Published in final edited form as:

Ear Hear. 2013 March ; 34(2): 160–167. doi:10.1097/AUD.0b013e31826a8ea7.

Aging and the perception of temporally-interleaved words

Karen S. Helfer, Ph.D.¹, Christine R. Mason, M.S.², and Christine Marino, B.S.¹

¹University of Massachusetts Amherst, Department of Communication Disorders

²Boston University, Department of Speech, Language, and Hearing Sciences

I. INTRODUCTION

The most problematic listening situations reported by older individuals are when they must comprehend one message in the presence of one or more competing talkers (Agus et al. 2009). An increasing number of research studies have verified these reports, with older listeners experiencing more difficulty than younger individuals in tasks of competing speech perception (Carhart & Tillman 1970; Duquesnoy 1983; Tun & Wingfield 1999; Tun et al. 2002; Li et al. 2004; Humes et al. 2006; Singh et al. 2008; Helfer & Freyman 2008; Agus et al. 2009; Rossi-Katz & Arehart 2009; Helfer et al. 2010; Ezzatian et al. 2011).

Competing speech can produce both energetic masking (when peripheral resources that are needed to respond to the target speech are not available because they are being recruited to process the masking speech) and informational masking. Informational masking can take a number of different forms. One aspect of this type of masking is related to confusion or uncertainty regarding which signal is to be attended and which is (or are) to be ignored. Research to date suggests that aging does not appear to make individuals more susceptible to this form of informational masking (e.g., Li et al. 2004; Helfer & Freyman 2008; Agus et al. 2009).

Then, why do older individuals have so much difficulty in these competing speech situations? The ability to understand a message in the presence of competing speech requires that the listener not only comprehend the message of interest but also separate the incoming auditory signal in streams, figure out what or to whom to attend, and ignore the background conversations. The hearing loss experienced by most older individuals undoubtedly is a strong contributor to problems understanding the target message, since it leads to increased susceptibility to energetic masking (e.g., Arbogast et al. 2002; Li et al. 2004; Moore 2007; Agus et al. 2009). Results of some studies suggest that only minimal differences between older and younger listeners in competing speech tasks are found once peripheral factors are taken into account (Schneider et al. 2000; Li et al. 2004; Murphy et al. 2006). However, given that the ability to attend to or suppress information takes cognitive mediation and that aging affects many cognitive functions, it is not surprising that vulnerability to energetic masking cannot entirely explain the problems that older adults have when speech is masked by other speech. For example, older listeners experience a greater decrement in speech understanding than younger listeners when comparing intelligible to non-intelligible speech maskers that are equated for energy (e.g., Rossi-Katz & Arehart 2009; Helfer et al. 2010),

Corresponding Author: Karen S. Helfer, University of Massachusetts Amherst, Dept. of Communication Disorders, 358 N. Pleasant St., Amherst, MA 01003, Phone: 413-545-4014, FAX: 413-545-0297, khelfer@comdis.umass.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

which cannot be explained by susceptibility to energetic masking. The finding that older individuals are less able to take advantage of the target and masking speech being spoken by opposite-sex talkers (e.g., Helfer & Freyman 2008), even when audibility for the older listeners is assured via spectral shaping (Humes et al. 2006; Humes & Coughlin 2009) also suggests that non-peripheral factors contribute to older adults' problems in complex listening situations.

The pattern of results to date, we believe, supports the idea that older listeners are susceptible to another type of informational masking: distraction caused by a meaningful speech signal. There is ample evidence to suggest that aging brings about a decline in the ability to inhibit meaningful auditory information (e.g., Sommers & Danielson 1999; Tun & Wingfield 1999; Tun et al. 2002; Helfer & Freyman 2008; Helfer & Vargo 2009). Problems with inhibition may lead to reduced ability to ignore distracting signals as well as difficulty deleting information that is no longer relevant (e.g., Lustig et al. 2001). Reduced working memory capacity may underlie problems ignoring background conversations, as individual differences in working memory are related to susceptibility to meaningful auditory distraction (Beaman 2004; Sorqvist 2010). Moreover, meaningful background speech may interfere with the encoding of information into memory, as processing resources must be expended on segregating the target message from the background, leaving fewer remaining resources for the encoding process (Heinrich et al. 2008). Research that has attempted to identify connections between memory/working memory and competing speech perception has led to no definitive answer (e.g., Humes et al. 2006; Neher et al. 2009, 2011).

The purpose of this study was to identify how aging affects susceptibility to the part of informational masking that involves ignoring or inhibiting competing speech. Separating energetic and informational masking in a competing speech task is complicated, as in most paradigms speech maskers cause both types of disruption. In the present study, we used a non-simultaneous competing speech task to emphasize the role of higher-level factors. The specific task used was one developed by Kidd et al. (2008) in which listeners hear two temporally interleaved five-word sentences. This task (which is an adaptation of a similar paradigm described in Broadbent, 1952) requires participants to repeat back every other word while ignoring the intervening words. Non-simultaneous presentation of the target and interfering words minimizes peripheral interference (Kidd et al. 2008; Best et al. 2011). Hence, age-related changes noted on this task, for the most part, cannot be explained by differences in susceptibility to energetic masking¹.

