



Changes in Temporal Processing of Speech Across the Adult Lifespan

Diane Kewley-Port¹, Larry E. Humes¹, Daniel Fogerty¹

¹Department of Speech and Hearing Sciences, Indiana University, Bloomington, Indiana
 kewley@indiana.edu, humes@indiana.edu, dfogerty@indiana.edu

Abstract

Speech is a rapidly varying signal. Temporal processing generally slows with age and many older adults experience difficulties in understanding speech. This research involved over 250 young, middle-aged and older listeners. Temporal processing abilities were assessed in numerous vowel sequence tasks, and analyses examined several factors that might contribute to performance. Significant factors included age and cognitive function as measured by the WAIS-III, but not hearing status for the audible vowels. In addition, learning effects were assessed by retesting two tasks. All groups significantly improved vowel temporal-order identification to a similar degree, but large differences in performance between groups were still observed.

Index Terms: speech perception, aging, temporal processing

1. Introduction

Aging generally has a negative impact on temporal processing. Speech understanding requires rapid processing of a complex signal and difficulties in speech understanding often increase with age. In an ongoing study at Indiana University (the Vantage project), a large battery of temporal-processing tasks is being completed by young, middle-aged and older adults using analogous methods in each of three senses: hearing, vision, and touch. All told, the battery consisted of three phases with over thirty measures, requiring approximately 60 hours to complete. To date, reports from this project [1, 2, 3] have consistently found age group differences on the vast majority of the temporal-processing measures, with the young adults performing the best, older adults performing the worst, and middle-aged adults somewhere in between.

The focus of the present report is on four auditory tasks from Phase II of the test battery that involve vowel sequences recorded in words from a male talker. Two-item or four-item sequences were presented with the amount of time between the vowels experimentally varied. The listener was asked to identify the vowels in the correct order. Thus, a direct measure of the ability to understand the short sequences of vowels was obtained for listeners from 18 to 87 years of age. Insight into the roles of cognitive function and learning can be inferred from the different levels of task difficulty, as well as observing changes in performance when two of the tasks were retested after substantial additional participation in the Vantage test battery.

The questions posed in the present report are: (1) What changes in temporal processing of vowel sequences are observed across the adult lifespan? (2) Given that age is a factor, how is performance improved with learning, i.e. could training improve the performance of older listeners to reach that of younger listeners?

2. Method

2.1. Participants

Young adults (N=76) Age: M = 22.0 yrs (18-30 yrs)
 Middle-aged adults (N=32) Age: M = 48.4 yrs (40-55 yrs)
 Older adult listeners (N=157) Age: M = 71.0 yrs (60-87 yrs)

All participants were screened on a variety of measures to ensure they could perform most of the tasks in the full Vantage battery (for details see [1]). This included that the listeners could identify the vowels in isolation at 90% correct. In addition, participants had standard audiometric tests and completed a full cognitive assessment using the WAIS-III.

2.2. Stimuli

The stimuli were recorded from a Midwestern male talker in carrier sentences containing the words “pit, pet, pot, put”. Words were digitally edited leaving only the voiced portions. These voiced words were modified by STRAIGHT to be exactly 70-ms long with an F0 of 100 Hz. To reduce the effects of high frequency hearing loss, they were then low-pass filtered at 1800 Hz. These resulting stimuli still sounded quite natural and in screening tasks for all subjects were identified with 90% or better accuracy [2]. These stimuli were equalized in RMS and calibrated for presentation at 83 dB SPL in ER-3A earphones.

2.3. Procedures

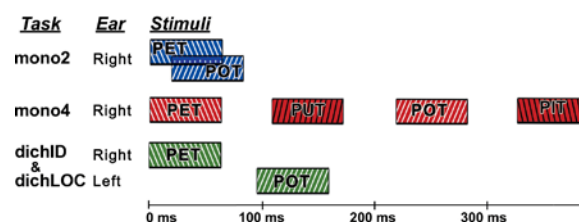


Figure 1: Schematic diagram of vowel sequences for one temporal-order trial. Words are located on the time axis to represent approximately SOA thresholds for the young listeners.

