) when speech rate is increased.

## Method

**Participants.** The older adults were 32 community-dwelling volunteers who had not taken part in either of the previous experiments—14 men and 18 women, ranging in age from 67 to 85 years ( $M = 74.5$  years). The group had a mean of 16.3 years of formal education ( $SD = 2.8$ ) and a mean Shipley vocabulary score of 15.9 ( $SD = 2.6$ ). The younger participants were 32 undergraduate and graduate students who had not taken part in either of the previous experiments—13 men and 19 women, with ages ranging from 18 to 22 years ( $M = 19.0$  years). The group had a mean of 13.0 years of formal education at time of testing ( $SD = 1.3$ ) and a mean Shipley vocabulary score of 14.8 ( $SD = 1.7$ ). Both groups were thus well-educated and had good vocabularies, with the older group having an average of 3.3 more years of formal education than the younger adults [ $t(62) = 6.05, p < .001$ ] and, again, a somewhat higher vocabulary score than the young adults [ $t(62) = 2.00, p < .05$ ]. All participants were native speakers of American English, with no known history of stroke, Parkinson's disease, or other neuropathology that might compromise cognitive function.

The older adults in this study had good hearing for their ages (Morrell et al., 1996). As in Experiments 1 and 2, the older group had a higher PTA (500, 1000, and 2000 Hz) across both ears than did the younger adults (older  $M = 20.9$  dB HL,  $SD = 2.2$ ; younger  $M = 7.5$  dB HL,  $SD = 6.5$ ) [ $t(62) = 11.05, p < .001$ ]. No individual participant in either age group, however, exceeded a cutoff of 25 dB, a level considered to be clinically normal for speech (Hall & Mueller, 1997). Both groups had approximately symmetrical hearing across both ears, with no individual participant showing an asymmetry greater than 5 dB. The older participants also had a higher average SRT (17.1 dB HL,  $SD = 6.7$ ) than did the younger adults (5.6 dB HL,  $SD = 4.9$ ) [ $t(62) = 7.84, p < .001$ ], but all participants were within limits considered to be clinically normal for speech (Hall & Mueller, 1997).

**Stimuli and Procedure.** The stimuli consisted of the same 16 paragraph-length passages of English prose as used in Experiment 1. The original passages, which had been spoken at 183 wpm, were time-compressed using the pitch-synchronous overlap (PSOLA) sampling method (SoundEdit 16; Macromedia Inc., San Francisco, CA) to 85% and 70% of their original speaking durations, corresponding to speaking rates of 215 and 261 wpm, respectively. The compression procedure deleted small (20 msec) segments at regular intervals from both speech and silent periods so as to maintain the relative temporal patterning of the speech, as well as maintenance of the original pitch contour.

Each participant heard all 16 passages, eight compressed to 85% of the original duration (215 wpm) and eight passages compressed to 70% of original duration (261 wpm). The particular passages heard at each of the two speech rates were counterbalanced across participants so that, by the end of the experiment, each passage had been heard an equal number of times at each speech rate and in each alternation condition. Stimulus presentations were blocked by presentation rate, with half of the participants receiving the passages compressed to 85% of the original duration first and the other half receiving the passages compressed to 70% of original duration first. In all other respects, the procedures were the same as those used in Experiment 1. Shadowing was again used as the response measure. The main experiment was preceded by a practice session with six passages to familiarize participants with the shadowing task and the sound of time-compressed and alternated speech.

## Results

The left panel of Figure 3 shows the intelligibility functions (percentage of words correct) for the younger and older adults when the speech was time compressed to 85% of the

original speaking duration. As can be seen from the open symbols on the left side of the panel, the younger adults showed superior performance in relation to the older adults for binaurally presented, nonalternating speech [ $t(62) = 3.49, p < .001$ ]. Consistent with the results of Experiments 1 and 2, a 7 (alternation rate)  $\times$  2 (age) ANOVA confirmed a significant main effect of alternation rate on intelligibility [ $F(6,434) = 6.36, p < .001$ ] and showed a main effect of age [ $F(1,62) = 149.38, p < .001$ ]. Although age had the effect of reducing the overall level of performance at all alternation rates, the point of minimal intelligibility appeared at the same point for both the younger and older adults, centered on 4 alternating cps, corresponding to 125 msec per ear. Consistent with this similarity, the ANOVA failed to reveal a significant alternation rate  $\times$  age interaction [ $F(6,434) < 1$ ].