We compared older and younger adults' understanding of interleaved words while manipulating a number of factors that are known to affect the processing of these stimuli. One variable that was examined in the present work was whether the to-be-ignored and to-be-attended words came from the same or from different spatial locations. Studies of simultaneous masking have found a reduction in older adults' ability to benefit from spatial separation of a target speech sound from a masking speech sound (e.g., Murphy et al. 2006; Marrone et al. 2008; Helfer et al. 2010). It appears that much of the reduced spatial release from masking obtained by older adults in competing speech tasks can be attributed to peripheral factors (e.g., Marrone et al. 2008). However, there is evidence suggesting that cognitive abilities may limit older adults' benefit from spatial separation of speech sources (e.g., Neher et al. 2009). In theory, spatial benefit obtained using non-overlapping

¹Based on the finding that performance obtained with no masker and with interleaved samples of noise is equivalent (and better than that obtained with interleaved speech) in young normally-hearing listeners, Kidd et al. (2008) concluded that it was unlikely that non-simultaneous energetic masking affects this task. It is possible that our older subjects did experience some peripheral interference due to increased susceptibility to forward masking. However, we believe that any effect of non-simultaneous masking should influence all masked conditions equally and therefore differences among the experimental manipulations, or between different types of benefit across groups, must be attributable to some other factor(s).

temporally-interleaved stimuli should not be affected by peripheral hearing loss, since there is little or no energetic masking to release. Hence, determining how spatial separation alters the perception of these stimuli will give insight into how non-peripheral factors affect spatial hearing in older adults.

We also examined the influence of two other variables: talker consistency (that is, whether the same talker was used for each target word within a trial, or whether a different talker was used for each word) and syntactically correct vs. random word order. Kidd et al. (2008) found both of these factors to be important determinants of younger adults' performance on this interleaved word task. There is reason to believe that aging may affect individuals' use of each of these variables. Evidence suggests that older adults are less able to encode and/or use information about the talker's voice (e.g., Naveh-Benjamin & Craik 1995; Yonan & Sommers 2000). Problems encoding or using voice information could lead to listeners' perception of different voices being more similar to one another. This would be expected to result in a smaller difference on talker-same vs. talker-different trials between our older and younger listeners. Difficulty using indexical information also might be expected to lead to reduced ability to use differences between the target and masking talker to help remember the words. If this is the case, we might expect to see larger group differences in talker-consistent trials vs. talker-varied trials.

Past research also has demonstrated that older adults rely more on top-down information (like sentence context) than do younger listeners (e.g., Pichora-Fuller et al. 1995; Wingfield 1996; Wingfield & Tun 2001). This would lead to the expectation that older listeners would be at a greater relative disadvantage than younger listeners when a trial consists of five random words vs. when those words are in correct syntactic order (and, therefore, create a meaningful sentence that provides syntactic structure). Both of these ideas are tested in the present paper.

The task used in this investigation may place a relatively heavy load on memory, as participants must recall five out of 10 words heard. Age-related decrements in memory span may therefore play an important role in differences obtained between our older and younger participants. We used both forward digit span (to measure short-term memory) and backward digit span (to assess working memory) in order to determine how these aspects of memory affect the ability to remember and repeat back some words while ignoring others. We anticipated finding significant associations between our digit span metrics and performance on the interleaved sentence task.

II. MATERIALS AND METHODS

A. Participants

Two groups of listeners ($n = 13/\text{group}$) participated in the experiment, which was completed in one 1–2 hour session. One subject group consisted of college-aged adults (range 19–24 years, mean age 21 years) with normal pure-tone thresholds (although one subject was found to have thresholds of 30 dB HL at 4 kHz and 40 dB HL at 6 kHz in one ear; we decided to retain her data because she was one of our best performers). The second group contained older individuals (range 61–89 years, mean age 70 years) with no more than a moderate hearing loss up to and including 4 kHz. As subjects were screened for a negative history of otologic disorder, neurologic problems, and significant noise exposure, the probable etiology of hearing loss was presbycusis. All participants had bilaterally normal tympanograms on the day of testing, suggesting normal middle-ear functioning. Composite audiograms for both participant groups are shown in Fig. 1. It is apparent that some of the older subjects had substantial hearing loss at 6 kHz and/or 8 kHz. Analysis of Variance on the better-ear high-frequency average (average of pure-tone thresholds at 2, 4, 6, and 8 kHz) showed that the

older group had significantly poorer hearing than the younger group ($F(1,24) = 33.2, p < .001$).