The Vantage project consisted of four sets of tests beginning with audiometric tests and WAIS-III intelligence tests, among other standard tests. The Vantage test battery was divided into three phases: Phase 1 sensory and gap detection thresholds [1]; Phase II temporal-order identification tasks; and Phase 3 masking tasks. The present paper involves only the auditory modality for the four temporal-order identification tasks from Phase II [2]: monaural two-item (mono2), monaural four-item

(mono4), dichotic two-item identification (dichID) and dichotic two-item location (dichLOC) or ear identification. Shown in Figure 1 are examples of single trials for each task. The right ear was used for listening in monaural tasks, and both ears in dichotic tasks (right or left ear first was randomly selected).

Listeners completed the four tasks in the order shown in Fig. 1. For each task, the time between the onsets of items (stimulus onset asynchrony, SOA) had 6 different SOA values that were selected as a range appropriate for each listener from familiarization tasks. Responses were collected using large buttons on a display labeled with the four words, “pit, pet, pot, put”, one column for each item in the sequence. Thus, listeners were required to correctly identify the vowels and the order in which they occurred (testing for dichLOC required listeners only to identify the ear presentation order regardless of vowel identity). SOA values at the 50% performance level (75% performance for dichLOC) for each block were recorded in milliseconds. Final average values were calculated for each individual over three blocks for each task. These are called temporal-order thresholds.

At the end of the original three phases of the Vantage test battery, a small number of tasks were selected for retesting to determine the effects of any learning that might have resulted from long-term exposure to the temporal-order tasks and the speech stimuli. The monaural and dichotic two-item tasks were selected for this comparison and were administered the second time following Phase III testing in all modalities.

3. Results

3.1. Temporal-order thresholds

Group results consist of 12 temporal-order thresholds for each group and task (3 age groups X 4 tasks). First individual temporal-order thresholds for each of the 12 conditions were examined. Nearly all conditions and groups had a small number of unusually high SOA values (outliers). Therefore, median values were calculated as the descriptive statistic for each of the 12 conditions. The median SOA thresholds for each age group and condition are shown in Table 1 and Figure 2.

Tasks	Young	Middle-Aged	Older	Average
mono2	17.3	30.1	49.0	31.1
mono4	89.3	165.4	126.6	127.1
dichID	90.2	104.0	115.6	103.3
dichLOC	107.1	109.9	119.1	112.0

Table 1. Temporal-order thresholds in milliseconds.

Examining Table 1, it is clear that the monaural two-item task had the lowest thresholds within each listener group. It is somewhat surprising that the two 70-ms long vowels in the sequence largely overlapped at threshold for most listeners, as can be seen schematically on Figure 1. That is, for the young listeners, they heard about 10 ms of the first vowel, 60 ms of overlapping vowels and 10 ms of the second vowel when they identified the vowel sequence 50% correct. While the older listeners needed considerable more time to perform the task, as

a group they also could identify the vowels correctly with overlap, namely 20 ms overlap.

From the average across age groups, Table 1 also shows that the monaural four-item task was vastly more difficult than the monaural two-item task by a factor of about four. As indicated in Figure 1, to identify a sequence of four words, they had to be separated widely in time.

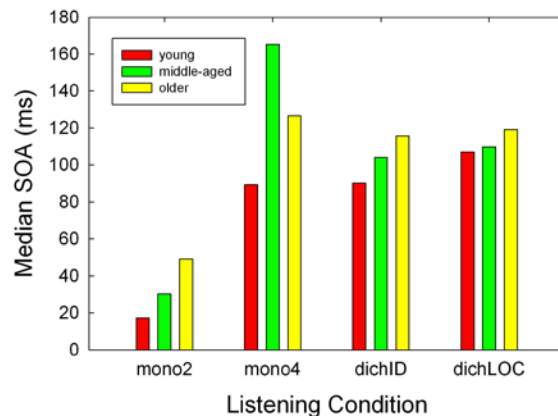


Figure 2: Performance of each of the three age groups in the four temporal-order vowel tasks. SOA is the time between onsets of each item in the sequence in milliseconds.

