Published in final edited form as:

Ear Hear. 2014 ; 35(2): 161–170. doi:10.1097/AUD.0b013e3182a830cf.

Competing speech perception in older and younger adults: Behavioral and eye movement evidence

Karen S. Helfer, Ph.D.¹ and Adrian Staub, Ph.D.²

¹University of Massachusetts Amherst, Department of Communication Disorders

²University of Massachusetts Amherst, Department of Psychology

I. INTRODUCTION

People often must communicate in situations where there is more than one person talking simultaneously. This can be a challenging task for all individuals, but often is especially problematic for older adults. Research into competing speech perception validates the anecdotal complaints of difficulty made by many older adults (e.g., Agus et al., 2009; Carhart and Tillman, 1970; Duquesnoy, 1983; Helfer, Chevalier, and Freyman, 2010; Helfer and Freyman, 2008; Humes, Lee, and Coughlin, 2006; Li et al., 2004; Rossi-Katz and Arehart, 2009; Singh, Pichora-Fuller, and Schneider, 2008; Tun, O'Kane, and Wingfield, 2002; Tun and Wingfield, 1999).

Successful communication in multi-talker environments requires both peripheral and cognitive resources. Competing speech can cause interference on both of these levels. Peripherally, energy in the masking speech can obscure components of the to-be-attended message, thereby producing energetic masking. When the background sound consists of speech, higher-level interference also can occur (see Mattys et al., 2012, for an extensive review of work in this area). This higher-level interference often is referred to as informational masking (e.g. Pollack et al., 1975), which encompasses problems with segregating the target message from the background messages, confusion about which is the to-be-attended stream, and distraction caused by an understandable competing message.

Age-related hearing loss undoubtedly contributes to older adults' difficulty in these situations, as peripheral hearing loss causes some speech sounds to be inaudible and brings about increased susceptibility to energetic masking (e.g., Arbogast, Mason, and Kidd, 2005; Dubno, Horwitz, and Ahlstrom, 2002). However, results of a number of research studies suggest that the problems experienced by older adults in situations in which there are competing messages cannot be entirely explained by degree of hearing loss (e.g., Helfer and

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.

The present study used an eyetracking paradigm to quantify differences between older and younger adults in the ability to attend to a sentence presented in one ear while ignoring speech or noise presented to the other ear. Participants were instructed to click on a word presented on a computer screen that was in the to-be-attended sentence. A second word displayed on the screen either was or was not from the to-be-ignored sentence. All three metrics of performance (percent correct, response time, and eye movement patterns) indicated that older adults experienced greater difficulty in competing speech conditions than did younger adults. These results support the idea that older adults have problems in competing speech situations for reasons beyond lack of audibility or susceptibility to energetic masking.

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.

Freyman, 2008; Humes, Lee, and Coughlin, 2006; Humes and Coughlin, 2009; Rossi-Katz and Arehart, 2009; Russo and Pichora-Fuller, 2008; Tun, O'Kane, and Wingfield, 2002). This is perhaps not surprising, as age-related changes in cognitive abilities (e.g., working memory, attention) may be expected to play a larger role when individuals must ignore competing speech, as compared to when the background competition consists of noise that does not contain linguistic content. In fact, research has shown that while amount of hearing loss drives performance in simple listening situations (such as with relatively steady background noise), cognitive abilities come into play in more complex listening situations, such as those that involve multiple talkers (e.g., Lunner, Rudner, and Ronnberg, 2009; Neher et al., 2009).

We are particularly interested in how attention mediates older adults' ability cope in competing speech situations. Selective attention comes into play when one must attend to a source (such as one talker) while inhibiting others. Strong links have been proposed between working memory (which includes functions that store, manipulate, and control the processing of information) and attention. High working memory capacity has been shown to be associated with better selective attention ability (e.g., Kane and Engle, 2003; Poole and Kane, 2009), suggesting that people with better working memory have stronger attentional control mechanisms. Results of a recent study (Ahmed and de Fockert, 2012) demonstrate that depletion of working memory resources leads to problems with maintaining attention on relevant (vs. irrelevant) information. Moreover, working memory (which appears to be important for speech understanding in noise -- see Akeroyd, 2008, for a summary of work in this area) is especially vulnerable to the negative effects of aging (e.g., Bopp and Verhaeghen, 2005; Verhaeghen and Salthouse, 1997).