B. Digit Span Test

Participants completed both forward and backward versions of a digit span test, administered via auditory presentation by presenting strings of digits of increasing length, beginning with two-digit strings and ending with a maximum of nine-digit strings. Digit span-forward is generally considered to be a measure of short-term memory, while digit span-backward is thought to provide insight into working memory (that is, the ability to both remember and manipulate items). An experimenter read lists of digits in a monotone voice at the rate of approximately 1/second using a normal-to-loud conversational vocal level. Subjects reported back the digits in the order heard (for digit span-forward) or in the reverse order (for digit span-backward). When a string was reported incorrectly, another string of the same length was presented; the test continued with a longer string if this trial was correct and was stopped if this trial was incorrect. An individual's digit span score was the length of the last correctly-reported string.

C. Interleaved speech recognition task

Speech recognition data were collected in a double-walled IAC sound chamber. Two loudspeakers (Realistic Minimus 7) were used to present the speech stimuli, one directly in front of the listener and the other at 60 degrees to the right of the participant. The loudspeakers were placed at a distance of 1.3 meters from the participant's head at a height of 1.2 meters from the floor, approximately ear-level height for the average seated listener. Stimuli were played out of a computer's sound card, attenuated (TDT PA4), amplified (TDT HBUF5), and then power amplified (TOA P75D) before being sent to the loudspeaker(s).

The stimuli consisted of a set of 40 monosyllabic words developed by Kidd et al. (2008) (see that paper for a full description of stimulus generation). A grid displaying the words is shown in Table I. This grid was available to the participants throughout data collection. The words are divided into five categories (names, verbs, numbers, adjectives, nouns). Each word was recorded in isolation from 16 talkers (eight male, eight female). The duration of each word ranged between 385 and 1051 ms (mean 624 ms). With the exception of the first two blocks of trials (see below), two five-word strings were presented per trial, with each sentence containing one word from each of the five categories. The two sentences were presented in interleaved fashion in which the ten words (five for each sentence) were concatenated but did not overlap in time. Participants were instructed to repeat back every other word beginning with the first word, and to ignore the intervening words. All words were equated for RMS amplitude and were played at a level of 60 dB A.

Three variables were manipulated in the present study. First, on half of all trials, all of the words were presented from the front loudspeaker (FF, for front-front); on the other half of trials, the to-be-attended five-word sentence (hereafter referred to as the target) was presented from the front and the to-be-ignored sentence (hereafter referred to as the masker) was presented from a loudspeaker located 60 degrees to the right of the listener (FR, for front-right). Second, on half of all trials, each of the two five-word utterances formed a syntactically-correct sentence; that is, the five words were presented in sequential order re: the response grid (see Table I). On the other half of trials, word order was random within the five-word sentences, with one word presented from each column. The third factor examined was talker consistency. On half of all trials, the target talker and the masker talker varied from word to word. During other trials, the same talker was used within a trial for the target, and a single (different) talker was used for the masker. Half of these talker-fixed trials used

same-sex talkers for the target and masker; the other half used different-sex talkers for target vs. masker.

Trial type was blocked, with 30 trials presented per block, and one block completed per condition. The first five trials in each block were considered practice and were not scored. Hence, each block of trials represents 125 scored items per subject ($25 \text{ trials} \times 5 \text{ words/trial}$). In order to measure baseline performance (that is, without interleaved distractor words) the first block consisted of single five-word sentences presented in syntactically-correct order; the second block was single five-word sentences presented in random order; both with silence in place of the distractor words. The remaining eight blocks (all combinations of the two spatial conditions, the two syntactic conditions, and the two talker consistency conditions) were presented in randomized order across listeners. A grid of the response alternatives was displayed throughout the test session. Participants repeated their responses, which were entered on-line by an experimenter. Although subjects were told to repeat the words in the order in which they were heard, responses were scored as correct even if they were said in incorrect order.

III. RESULTS

A. Digit Span

Average digit span-forward was 7.5 in our younger subjects and 7.1 in our older subjects, and the mean values for digit-span backward were 4.8 for the younger participants and 5.5 for the older subjects. Available norms for digit-span forward suggest that both groups of participants had short-term memory typical of individuals in their age range (with reported digit span-forward averages for younger individuals of 7.6 and for older adults, 7.1; Bopp & Verhaeghen 2005). ANOVA on the digit span scores indicated that the groups did not differ significantly for either the forward ($F(1,24) = .64$, $p = .432$) or the backward ($F(1,24) = 1.11$, $p = .301$) versions.

B. Recognition of interleaved words

Analyses of variance (ANOVA) on the percent-correct scores were used to analyze differences between groups and between conditions (spatially-coincident vs. spatially-separated, consistent target talker vs. varied target talker, and syntactically correct word order vs. random word order). It should be noted that analyses also were conducted with these data transformed into RAU (Studebaker 1985). Results of the two sets of data analyses (on percent-correct scores and on scores converted to RAU) were identical in terms of significant vs. non-significant findings.