The right panel shows the same data, but for speech that had been time compressed to 70% of the original speaking time. The point of minimal intelligibility is aligned at the same position on the abscissa for both the younger and older adults. In this case, however, this point is shifted to a faster alternation rate (5 cps) than these participants showed for speech that was compressed to 85% of the original duration ( $p < .05$  for younger and older adults alike).

As in the previous cases, there was a significant age difference for the binaurally presented, nonalternating speech [ $t(62) = 4.31, p < .001$ ]; an ANOVA conducted on

the alternated speech data showed significant main effects of alternation rate [ $F(6,434) = 8.23, p < .001$ ] and age [ $F(1,62) = 94.03, p < .001$ ]. Figure 3, however, shows a difference in the recovery patterns for the younger and older adults when alternation rates were increased beyond the critical point of minimal intelligibility at 5 alternating cps. In the case of the younger adults, intelligibility continued to rise when the alternation rates were increased beyond this point. The older adults' performance, however, failed to improve progressively when the alternation rates were increased beyond 5 alternating cps.

This difference in recovery pattern was not revealed by a significant alternation rate  $\times$  age interaction [ $F(6,434) = 1.26, n.s.$ ], but this absence of an interaction effect may have been heavily influenced by the parallel declines in intelligibility for both the younger and older adults over the first four alternation rates. To test this possibility, we performed an ANOVA on only the final three alternation rates (5, 8, and 16 alternating cps). This analysis showed main effects of both alternation rate [ $F(2,186) = 7.85, p < .001$ ] and age [ $F(1,186) = 59.37, p < .001$ ], but a significant alternation rate  $\times$  age interaction [ $F(2,186) = 3.10, p < .05$ ] as well.

**Discussion**

The use of time-compressed speech allowed us to vary independently the duration of the speech segment per ear