Results of investigations of selective visual attention support the idea that older adults have more difficulty than younger individuals in tasks incorporating visual distraction (see Guerreiro, Murphy, and Van Gerven, 2010, for a thorough review of this body of work). Previous research on age-related changes in the ability to ignore auditory distractors has led to mixed conclusions. One challenge in conducting this type of research is that the majority of older individuals have some degree of hearing loss, which, in and of itself, leads to reduced performance on many tasks used to measure auditory attention. Some studies suggest no evidence of age-related problems inhibiting competing speech once peripheral hearing loss is controlled (e.g., Li et. al, 2004; Murphy, Daneman, and Schneider, 2006; Schneider et al., 2000). However, as mentioned above, other work suggests that age-related differences in competing speech perception cannot be explained solely by degree of threshold elevation or by difficulty in determining which talker should be attended to (e.g., Helfer and Freyman, 2008; Helfer, Mason, and Marino, 2013; Helfer and Vargo, 2009; Sommers and Danielson, 1999; Tun and Wingfield, 1999; Tun et al., 2002). This suggests that age-related changes in selectively attending to one message in the presence of other messages may play a role in older adults' difficulty in complex listening situations.

Our goal for this study was to identify age-related differences in attentional patterns in a competing speech task. One challenge in designing this experiment was determining whether any age group differences eventually found were related to hearing loss in our older subjects or to true attentional influences. That is, if an older adult has difficulty understanding speech in the presence of competing speech, to what extent is this due to inability to hear or distinguish speech sounds (from decreased audibility and/or increased susceptibility to energetic masking) vs. changes in the ability to allocate attention? The present study attempted to control for the effects of hearing loss by presenting the to-be-attended and to-be-ignored speech in different ears, which should essentially eliminate energetic masking. We used an eye tracking technique to examine the time course of allocation of attention to both to-be-attended and to-be-ignored speech streams. Our working

hypothesis was that age-related differences in the processing of competing speech would be revealed as differences in eye movement patterns between younger and older individuals. Coupled with the dichotic listening paradigm used here, results of this study would provide additional evidence that older adults have particular difficulty in competing speech situations, above and beyond that caused by inaudibility issues from hearing loss.

Eye tracking techniques have been used for many years to study on-line processing of visual information, such as reading (see, e.g., Rayner, 2009; Staub and Rayner, 2007 for reviews). More recently researchers have used this tool to examine the processing of spoken language. One of the first studies to do so was conducted by Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy (1995; cf. Cooper, 1974) who monitored the eye movements of listeners while they responded to spoken instructions to move or touch items on a display (e.g., “Touch the starred yellow square”). They found that individuals’ eye movements to the displayed objects were closely time-locked to the onset of the spoken words that potentially referred to those objects. Since then, eye tracking research has provided insight into the detailed time-course of language comprehension by exploring issues such as the activation of alternative syntactic parses (e.g., Tanenhaus et al., 1995), the activation of phonological competitors during word recognition (e.g., Allopenna, Magnuson, and Tanenhaus, 1998), and the activation of words that are semantically related to a spoken word (e.g., Yee and Sedivy, 2006).

Traditionally, eye tracking studies use a “visual world” paradigm where the participant looks at a visual display containing pictures while listening to speech. In the typical experiment, the display includes a picture of an object that is named in the spoken stimulus, one or more objects that share some property (e.g., initial sounds) with the target, and one or more irrelevant distractor objects. More recently, experimenters have used printed words in the visual display instead of pictures (e.g., Huettig and McQueen, 2007; McQueen and Viebahn, 2007). This modification allows for the use of stimuli that are not easily portrayed by pictures and eliminates potential differences among subjects in the labeling of possibly ambiguous images.