The first two blocks of trials required participants to repeat back 5-word utterances without interleaved words. In the first block (in which the five words were presented in correct syntactic order), performance was near ceiling for both groups (younger: 99.2% correct; older: 98.6% correct). Performance on the second block, in which five words were presented in random order, was similarly high (younger: 97.3% correct; older: 95.3% correct). Analyses of variance showed non-significant group differences for both of these blocks (correct syntactic order: $F(1,24) = 1.23$, $p = .278$; random syntactic order: $F(1,24) = 1.44$, $p = .242$).

Performance for the interleaved stimuli is shown in Fig. 2. These data were analyzed using repeated-measures ANOVA with subject group as a between-subjects factor and syntactic condition (syntactically correct vs. random word order), talker consistency (fixed vs. variable talkers), and spatial condition (spatially coincident/FF vs. spatially separated/FR) as within-subjects factors. Results of the ANOVA showed significant main effects for all three within-subjects factors (syntactic condition: $F(1,24) = 83.38$, $p < .001$; talker consistency:

$F(1,24) = 46.01, p < .001$; spatial condition: $F(1,24) = 205.51, p < .001$). The main effect of subject group also was significant ($F(1,24) = 5.45, p = .029$) as were three interactions: spatial condition $\times$ group ($F(1,24) = 11.85, p = .002$), spatial condition $\times$ talker consistency ($F(1,24) = 11.03, p = .003$), and talker consistency $\times$ syntactic condition ($F(1,24) = 11.98, p = .002$). It is worth noting that the other two-way interactions involving group were marginally significant: talker consistency $\times$ group ($F(1,24) = 3.78, p = .064$) and syntactical condition $\times$ group ($F(1,24) = 3.49, p = .074$).

In order to explore the significant interaction between spatial condition and subject group, post-hoc t-tests were conducted with percent-correct scores averaged for all FF conditions and for all FR conditions. This analysis showed that group differences were statistically significant for the spatially-coincident FF condition ($t = 2.76, p = .011$) but not for the FR condition ($t = 1.62, p = .118$).

Simple difference scores were calculated for each of the within-subjects variables (spatial configuration, talker consistency, and syntactic structure) in order to further examine interaction effects. These difference scores are displayed in Fig. 3. The top panel shows spatial benefit, which was the simple difference between performance in the spatially-coincident F-F condition and the spatially-separated F-R condition. It is apparent that older listeners obtained greater spatial benefit than younger listeners. Moreover, spatial benefit was greater when the words were presented by multiple talkers within a trial (talker varied conditions) than when they were spoken by a single talker (talker fixed conditions).

The middle panel of Fig. 3 depicts the benefit of talker consistency (talker consistent scores – talker variable scores). It can be observed that the significant spatial condition $\times$ talker consistency interaction was influenced by the larger talker consistency benefit for spatially-coincident FF trials vs. spatially separated trials. Also of note is that older participants had a larger benefit from a consistent talker than younger subjects in all conditions, although recall that the interaction of talker consistency $\times$ group just missed statistical significance.

Syntactic benefit (performance on syntactically-correct conditions - random word order conditions) can be seen in the bottom panel of Fig. 3. Providing syntactic structure was more important when the talker varied from word to word than when the same talker was used within a trial. And although again the interaction with group just missed statistical significance, it can be observed that younger listeners derived more benefit in each condition from syntactically-correct word order than did the older adults.

C. Error types

Participants could make several kinds of responses: they could correctly report a target word; they could respond with a word from the masker; they could reply with a word that was in neither the target nor the masker; or they could omit a word from their response (i.e., their response to that trial contained fewer than five words). We analyzed the types of error responses by comparing masker errors (when listeners responded with a to-be-ignored word) and non-masker errors (when subjects replied with a word that was not presented during that trial or responded with fewer than five words). We speculate that if older adults had more difficulty than younger listeners in suppressing irrelevant information, there would be greater group differences in the proportion of masker errors than in the proportion of non-masker errors. Non-masker errors produced in simultaneous speech-on-speech masking tasks generally are attributed to energetic masking (e.g., Brungart 2001, Kidd et al. 2005, Ihlefeld & Shinn-Cunningham 2008, Iyer et al 2010). In the present task (for which energetic masking is not a primary issue), these errors likely reflect a strategy based on guessing.

Percentage of the two types of error responses are depicted in Fig. 4 for each group, with errors averaged across conditions. Although both groups of participants made more masker errors than non-masker errors, this was especially true of the older subjects. T-tests on the non-masker and masker error proportions showed a significant difference between groups for masker errors ($t = 2.37$, $p = .026$) but a non-significant difference for non-masker errors ($t = 1.92$, $p = .07$).