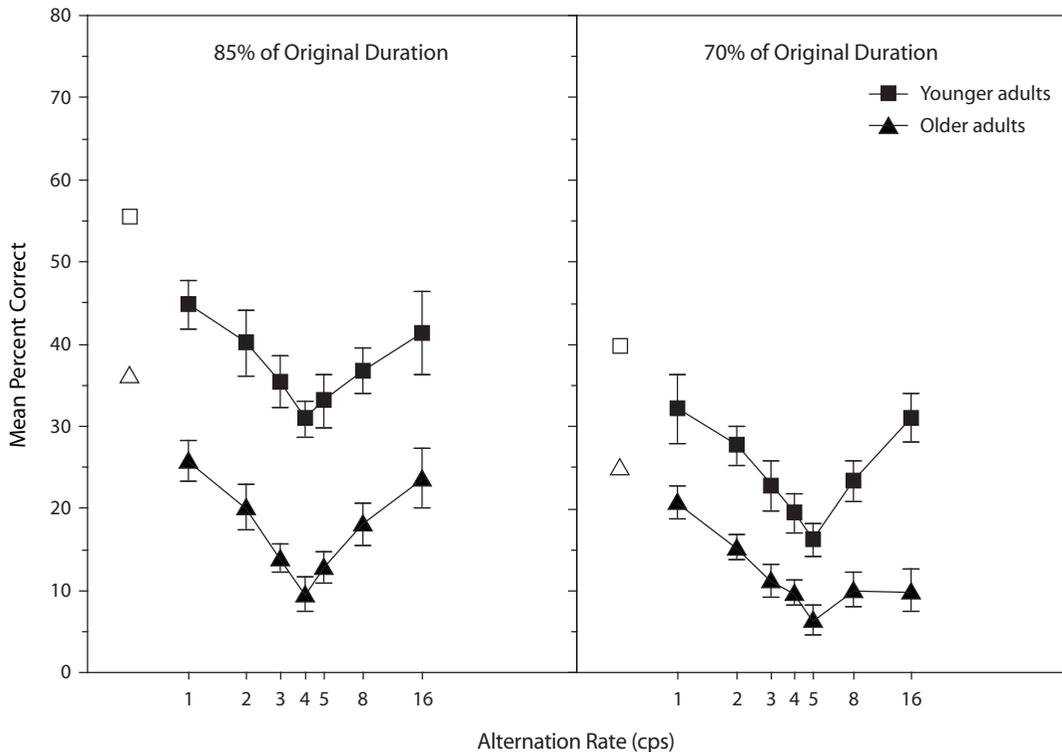


Figure 3. Percentage of words correctly repeated by younger and older adults for speech alternated between the two ears as a function of the rate of alternation when speech was time compressed to 85% of its original duration (left panel) and 70% of its original duration (right panel). The abscissa is shown on a logarithmic scale. Open symbols on the left side of each panel show mean accuracy for younger and older adults for binaurally presented, nonalternating speech heard at each speech rate. Error bars represent  $\pm 1 SE$ .

versus the amount of speech content in each segment in each ear. Following Huggins's (1964) argument, if the critical parameter determining the point of minimal intelligibility is absolute time per ear, the points of minimal intelligibility for both speech rates should be aligned at the same point on the abscissa. In contrast, if the amount of speech contained in each ear segment is the critical parameter, the point of minimal intelligibility should shift to the right on the abscissa when speech rate is increased. This would result from an increase in the amount of speech content available in each ear's dwell time, a finding reported by Huggins (1964) and Wingfield and Wheale (1975b) for younger adults. This predicted shift appeared in the present experiment, with a comparable shift in the point of minimal intelligibility for both younger and older listeners.

It is interesting to note that, when speech was time compressed to 70% of its original duration (261 wpm), the older adults did not show the same progressive recovery at rates faster than the point of minimal intelligibility, as seen in Experiments 1 and 2 with noncompressed speech or in this experiment when speech was compressed to 85% of its original duration (215 wpm). A possible account for this difference might be that, when interaural alternation rates become very rapid, listeners begin to give an attentional bias to a single ear, basing identification primarily or solely on the interrupted "picket fence" signal to that ear (cf. Cherry & Taylor, 1954; Miller & Licklider, 1950; Wingfield, 1977). It might thus be that the older adults found the combined perceptual challenge of this reduced information, exacerbated by the high degree of time compression, a challenge too difficult to attempt. Such effects have previously been found for older adults under difficult listening conditions (Bergman, 1980, p. 77).

Should this have been the case, it would suggest that this difficulty threshold had not yet been reached by the older adults for speech compressed to 85% of its original duration (i.e., a 15% increase in speech rate). Consistent with this argument is the observation that this lesser degree of compression did not affect the older adults' intelligibility recovery, nor was this smaller degree of compression sufficient to shift the point of minimal intelligibility relative to noncompressed speech as observed in Experiment 2.

## GENERAL DISCUSSION

The results of all three experiments reproduce the V-shaped intelligibility function for alternated speech first described by Cherry and Taylor (1954). It has been argued that the mechanisms that produce perceptual separation of interleaved acoustic signals, such as that which occurs with rapid tone sequences that alternate in frequency, are determined by bottom-up sensory processes independent of any need for attentional resources (Sussman et al., 2001; but see also Macken, Tremblay, Houghton, Nicholls, & Jones, 2003). In contrast to stream segregation, successful perception of alternated speech implies a fusion of the speech alternating between the two input sources. The present results suggest that this fundamental process is unaffected by adult aging, in spite of age-related

factors such as decline in auditory acuity (Morrell et al., 1996) and reduced speed and efficiency in rapid attention switching (e.g., Verhaeghen & Basak, 2005). The apparent immunity of this process to developmental change is characteristic of attention-free automatic operations (W. Schneider, Dumais, & Shiffrin, 1984).