Very few studies have used eye tracking to assess the processing of spoken stimuli in older adults. Ben-David et al. (2011) employed a visual world paradigm to examine the processing of speech in younger and older individuals. In their study, stimuli were presented with or without background noise, which was used to equate the performance of older and younger adults by manipulating the signal-to-noise ratio. Participants viewed pictures of the target words and competitors that either overlapped in onset (e.g., “tower” vs. “towel”) or rhymed (e.g., “sandal” vs. “candle”). Results of this study uncovered age-related differences in eye movements when target words were presented with rhyming competitors, particularly in conditions with background noise. Age effects were not noted when target and distractor words had the same onset sounds. Revill and Spieler (2012) used an eyetracking paradigm to examine the effect of lexical frequency on word recognition by older and younger individuals. Their results suggest that while lexical frequency does appear to play a greater role for older individuals, there was little evidence of increased competition from failure to inhibit on the part of older participants.

In the present study, participants were instructed to use a computer mouse to click on a visually-presented word that was heard in a to-be-attended speech stream, played to one ear. A foil word was also presented on the screen. Within subjects, we manipulated whether the stimulus heard in the contralateral ear was silence, steady-state noise, or a to-be-ignored speech stream. When to-be-ignored speech was played in the contralateral ear, we also manipulated whether the foil word presented on the computer screen was a word that occurred in this second speech stream, or a random word heard in neither stream. This

manipulation was included in order to determine whether older adults would be at an additional disadvantage when they also needed to inhibit attention to the visual form of a word they were instructed to ignore. Comparing differences between groups when the foil was vs. was not heard would give insight into the nature of inhibitory problems in terms of modality. Finding group differences in eye movement patterns in the presence of competing speech would suggest age-related changes in auditory attention. Uncovering more significant effects when the foil was the to-be-ignored word could be interpreted as problems inhibiting visually-presented stimuli. Based on previous work on age-related changes in visual attention, we expected that the presence of a foil containing a to-be-ignored word would have a larger negative effect on our older (vs. younger) participants.

II. MATERIALS AND METHODS

A. Participants

This experiment used two groups of participants, with 23 individuals per group: college-aged adults with normal hearing (mean age 21.4 years, range 20-25 years) and older adults (mean age 68.6 years, range 60-81 years) with no more than a mild-to-moderate hearing loss. Audiometric thresholds were obtained from each participant for octave frequencies between 250 Hz and 8000 Hz. A composite audiogram for the older subjects is shown in Figure 1. The mean high-frequency pure-tone average (average for thresholds at 2 kHz, 3 kHz, 4 kHz, and 6 kHz) for the older listeners was 32 dB HL in the both the right and left ear. In order to rule out conductive hearing loss, each participant also had normal (Type A) tympanograms in each ear on the day of testing. All participants were native-English speakers and with no history of significant otologic disease or neurologic disorder. Each participant had corrected-to-normal vision by self-report.

B. Stimuli

The auditory stimuli for the studies described in this paper were pairs of sentences from the TVM (Theo, Victor, Michael) corpus (Helfer and Freyman, 2009). These sentences take the form “*Name* discussed the ____ and the ____ today” where *Name* is one of three cues names (*Theo*, *Victor*, or *Michael*) and underlines correspond to key words used for scoring. Sentences were recorded from three male talkers (see Helfer and Freyman, 2009). For the current studies, sentences were selected in which the key words consisted of one-syllable nouns. Pairs of TVM sentences were created in which none of the four key words within the pair started with the same phoneme or was semantically related. Stimuli were edited so that the first key word in each of the two sentences of the pair began at the same time by extending the length of naturally-occurring silent intervals within the sentence. The same was done for the second key word within each sentence pair. Each pair contained one sentence beginning with *Theo* (which was designated as the to-be-attended sentence) and another starting with *Victor* or *Michael*. Each sentence within a pair was spoken by a different talker, with the target talker randomized from trial to trial.