D. Correlations among measures

We calculated Pearson r correlations to examine the connections among digit span scores, high-frequency hearing loss (the average of better-ear thresholds at 2 kHz, 4kHz, 6 kHz, and 8 kHz), and recognition of interleaved words for each of the two groups. We speculated that memory would be more important for performance in the random word order conditions (in which the target utterance did not have syntactic support to help the listeners remember the words) than in the syntactically-correct word order trials. To test this idea, scores across conditions were averaged to get a mean percent-correct score for syntactically correct presentation and an average metric for random word order conditions. For the younger listeners, there were no significant associations between either digit span score or audiometric thresholds and performance on the interleaved word task. However, a different picture emerged for the older participants; results of this correlation analysis can be seen in Table II. Here, digit span-forward scores were strongly correlated with performance in both the random word order condition ($r = .67$, $p < .05$) and in the syntactically correct word order condition ($r = .70$, $p < .01$). These associations are depicted in the scatterplots shown in Fig. 5. It appears that short-term memory contributed to older subjects' ability to perform this task regardless of whether or not the target utterance provided syntactic structure. Neither digit span-backward scores nor the high-frequency pure-tone average was significantly associated with any other variable.

IV. DISCUSSION

Older participants in the present study experienced more difficulty with this interleaved speech task, as compared to younger listeners. Although there was minimal difference in performance accuracy between groups for the five-word strings presented without interleaved maskers, significant differences in scores were noted when the words were presented in interleaved fashion, especially when all words were presented from the same spatial location. These results provide evidence of a deficit in the ability to alternately attend to and ignore sequentially-presented words, likely related to selective attention and/or short-term memory deficits. Since the task used in this study greatly reduced energetic masking, the present results support the idea that problems experienced by older adults in competing speech tasks (e.g., Carhart & Tillman 1970; Duquesnoy 1983; Tun & Wingfield 1999; Tun et al. 2002; Li et al. 2004; Humes et al. 2006; Singh et al. 2008; Helfer & Freyman 2008; Agus et al. 2009; Rossi-Katz & Arehart 2009; Helfer et al. 2010) are not due entirely to increased susceptibility to energetic masking. This is also supported by the fact that different experimental conditions showed greater or lesser deficits in older adults relative to the younger group.

The finding of overall poorer performance by our older participants on this speech perception measure was not surprising. However, some of the other results found in the present study were unexpected. One of these concerns the finding of greater spatial benefit by our older subjects (see Fig. 3). A body of previous work has shown reduced spatial release from masking (SRM) in older adults during simultaneous speech-on-speech masking paradigms (e.g., Murphy et al. 2006; Marrone et al. 2008; Helfer et al. 2010). One question that has been raised is whether this reduced SRM in older adults is driven by threshold elevation or by other factors related to aging. The fact that peripheral hearing loss likely had

little influence on our non-simultaneous task (as indicated by the lack of a significant correlation between pure-tone hearing loss and recognition performance) suggests that reduced SRM in older listeners found in simultaneous masking tasks is related to primarily peripheral factors. However, results of a recent study by Best et al. (2011) using the same stimuli employed here may complicate this interpretation. In that experiment, young hearing-impaired listeners obtained less spatial release from masking than their normally-hearing counterparts. It should be noted, however, that there are a number of procedural differences between these two studies. For example, Best et al. (2011) used headphone presentation and interaural delay of the masker to produce spatial separation, while actual physical separation was used in the present investigation. Perhaps even more significant, the participants in Best et al. (2011) had more hearing loss than the older adults in our study, and the probable etiology of their hearing loss was different as well. Regardless, the present study's results suggest that older adults do not necessarily experience a breakdown in the ability to use spatial cues.

We had speculated that older participants might receive less benefit from having the same talker say each target word on a given trial if they had reduced ability to encode information about the talkers' voices. The effect of talker consistency did not differ significantly between groups, and it is clear that older adults were able to use this information at least as well as younger listeners (see Fig. 3). It is feasible that talker consistency effects reflected how listeners were able to benefit from reduced variability within a trial rather than the degree to which they were able to use voice information. Having a consistent talker from word to word might have served to reduce the uncertainty within a trial, which perhaps was able to override any age-related reduction in the ability to use information about talkers' voices. This argument is supported by the finding that the benefit of talker consistency was greater in the spatially-coincident FF conditions than in spatially-separated FR conditions. In other words, when listeners were given spatial cues that clarified the target words, the effect of talker consistency was much less important. When all of the words came from the same location, listeners' ability to benefit from a consistent talker took on greater importance. Moreover, differences in performance between younger and older listeners were larger for talker-fixed (as compared to talker-varied) trials, a finding that could be interpreted as indicating that older adults are less able to use consistent voice information to help remember the target words. It is clear that additional work needs to be done in order to understand how aging affects listeners' use of indexical information in tasks of speech understanding.