Cherry and Taylor (1954) offered a two-process account for the V-shaped intelligibility function: The initial progressive decrement in intelligibility as alternation rates are increased results from information loss due to a lag in shifting attention, with the point of minimal intelligibility occurring when the switching signal and the time taken to switch attention combine to put the listener's attention continually out of phase with the shifting signal. The improvement in intelligibility beyond this point results from the perceptual system's focusing primarily on the interrupted signal from just one ear. As the alternation rate becomes very rapid, the input to either ear would be speech interrupted by progressively shorter silent gaps, progressively easier to perceptually bridge.

In the five decades since Cherry and Taylor's (1954) discovery, a number of experiments have replicated their basic finding of reduced intelligibility when speech is alternated between the two ears at 3–5 alternating cps. These subsequent studies have suggested, however, that Cherry and Taylor's initial account was at best incomplete. Most notable was Huggins's (1964) demonstration that the point of minimal intelligibility is affected by speech rate. This would not be expected were the point of minimal intelligibility a consequence of the absolute time per ear of the shifting signal independent of speech content in that ear. Huggins's initial argument that minimal intelligibility results specifically from the disruption of syllabic units was, however, not sustained in subsequent studies that interrupted speech between and at various points within syllables (cf. Huggins, 1967; Rupf et al., 1971). A further challenge to Huggins's syllabic disruption hypothesis came with Samuel's (1991) demonstration that successive same-different recognition judgments for musical sequences alternated between the ears, showing minimal recognition accuracy at approximately 3–4 alternating cps, this point being affected by the rate of presentation of the melodies. Although such findings discourage a purely syllabic disruption interpretation, they do sustain the principle exemplified in the present Experiment 3, that the amount of speech (or melody) heard in a particular ear segment moderates the effect.

Successful perception of alternating speech requires the registration of the signal from each input source and subsequent integration of these signals into a single fused stream for analysis, presumably in language-processing regions in the dominant (usually left) hemisphere (Hickok & Poeppel, 2000; Scott & Johnsrude, 2003). The combination of reduced peripheral acuity and potentially reduced efficiency in higher level auditory processing is sufficient to account for the older adults' reduced intelligibility scores at all alternation rates in all three experiments. The important question relates not to the age difference in overall performance level but to the observed invariance across age in the alternation rates that produce minimal intelligibility.

If the time it takes to switch attention between the two ears accounts for the point of minimal intelligibility, one would expect to see a significant age difference in the alternation rate that yields this minimal intelligibility point. That is, although linguistic and morphological knowledge is generally preserved in old age (see Wingfield & Stine-Morrow, 2000, for a review), older adults show significant slowing in a variety of perceptual and cognitive tasks (Cerella, 1994; Salthouse, 1991, 1996), to include the speed of switching attention (Panek & Rush, 1981; Verhaeghen & Basak, 2005; Waugh & Barr, 1980). To the extent that this is so, the similarity in the points of minimal intelligibility for the younger and older adults in each of the three experiments runs counter to a time-based attentional lag account for the occurrence of the point of minimal intelligibility.

Speech recognition is known to be dependent on context, with the identification of any single speech sound heavily influenced by neighboring elements. Were a preliminary analysis of the successive parts of the originally complementary speech segments conducted in the cerebral hemisphere contralateral to their ear of arrival, it would have to be conducted on speech segments deprived of their linguistic surround. The result would be speech recognition based on an attempted integration of poorly encoded representations of what would ordinarily have been juxtaposed segments in nonalternated speech. This contextual deprivation argument could account for the reduced perceptual performance observed when speech is alternated between two sources.

This characterization would also account for the importance of speech rate because a faster speech rate would allow a larger contextual window for any analysis of an ear's segment than speech heard at a slower rate. To the extent that speech rate affects the point of minimal intelligibility, increasing the speech rate would have the effect of shifting this point to a faster alternation rate (Huggins, 1964; Wingfield & Wheale, 1975b). Experiment 3 shows this to be true to an identical degree for both younger and older listeners.

When the alternation rates become very rapid, individual speech segments again become recognizable. This would be so because, although the speech segments to each ear would be too short to be recognized in isolation, increasingly rapid alternation would result in short speech segments separated by progressively smaller intervening gaps perceptually bridged to be heard as single acoustic events with individual elements restored within their contextual fields (e.g., Huggins, 1971).