To examine group differences in more detail, a non-parametric Kruskal-Wallis test, similar to an analysis of variance, was conducted on the SOA values. The effects of age group were significant for each of the four temporal-order tasks. Follow-up paired-comparison analyses were performed using the Mann-Whitney non-parametric test. The general trend seen in Figure 2 was for best performance by the young group, worst by the older group with the middle-aged in between. Exceptions included mono4 where the middle-aged group performed worse and the dichLOC task where young and middle-aged groups had similar results. It is notable that for the mono2 task with the lowest thresholds, thresholds increased rather linearly across age groups with older listeners having 183% higher thresholds than young adults. Threshold elevation with age was considerably reduced for the other conditions, such that on average thresholds of older adults increased by 66% relative to young adults.

The general trend for task differences was for the mono2 to have the shortest SOAs, mono4 the longest, and the two dichotic tasks to have similar SOAs between mono2 and mono4. Significant inter-task differences were found between mono2 and mono4 as well as the mono2 and dichotic tasks by means of a Wilcoxon Signed-Ranks Test and indicate that there are different auditory processing demands across the tasks. Listeners can perform fastest when there are only two vowel sounds in one ear. Doubling the number of vowel sounds quadruples the amount of time needed to process the vowels. This large processing time difference is likely associated with additional cognitive processing, and may include a memory component given that four rather than two vowels must be correctly identified. The difference between the SOAs for the mono2 and mono4 tasks was not uniform over age. The middle-aged group needed considerably more time than either the young or older groups to identify the

vowels in the mono4 task. This could be an effect of sample size because the middle-aged group was much smaller than the others. Therefore, group differences for mono4 may change as ongoing enrollment in the Vantage project increases the number of middle-aged listeners.

Considering differences in ear presentation, having two vowels, but presenting them to both ears dichotically (dichID) instead of one (mono2), at least doubles the amount of processing time. This doubling occurred even though the listener's task was the same, and indicates that engaging the central auditory system in the dichotic task significantly slows speech processing.

Although group differences have been described thus far, the series of Vantage experiments was designed to examine the effects of aging in a large individual-differences paradigm [3]. Therefore, data on listeners' hearing status and cognitive function were also collected. Hearing should not have been a factor given that older listeners' hearing was near normal in the frequency range of the stimuli (low-pass filtered vowels at 1800 Hz). In fact, correlation between average hearing thresholds and temporal-order identification was not significant suggesting audibility differences in the older listeners did not affect vowel-sequence processing.

Cognitive function was assessed using the fifteen subtests from the WAIS-III. Regression analysis was applied to examine the contribution of age and cognitive factors to temporal processing individually for older listeners. The goal was to obtain insights into reasons for the high variability seen among the older listeners. First it was necessary to reduce the 15 cognitive measures to a smaller set reflecting overlap between WAIS-III subtests. Using factor analysis for the older listeners, three principal components accounted for most of the variance. These were interpreted as general cognitive processing, cognitive processing speed and incidental learning. Step-wise linear regression was performed, with age and the three cognitive factors as predictors of the four temporal-order tasks measures. Results indicated that two of the cognitive factors, general processing and processing speed, served as primary predictors for all four experimental tasks [2]. Age was not a factor within the older group. Therefore, cognitive factors appear to account for most of the variance in performance among older adults in these temporal-order tasks.

3.2. Learning effects in temporal-order thresholds

The second purpose of this report is to explore potential age-related differences in learning for these temporal-order tasks. The full protocol for all tasks in the multi-sensory Vantage project required more than 60 hours of testing. Therefore, there was considerable practice on temporal-processing tasks, in general, and substantial exposure to the vowel stimuli as well. The question naturally arose as to how much improvement in performance would be observed during the course of the project and would there be age differences in the amount of learning. As noted, the monaural two-item (mono2) and the dichotic two-item (dichID) tasks were selected for retest after project completion. In the retesting, the same participants, stimuli and procedures were used as described in the Method section above. The average duration of time between test and retest differed across groups: young, interval = 8.0 mo; middle-aged, interval = 6.3 mo; older, interval = 15.4 mo. However, the same number of tasks (auditory, visual

and tactile) occurring in the test-retest interval was the same for all listeners.