Three types of listening conditions were used: pairs of TVM sentences (as described above) with one sentence presented to each ear; a single sentence presented in one ear; and one sentence presented in one ear with steady-state noise presented in the other. Equal-level stimuli were presented in each ear at peak SPL of approximately 80 dBA. Dichotic presentation was used for two related reasons. First, the eye tracking paradigm yields most useful information when performance levels are high. Because dichotic presentation is likely to essentially eliminate energetic masking (with the exception of sound energy that might cross over via bone conduction), it increases the chance of obtaining high levels of performance even in older listeners, for whom simultaneous presentation of competing sentences in the same ear would likely be negatively impacted by peripheral hearing loss. In

a similar vein, dichotic presentation also would allow us to isolate the effects of informational masking without the confounding factor of increased susceptibility to energetic masking that occurs with age-related hearing loss. Other studies have used dichotic presentation of competing speech as a way to examine the effects of informational masking (Brungart and Simpson, 2002; Wightman, Kistler, and O'Bryan, 2010).

C. Procedures

On each trial, two words were displayed on a computer screen directly in front of the participant, above and below the center of the screen. One of these was always one of the two key words heard in the target sentence. The to-be-attended ear was randomized across items, as was the position of the target word on the screen. During single-sentence presentations (in quiet and with contralateral noise), the foil word was a random word that was not phonetically or semantically related to either of the key words in the to-be-attended sentence. Two different visual display conditions were used during competing speech trials. On half of these trials, the visually-presented foil word was the key word in the to-be-ignored sentence that played synchronously with the target word; e.g., if the first key word from the to-be-attended sentence was presented on the screen, the displayed word from the to-be-ignored sentence was also the first key word in that sentence. On the other half of trials, the foil word was a random word that was not presented in the target/masking complex. None of the displayed words were used more than once within the experiment. The words were presented on a 19-inch monitor positioned 70 cm from the participant. Screen resolution was 1024×768 pixels, and the words were presented centered 184 pixels above and below the center of the screen, in 36-point Monaco font, in black against a light gray background.

Eye movements were recorded using an EyeLink 1000 (SR Research, Ltd.) system. This system uses a remote desktop camera to monitor eye movements; the sampling rate was 500 Hz. Before testing was initiated a calibration was completed in which the participant visually followed a series of 13 points displayed on the computer screen. This was used to calculate an average gaze error, which was between 0.5° and 1.0° of visual angle. Calibration was repeated if necessary during the experiment.

The sequence of events on each trial was as follows. First, the two words were displayed above and below a red dot in the center of the screen. After a 1000 ms delay, the dot turned green, signaling to participants to click on the dot to begin the trial. The subject's click on the dot was followed by another 250 ms delay, at which point the auditory stimulus was played. Participants were instructed to click on the word displayed on the screen that they heard in the sentence beginning with "Theo". In the quiet and noise conditions, this was the only sentence the subject heard, while in the dichotic condition, one sentence began with "Theo", and the other with "Michael" or "Victor".

Participants completed a total of 90 trials, comprising three blocks each consisting of two practice trials and 28 experimental trials. In the initial block a single sentence was presented in quiet, with the order of the other two conditions (single sentence in noise and dichotic listening) counterbalanced across subjects. The dichotic listening trials were evenly divided between trials on which the visually-presented foil word did occur in the to-be-ignored stream, and trials on which it did not, with these two trial types randomly intermixed.

III. RESULTS

A. Percent-correct performance

Data from one younger subject were eliminated as her overall accuracy (82.9%) was more than three standard deviation (sds) below the mean (96.3%; $sd = .039\%$). Hence, results

from 22 younger individuals and 23 older individuals will be presented. We note that in this and all subsequent analyses, we collapse the trials in which the critical word was presented first and second in the to-be-attended stream, as this variable did not result in significant effects, and did not interact with age group.