The challenging question that remains is this: Why is intelligibility poorest specifically at 3–5 alternating cps but better when speech is alternated at rates slower or faster than this region? The alignment of critical points for younger and older adults suggests that the answer is not purely a time-based consequence of lagging attention, nor, as Rupf et al. (1971) and Samuel (1991) have shown, due specifically to a point of maximal disruption of syllables. Rather, the point of minimal intelligibility may reflect an interaction of three factors: (1) the rate of disruption induced by breaking the speech stream between two sources, (2) the amount of contextual information per ear, and (3) the size of the silent gaps separating the speech

elements that must be perceptually bridged. What is clear is that these mechanisms operate in a similar fashion for both younger and older adults, suggesting that perceptual fusion and lexical identification operate at an automatic level and are unaffected by attentional slowing and other sensory and cognitive changes.

#### AUTHOR NOTE

This research was supported by NIH Grant AG19714 from the National Institute on Aging. We gratefully acknowledge support from the W. M. Keck Foundation. We thank Arthur Samuel and an anonymous reviewer for helpful suggestions on an earlier version of this article, and Hiram Brownell for statistical advice. Correspondence concerning this article should be addressed to A. Wingfield, Volen National Center for Complex Systems, MS-013, Brandeis University, Waltham, MA 02454-9110 (e-mail: wingfield@brandeis.edu).

#### REFERENCES

- ANDERSON, N. D., CRAIK, F. I. M., & NAVEH-BENJAMIN, M. (1998). The attentional demands of encoding and retrieval in younger and older adults: 1. Evidence from divided attention costs. *Psychology & Aging, 13*, 405-423.
- BENJAMIN, B. J. (1997). Speech production of normally aging adults. *Seminars in Speech & Language, 18*, 135-141.
- BERGMAN, M. (1980). *Aging and the perception of speech*. Baltimore: University Park Press.
- BREGMAN, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- CERELLA, J. (1994). Generalized slowing in Brinley plots. *Journal of Gerontology, 49*, P65-P71.
- CHERRY, E. C., & TAYLOR, W. K. (1954). Some further experiments upon the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 26*, 554-559.
- GORDON-SALANT, S., & FITZGIBBONS, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech & Hearing Research, 36*, 1276-1285.
- HALL, J. W., & MUELLER, H. G. (1997). *Audiologists' desk reference* (Vol. 1). San Diego: Singular.
- HICKOK, G., & POEPEL, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences, 4*, 131-138.
- HUGGINS, A. W. F. (1964). Distortion of the temporal pattern of speech: Interruption and alternation. *Journal of the Acoustical Society of America, 36*, 1055-1064.
- HUGGINS, A. W. F. (1967). Distortion of the temporal patterns of speech by syllable-tied alternation. *Language & Speech, 10*, 133-140.
- HUGGINS, A. W. F. (1971, August). *Perception of temporally segmented speech*. Paper presented at the 7th International Congress of Phonetic Sciences, Montreal.
- KEPPEL, G. (1991). *Design and analysis: A researcher's handbook* (3rd ed.). Englewood Cliffs, NJ: Prentice Hall.
- MACKEN, W. J., TREMBLAY, S., HOUGHTON, R. J., NICHOLLS, A. P., & JONES, D. M. (2003). Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. *Journal of Experimental Psychology: Human Perception & Performance, 29*, 43-51.
- MILLER, G. A., & LICKLIDER, J. C. R. (1950). The intelligibility of interrupted speech. *Journal of the Acoustical Society of America, 22*, 167-173.
- MORRELL, C. H., GORDON-SALANT, S., PEARSON, J. D., BRANT, L. J., & FOZARD, J. L. (1996). Age- and gender-specific reference ranges for hearing level and longitudinal changes in hearing level. *Journal of the Acoustical Society of America, 100*, 1949-1967.
- NEISSER, U. (1966). *Cognitive psychology*. New York: Appleton-Century-Crofts.
- PANEK, P. E., & RUSH, M. C. (1981). Simultaneous examination of age related differences in the ability to maintain and reorient auditory selective attention. *Experimental Aging Research, 7*, 405-416.
- PICHORA-FULLER, M. K., & SOUZA, P. E. (2003). Effects of aging on auditory processing of speech. *International Journal of Audiology, 42*, 2S11-2S16.