Individual SOA thresholds for test and retest tasks were measured in the same way (see Procedures). Group thresholds (means) for the three age groups and the two tasks were calculated and are shown in Figure 3. Performance for all age groups improved significantly from test to retest by means of a paired t-test for both the monaural and dichotic tasks. All groups improved by about the same amount, 12 ms. There was no significant difference in the size of this improvement across groups. However, in terms of percentage improvement, because the test thresholds were significantly lower for the young group than for the middle-aged who, in turn, had lower thresholds than the older group, the relative improvement in performance, expressed as a percentage, was rather different. In the monaural task, the percentage improvement was 51%, 52% and 33% for young, middle-aged and older groups, respectively. The percentages were lower for the dichotic task, but the trend was similar with 10%, 9% and 6% improvement for young, middle-aged and older groups. Thus, with additional exposure, all groups improved, but the rate of improvement of the older group is lower indicating that even with training procedures to promote learning older listeners would not catch up with young listeners.

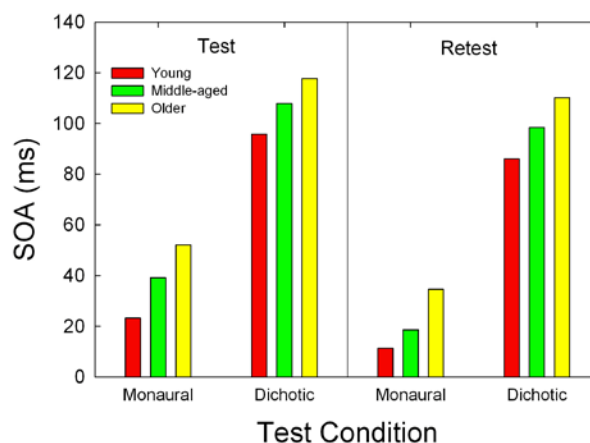


Figure 3. Performance of each of three age groups in two temporal-order tasks with two-item vowel sequences with either monaural or dichotic presentation. The panel on the left displays thresholds obtained during the initial test time, and the panel on the right displays thresholds retested at the completion of the main project.

Given that all groups improved about 12 ms on average, an additional question was whether this improvement was relatively constant for individual listeners. This was tested by means of correlation between the test and retest thresholds for all listeners in either the monaural or dichotic condition. The scatterplot for the monaural condition is shown in Figure 4.

The best-fitting linear regression line (dotted line) in Figure 4 for all listeners in the monaural condition has a test-retest correlation of $r = 0.8$. This is the same correlation that was found for the dichotic listening condition, $r = 0.8$. Thus in both conditions retested, the relative rank-ordering of subject performance was stable. However, as shown in Figure 4 for the monaural 2-item task, the regression fit indicates that those

with larger initial SOAs, regardless of age, tended to improve more than those with smaller initial SOAs. To address whether the aforementioned age group differences in test-retest intervals affected thresholds, additional correlational analyses were calculated. Poor correlations were obtained between the duration of the test-retest interval and change in SOA thresholds for both mono2 ($r^2 = 0.03$) and dichID ($r^2 = 0.01$) indicating there was no dependence between improved thresholds and size of the interval.

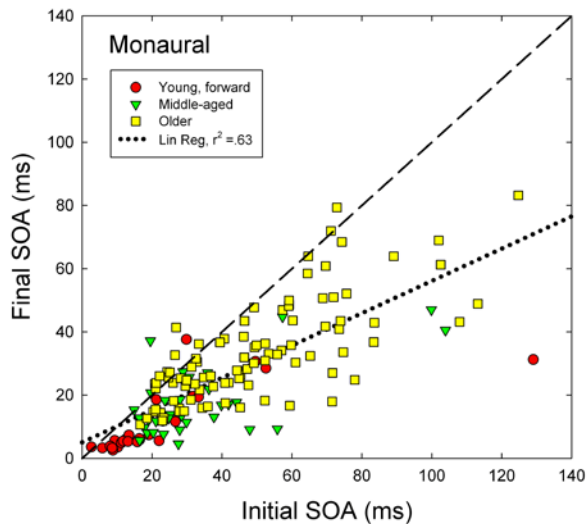


Figure 4: Scatterplot of individual retest (“final”) SOA values as a function of test (“initial”) SOA values for the monaural two-item task. The dashed line shows equivalent test-retest performance. The actual regression line is indicated by the dotted line for a correlation of $r = 0.8$.