Figure 2 shows performance in terms of accuracy of identification of the key words for each group, in each condition. Accuracy was at ceiling for both groups of participants when single sentences were presented in quiet or in the presence of steady-state noise in the opposite ear. However, performance declined for both groups when a competing sentence was presented in the opposite ear. Moreover, the size of this decrement was influenced by whether the non-target word displayed on the screen was a random word or came from the masking sentence. For the younger group, mean accuracy was 98.7% when the word displayed on the screen was a random word and 93.0% when the displayed word came from the masking sentence. For the older listeners, mean accuracy was 91.6% in the condition with a random displayed word and 78.1% when the second word displayed was from the to-be-ignored sentence.

Because of the high accuracy levels, the percent correct scores were transformed into Rationalized Arcsine Units (RAU; Studebaker, 1985) prior to analysis. ANOVAs were completed on RAU scores from the two competing speech conditions with display condition (displayed word random or from the masker) as a within-subjects factor and subject group as a between-subjects factor. Results showed significant main effects of display condition ($F(1, 43) = 15.52, p < .001$) and subject group ($F(1, 43) = 32.88, p < .001$). The interaction between these two factors was not significant ($F(1, 43) = 1.03, p = .32$). We note that though there appears to be an interaction when the data are considered as untransformed proportions, the appropriate transformation of these proportional data increases the relative size of condition differences when performance is nearer to ceiling (i.e., for the younger subjects), eliminating the appearance of an interaction.

B. Response time

Response time (RT) was defined as the time in milliseconds from the onset of the target word in the to-be-attended stimulus until a word was clicked on the screen. We restrict RT analyses to trials on which the correct word was ultimately clicked. Mean response time for each group is shown in Figure 3. Several trends are of note. First, RT was generally slower in the competing speech conditions. Second, this trend was more pronounced for older than younger subjects; older subjects were no slower than younger subjects in the quiet and noise condition, but were clearly slower in the competing speech conditions. To test these effects, we first computed an ANOVA on the RT data with condition as a two-level factor (quiet and noise vs. speech), and with age group as a between-subjects factor (as the quiet and noise conditions did not differ statistically on any measure, we aggregated them in this and subsequent analyses). This analysis yielded no main effect of group ($F(1, 43) = .90, p = .35$), but a significant effect of condition ($F(1, 43) = 39.43, p < .001$), and a significant interaction ($F(1, 43) = 8.85, p < .01$). We also assessed whether there was a difference in RT patterns for the two competing speech conditions. ANOVA on the RT data for the competing speech conditions (with trial type as a within-subjects factor and group as the between-subjects factor) showed a nonsignificant main effect of age group ($F(1, 43) = 2.83, p = .099$), a significant effect of condition ($F(1, 43) = 8.20, p < .01$), and a nonsignificant interaction ($F(1, 43) = 2.06, p = .16$). Note that the effect of trial type was in an unexpected direction, as RTs were actually faster when the foil word displayed on the screen was heard in the to-be-ignored speech stream.

Summing up the behavioral data, it appears that older subjects were no slower and were no less accurate than younger subjects when a sentence was presented in isolation to one ear or

when noise was heard in the contralateral ear. Both groups were slower and less accurate when speech was presented in the contralateral ear. However, in terms of both accuracy and RT, older subjects were more affected than younger subjects by competing speech. In addition, subjects' responses were less accurate when the foil word presented on the screen was heard in the competing speech stream. This effect did not significantly interact with age, and was not evident in RTs; indeed, RTs were faster when the foil word presented on the screen was heard in the competing speech stream.

C. Eye movement patterns

Because the response words were presented in large type and without lateral masking, it was likely that some eye fixations for a processed word would not fall directly on that word. Visual inspection of the spatial distribution of all eye fixations in the experiment revealed a concentration of fixations around each word location that was contained in a rectangle 100 pixels in height and 168 pixels in width, centered at the word's centroid. Thus, any fixation within this rectangle was counted as a fixation on the word. As for RTs, analysis of eye movement data was restricted to correct trials. Figure 4 illustrates the basic eye movement patterns. Consistent with the behavioral data, it is clear that there is little difference between younger and older subjects' eye movements in the quiet and noise conditions. However, it appears that in the competing speech conditions, older subjects generally took longer to fixate the target word than did younger subjects. There is also a hint that when the distractor word was displayed, older subjects showed more persistent fixations on this word than did younger subjects.