- RUPF, J. A., HUGHES, G. W., & HOUSE, A. S. (1971). Speech-synchronized versus periodic interaural switching of speech. *Journal of the Acoustical Society of America*, **49**, 608-610.
- SALTHOUSE, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.
- SALTHOUSE, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, **103**, 403-428.
- SAMUEL, A. G. (1991). Perceptual degradation due to signal alternation: Implications for auditory pattern processing. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 392-403.
- SCHNEIDER, B. A., & PICHORA-FULLER, M. K. (2000). Implications of perceptual deterioration for cognitive aging research. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (2nd ed., pp. 155-220). Mahwah, NJ: Erlbaum.
- SCHNEIDER, W., DUMAIS, S. T., & SHIFFRIN, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 1-27). San Diego: Academic Press.
- SCHUBERT, E. D., & PARKER, C. D. (1955). Addition to Cherry's findings on switching speech between the two ears. *Journal of the Acoustical Society of America*, **27**, 792-794.
- SCOTT, S. K., & JOHNSRUDE, I. S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, **26**, 100-107.
- SMITH, B. L., WASOWICZ, J., & PRESTON, J. (1987). Temporal characteristics of the speech of normal elderly adults. *Journal of Speech & Hearing Research*, **30**, 522-529.
- SPEAKS, C., & TROOIJEN, T. T. (1974). Interaural alternation and speech intelligibility. *Journal of the Acoustical Society of America*, **56**, 640-644.
- SUSSMAN, E., ČEPONIENĖ, R., SHESTAKOVA, A., NÄÄTÄNEN, R., & WINKLER, I. (2001). Auditory stream segregation processes operate similarly in school-aged children and adults. *Hearing Research*, **153**, 108-114.
- TUN, P. A., WINGFIELD, A., STINE, E. A. L., & MECASAS, C. (1992). Rapid speech processing and divided attention: Processing rate versus processing resources as an explanation of age effects. *Psychology & Aging*, **7**, 546-550.
- VERHAEGHEN, P. (2003). Aging and vocabulary scores: A meta-analysis. *Psychology & Aging*, **18**, 332-339.
- VERHAEGHEN, P., & BASAK, C. (2005). Ageing and switching of the focus of attention in working memory: Results from a modified N-Back task. *Quarterly Journal of Experimental Psychology*, **58A**, 134-154.
- WAUGH, N. C., & BARR, R. A. (1980). Memory and mental tempo. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging: Proceedings of the George A. Talland Memorial Conference* (pp. 251-260). Hillsdale, NJ: Erlbaum.
- WINGFIELD, A. (1977). The perception of alternated speech. *Brain & Language*, **4**, 219-230.
- WINGFIELD, A., & STINE-MORROW, E. A. L. (2000). Language and speech. In F. I. M. Craik & T. A. Salthouse (Eds.), *Handbook of aging and cognition* (2nd ed., pp. 359-416). Mahwah, NJ: Erlbaum.
- WINGFIELD, A., TUN, P. A., KOH, C. K., & ROSEN, M. J. (1999). Regaining lost time: Adult aging and the effect of time restoration on recall of time-compressed speech. *Psychology & Aging*, **14**, 380-389.
- WINGFIELD, A., & WHEALE, J. L. (1975a). Interaural alternation, information load, and speech intelligibility. *Journal of the Acoustical Society of America*, **57**, 1219-1220.
- WINGFIELD, A., & WHEALE, J. L. (1975b). Word rate and intelligibility of alternated speech. *Perception & Psychophysics*, **18**, 317-320.
- ZACHARY, R. (1986). *Shipley Institute of Living Scale* (Rev. ed.). Los Angeles: Western Psychological Services.

(Manuscript received March 23, 2007;  
revision accepted for publication September 11, 2007.)