4. Discussion

Results from this research measured how temporal processing of vowel sequences declines with age for a large group of listeners spanning 18 – 87 years of age. Previous studies have found age-related deficits in temporal processing for a variety of auditory stimuli [4, 5], but none previously has employed speech stimuli. The tasks here involved the presentation of naturally sounding vowels in sequences that required the correct identification of the vowels and their order. Fluent speech also requires rapid identification of vowels (and consonants). The common complaint of older people that someone cannot be understood because they speak too fast may be related to the changes in auditory processing speed shown here.

In the mono2 temporal-order task with the lowest task thresholds, older listeners required twice the processing time that young listeners did. As the task thresholds increased, the gap in extra processing time between young and old narrowed considerably at the group level. However, it is important to note that many middle-aged and older adults performed as well as young listeners in several tasks. This can be seen in the scatterplot, Figure 4. Looking at just the thresholds for the initial SOA values on the x-axis, there is considerable overlap between groups. Of special interest is the large number of older listeners (yellow squares) who have thresholds in the upper range (20 – 35 ms) of the young listeners (red circles). This is the same range where nearly all the middle-aged data

fall. Thus, even though there is a large and significant decrease in performance with age for temporal-order identification of vowels, many older listeners continue to perform as well as younger listeners, even into their 80s.

Nonetheless, in Figure 4 it can also be seen that nearly one-half of the older listeners’ thresholds exceeded those of the majority of young and middle-aged listeners, e.g. had SOAs greater than 50 ms. A subset of tasks were retested after considerable exposure to other temporal experiments and vowel stimuli and these results shed light on whether older listeners’ performance might approach that of younger listeners with training. Results showed that all three age groups improved by a similar amount of about 12 ms. Further analysis showed that this meant that the rate of improvement for the older listeners was smaller than that for younger listeners. Therefore, an implication of the test-retest experiments is that older listeners who initially demonstrate slower auditory processing of the speech sounds are also slower to improve performance with additional training.

5. Conclusions

The ability of listeners to rapidly process information in vowel sequences worsened substantially with age, about 66% in our experiments. Slowing due to cognitive factors was observed at several levels (across task and individually). Although improved performance was observed for additional exposure to the stimuli and temporal-order tasks, group results for older listeners indicated that even with additional training older adults would not achieve the performance of younger adults. It is notable, however, that individually there were many older adults (60-87 yrs) who performed as well as many young adults (18-30 yrs.)

6. Acknowledgements

The authors would like to thank Tom Busey and Jim Craig for their significant contributions to this project. We also thank Dana Kinney, as well as several graduate research assistants involved with this project. This work was supported, in part, by NIA R01 AG022334 and by NIH-NIDCD Training Grant No. T32-DC00012.

7. References

- [1] Humes, L. E., Busey, T. A., Craig, J. C., and Kewley-Port, D., “The effects of age on sensory thresholds and temporal gap detection in hearing, vision, and touch,” *Percept. Psychophys.* 71:860–871, 2009.
- [2] Fogerty, D., Humes, L.E., Kewley-Port, D., “Auditory temporal-order processing of brief vowel sequences by young and elderly listeners”. *J. Acoust. Soc. Am.*, 127(4):2509-2520, 2010.
- [3] Humes, L. E., Kewley-Port, D., Fogerty, D., and Kinney, D. “Measures of hearing threshold and temporal processing across the adult lifespan,” *Hear. Res.* 264:30-40 2010.
- [4] Trainor, L. J., and Trehub, S. E., “Aging and auditory temporal sequencing: Ordering the elements of repeating tone patterns,” *Percept. Psychophys.*, 45: 417–426, 1989.
- [5] Fitzgibbons, P. J., and Gordon-Salant, S., “Auditory temporal processing in elderly listeners,” *J. Am. Acad. Audiol.*, 7: 183–189, 1996.