Another unexpected result was that younger listeners were able to use correct syntactic word order at least as well as older participants. A good deal of previous research suggests that older adults rely on higher-level information like context more than younger listeners, leading to greater contextual benefit as compared to that obtained by younger listeners (e.g., Pichora-Fuller et al. 1995; Wingfield 1996; Wingfield & Tun 2001). We offer two possible explanations for our findings. First, context in the present task involved supplying a syntactic framework, while context in other studies of speech understanding typically means giving the listener semantic cues within the sentence to help decipher a target word or words. Although it is true that providing a syntactical framework in the present study also served to give some semantic structure, it is perhaps not surprising to find differential effects of aging among tasks in which what is meant by "provision of context" differs substantially. Another possible explanation for our finding that older adults did not receive more benefit from context lies in the nature of target and masker presentation. During syntactically correct trials, both the target and masking words created syntactically plausible sentences. It is possible that older listeners had a more difficult time ignoring the syntactically-correct masker during syntactic trials (although, at least for younger listeners, linkage variables applied to the masker appear to have little effect on performance—see Kidd et al. 2008). Our

results suggest that the assumption that older adults make better use of context than younger listeners cannot be extended to all listening tasks.

Perhaps the most unexpected finding in the present study was the fact that digit span scores did not differ significantly between our younger and older groups of listeners. This suggests that our older listeners were likely quite high-functioning in terms of memory, compared to their age peers. Yet, it is interesting to note that forward digit span scores were significantly associated with performance on our speech recognition task, suggesting that age-related decline in short-term memory played a role in our findings. We had anticipated that memory would be important for this task, but were surprised to note that forward digit span, but not backward digit span, was significantly associated with recognition of interleaved words. It is not clear to us why this pattern of results was obtained. It could be argued that successful performance on the primary task required both short-term memory skills (in order to remember the target words) as well as working memory ability (as some manipulation of the to-be-remembered words was required in order to sort out which should be attributed to the target talker). In fact, the finding that older adults produced substantially more masker errors than younger adults supports the contention that our older participants had difficulty determining whether to-be-heard words were from the target or the masker. This result, however, also could be an indication of reduced ability to inhibit the to-be-ignored words. Currently, we are at a loss to explain why digit-span forward was more closely associated with task performance than was digit-span backward.

Of course, the task used in this study is not one that is encountered in day-to-day listening, as people rarely need to ignore every other word in actual communication. Previous research suggests that sequential streaming is related to the ability to separate the voices of simultaneous talkers (e.g., Mackersie 2003), giving credence to the relevance of this task to speech understanding in complex listening environments. The pattern of results shown in this study (significantly poorer perception of interleaved speech and a greater proportion of masker errors found in older vs. younger adults, combined with performance related to short-term memory but not degree of hearing loss) supports the idea that non-peripheral factors in part mediate age-related changes in speech understanding. However, it also is possible that subtle changes in supra-threshold abilities (which could be either peripherally or centrally mediated) influenced performance on this task. Irrespective of the source of the problems, our results support the idea that solutions to age-related speech understanding deficits need to look beyond treating presbycusis via methods designed only to increase audibility. Future work should be conducted to clarify the relative importance of understanding the message of interest vs. inhibiting distracting speech in order to make headway in improving how we remediate age-related problems with communicating in complex environments.

Acknowledgments

Portions of these data were presented at the American Auditory Society Conference, Scottsdale, AZ, March 2010. This work was supported by NIDCD R01 DC01623 (University of Massachusetts) and NIDCD P30 DC004663 (Boston University). The authors would like to thank Timothy Streeter for providing programming support for this project.

References

- Agus TR, Akeroyd MA, Gatehouse S, et al. Informational masking in young and elderly listeners for speech masked by simultaneous speech and noise. *J Acoust Soc Am*. 2009; 126:1926–1940. [PubMed: 19813805]
- Arbogast TL, Mason CR, Kidd GK Jr. The effect of spatial separation on informational and energetic masking of speech. *J Acoust Soc Am*. 2002; 112:2086–2098. [PubMed: 12430820]

- Beaman CP. The irrelevant sound phenomenon revisited: What role for working memory capacity? *J Exp Psychol Learn Mem Cogn*. 2004; 30:1106–1118. [PubMed: 15355139]
- Best V, Mason CR, Kidd G. Spatial release from masking in normally hearing and hearing-impaired listeners as a function of the temporal overlap of competing talkers. *J Acoust Soc Am*. 2011; 129:1616–1625. [PubMed: 21428524]
- Bopp KL, Verhaeghen P. Aging and verbal memory span: a meta-analysis. *J Gerontol B Psychol Sci Soc Sci*. 2005; 60B:P223–P233. [PubMed: 16131616]
- Broadbent DE. Failures of attention in selective listening. *J Exp Psychol*. 1952; 44:428–433. [PubMed: 13000090]
- Brungart DS. Informational and energetic masking effects in the perception of two simultaneous talkers. *J Acoust Soc Am*. 2001; 109:1101–1109. [PubMed: 11303924]
- Brungart DS, Simpson BD. Effects of target-masker contextual similarity on the multimasker penalty in a three-talker diotic listening task. *J Acoust Soc Am*. 2010; 128:2998–3010. [PubMed: 21110595]
- Carhart R, Tillman TW. Interaction of competing speech signals with hearing loss. *Arch Otolaryngol*. 1970; 91:273–279. [PubMed: 5414080]
- Duquesnoy AJ. Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons. *J Acoust Soc Am*. 1983; 74:739–743. [PubMed: 6630729]
- Ezzatian P, Li L, Pichora-Fuller K, Schneider B. The effect of priming on release from informational masking is equivalent for younger and older adults. *Ear Hear*. 2011; 32:84–96. [PubMed: 21178568]
- Heinrich A, Schneider BA, Craik FIM. Investigating the influence of continuous babble on auditory short-term memory performance. *Q J Exp Psychol*. 2008; 61:735–751.
- Helfer KS, Chevalier J, Freyman RL. Aging, spatial cues, and single- vs. dual-task performance in competing speech perception. *J Acoust Soc Am*. 2010; 128:3625–3633. [PubMed: 21218894]
- Helfer KS, Freyman RL. Aging and speech-on-speech masking. *Ear Hear*. 2008; 29:87–98. [PubMed: 18091104]
- Helfer KS, Vargo M. Speech recognition and temporal processing in middle-aged women. *J Am Acad Audiol*. 2009; 20:264–271. [PubMed: 19927698]
- Humes LE, Coughlin MP. Aided speech-identification performance in single-talker competition by older adults with impaired hearing. *Scand J Psychol*. 2009; 50:485–494. [PubMed: 19778396]
- Humes LE, Lee JH, Coughlin MP. Auditory measures of selective and divided attention in young and older adults using single-talker competition. *J Acoust Soc Am*. 2006:2926–2937. [PubMed: 17139749]
- Ihlefeld A, Shinn-Cunningham B. Spatial release from energetic and informational masking in a selective speech identification task. *J Acoust Soc Am*. 2008; 123:4369–4379. [PubMed: 18537388]
- Iyer N, Kidd G, Arbogast TL, Mason CR, Gallun FJ. The advantage of knowing where to listen. *Acoust Soc Am*. 2005; 118:3804–3815.
- Kidd G, Best V, Mason CR. Listening to every other word: Examining the strength of linkage variables in forming streams of speech. *J Acoust Soc Am*. 2008; 124:3793–3802. [PubMed: 19206805]
- Li L, Daneman M, Qi JG, et al. Does the information content of an irrelevant source differentially affect spoken word recognition in younger and older adults? *J Exp Psychol Hum Percept Perform*. 2004; 30:1077–1091. [PubMed: 15584816]
- Lustig C, May CP, Hasher L. Working memory span and the role of proactive interference. *J Exp Psychol Gen*. 2001; 130:199–207. [PubMed: 11409099]
- Mackersie CL. Talker separation and sequential stream segregation in listeners with hearing loss: Patterns associated with talker gender. *J Speech Lang Hear Res*. 2003; 46:912–918. [PubMed: 12959469]
- Marrone N, Mason CR, Kidd G. The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *J Acoust Soc Am*. 2008; 124:3064–3075. [PubMed: 19045792]

- Moore, BCJ. Cochlear Hearing Loss: Physiological, Psychological, and Technical Issues. 2nd ed.. West Sussex, England: John Wiley & Sons, Ltd; 2007.
- Murphy DR, Daneman M, Schneider BS. Why do older adults have difficulty following conversations? *Psychol Aging*. 2006; 21:49–61. [PubMed: 16594791]
- Naveh-Benjamin M, Craik FIM. Effects of perceptual and conceptual processing on memory for words and voices: different patterns for young and old. *Q J Exp Psychol*. 1996; 49A:780–796.
- Neher T, Behrens T, Carlile S, et al. Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, cognition. *Int J Audiol*. 2009; 48:758–774. [PubMed: 19951144]
- Neher T, Laugesen S, Jensen NS, et al. Can basic auditory and cognitive measures predict hearing-impaired listeners' localization and spatial speech recognition abilities? *J Acoust Soc Am*. 2011; 130:1542–1558. [PubMed: 21895093]
- Pichora-Fuller MK, Schneider BA, Daneman M. How young and old adults listen to and remember speech in noise. *J Acoust Soc Am*. 1995; 97:593–608. [PubMed: 7860836]
- Rossi-Katz J, Arehart KH. Message and talker identification in older adults: Effects of task, distinctiveness of the talkers' voices, and meaningfulness of the competing message. *J Speech Lang Hear Res*. 2009; 52:435–453. [PubMed: 19064902]
- Schneider, BA.; Pichora-Fuller, MK. Implications of perceptual deterioration for cognitive aging research. In: Craik, FIM.; Salthouse, TA., editors. *The Handbook of Aging and Cognition*. 2nd ed.. Mahwah, NJ: Lawrence Erlbaum; 2000. p. 155-219.
- Singh G, Pichora-Fuller M, Schneider BA. The effect of age on spatial attention in conditions of real and simulated spatial separation. *J Acoust Soc Am*. 2008; 124:1294–1305. [PubMed: 18681615]
- Sommers MS, Danielson SM. Inhibitory processes and spoken word recognition in younger and older adults: The interaction of lexical competition and semantic context. *Psychol Aging*. 1999; 14:458–472. [PubMed: 10509700]
- Sorqvist P. The role of working memory capacity in auditory distraction: A review. *Noise Health*. 2010; 49:217–223. [PubMed: 20871176]
- Studebaker GA. A “rationalized” arcsine transform. *J Speech Hear Res*. 1985; 28:455–462. [PubMed: 4046587]
- Tun PA, O’Kane G, Wingfield A. Distraction by competing speech in young and older listeners. *Psychol Aging*. 2002; 17:453–467. [PubMed: 12243387]
- Tun PA, Wingfield A. One voice too many: Adult age differences in language processing with different types of distracting sounds. *J Gerontol B Psychol Sci Soc Sci*. 1999; 54B:P317–P327. [PubMed: 10542824]
- Wingfield A. Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory. *J Amer Acad Audiol*. 1996; 7:175–182. [PubMed: 8780990]
- Wingfield A, Tun PA. Spoken language comprehension in older adults: interactions between sensory and cognitive change in normal aging. *Semin Hear*. 2001; 22:287–301.
- Yonan CA, Sommers MS. The effects of talker familiarity on spoken word identification in younger and older listeners. *Psychol Aging*. 2000; 15:88. [PubMed: 10755292]

Short Summary

Modulation-based noise reduction (MBNR) is used in hearing aids to reduce noise and improve listening comfort. The main goal in the present study is to investigate the effects of MBNR on noise and speech using acoustic measures. The inversion technique was used to separate speech and noise post-recording. Fricatives /s/ and /z/ were extracted from the retrieved speech file for detailed acoustic analysis. Some clinically-available implementations of MBNR have measurable effects on the acoustics of fricatives. Implications for pediatric amplification are discussed.

Figure 1.
Composite audiograms (means and standard errors) for the younger (left panel) and older (right panel) participants. The dotted lines represent the range of pure-tone thresholds.

Figure 2.

Performance on the interleaved sentence tasks. FF = target and masker presented without spatial separation, both from a front loudspeaker; FR = target and masker presented in a spatially separated fashion, with the target words coming from the front and the masker words delivered by a loudspeaker located 60 degrees to the right of the participant. Errors bars represent the standard error.

Figure 3.

Difference scores derived for the three within-subjects variables. The upper panel shows spatial benefit (spatially-separated FR trials – spatially coincident FF trials). The middle panel displays talker consistency benefit (talker consistent trials – talker variable trials). The bottom panel represents syntactic benefit (syntactically-correct word order trials – random word order trials). Error bars represent the standard error.

Figure 4.

Error responses produced by participants, averaged across all conditions. Masker errors were when participants responded with a word from the to-be-ignored stream. Non-masker errors were responses that were words not played on that trial, or omissions. Error bars represent the standard error.

Figure 5. Scatterplots depicting the associations between digit-span forward and performance in syntactically-correct trials (top panel) and random word order trials (bottom panel) for the 13 older participants.

Table I

Response grid displayed during this study.

Bob	bought	2	big	bags
Gene	found	3	blue	cards
Jane	gave	4	cold	gloves
Jill	held	5	hot	hats
Lynn	lost	6	new	pens
Mike	saw	8	old	shoes
Pat	sold	9	red	socks
Sue	took	10	small	toys

Correlation coefficients for the Pearson r analysis conducted on the older participants' data. Digfor = forward digit span; digback = backward digit span; hfpta = average of better-ear thresholds at 2 kHz, 4 kHz, 6 kHz, and 8 kHz; syn = averaged scores on syntactically-correct trials; ran = averaged scores on trials with random word order

Table II

	digfor	digback	hfpta	syn	ran
digfor	----	.67*	-.01	.70**	.67*
digback		----	.11	.13	.15
hfpta			----	.02	.27
syn				----	.95**
ran					----

** significant at the .01 level;

* significant at the .05 level.