We assess these patterns statistically by examining three eye movement measures that reflect the degree to which subjects' visual attention was allocated to the foil word. First, we report the probability of initiating an eye fixation on the foil word at any time between the critical point in the to-be-attended speech stream (i.e., the onset of the target word) and the mouse click response. Second, we report the total time spent fixating the foil word, starting from the critical point in the to-be-attended stream and ending with the mouse click response. If the subject did not fixate the foil during this period, this measure would receive a value of 0 for that trial. If the subject fixated the foil once, this value would be equal to the duration of that single eye fixation. If the subject fixated the foil multiple times, the durations of these fixations would be summed. Finally, we measured the latency between the critical point in the to-be-attended stream and the beginning of the participant's first eye fixation on the target word.

For each measure, we first computed an ANOVA comparing the non-speech and speech conditions (for all these measures, as for the behavioral measures above, the quiet and noise conditions were very similar), and then computed an ANOVA comparing the two speech conditions.

Figure 5 shows that the probability of fixating the foil was greater in the speech conditions, and as was the case with accuracy and RT, this effect was numerically larger for older subjects. The proportion measures were converted to rationalized arcsin units for statistical analysis. The ANOVA comparing the speech to non-speech conditions revealed only a significant effect of speech vs. non-speech ($F(1, 43) = 21.00, p < .001$), while the effect of age group did not reach significance ($F(1, 43) = 1.25, p = .27$), nor did the interaction effect ($F(1, 43) = .55, p = .46$). The ANOVA for the speech conditions revealed no significant effect of age group ($F(1, 43) = 1.77, p = .19$), with no significant effect of condition ($F(1, 43) = 2.84, p = .10$), and no significant interaction ($F(1, 43) = 1.27, p = .27$).

The total fixation time spent on the foil is shown in Figure 6. The ANOVA comparing the speech to non-speech conditions resulted in a significant effect of condition ($F(1, 43) =$

28.32, $p < .001$), but no significant effect of age group ($F(1, 43) = 1.84$, $p = .18$) and no significant interaction ($F(1, 43) = 1.33$, $p = .25$). The ANOVA comparing the two speech conditions also revealed a significant effect of condition ($F(1, 43) = 8.58$, $p < .01$), no significant age effect ($F(1, 43) = 2.63$, $p = .11$), and no significant interaction ($F(1, 43) = .87$, $p = .35$).

Figure 7 shows the latency to initiate a fixation on the target word, starting at the onset of the critical word in the to-be-attended stream. This analysis was restricted to trials on which an eye fixation on the target word was in fact recorded after the critical point in the auditory stimulus, which was 84% of all correct trials. The ANOVA comparing speech to non-speech conditions revealed a significant effect of condition ($F(1, 43) = 17.12$, $p < .001$). The effect of age group was not significant ($F(1, 43) = .20$, $p = .65$), but the interaction was significant ($F(1, 43) = 6.37$, $p = .015$). Comparing the speech conditions only, the ANOVA indicated no significant effects (all p values $> .15$).

In sum, the presence of competing speech resulted in more looks to the foil word, more time spent fixating the foil word, and a longer delay before fixating the target word. Figures 5-7 all indicate that numerically, eye movements of older subjects were more affected by the presence of competing speech than were eye movements of younger subjects. However, this numerical interaction reached statistical significance only in the latency-to-target measure. This is consistent with Figure 4, where it is clear that in the competing speech conditions, but not in the quiet and noise conditions, older subjects' fixations on the target increased more slowly than did those of younger subjects. The only evidence that eye movements were influenced by whether the foil word was present in the competing speech stream was that subjects spent more total time fixating the foil word when it was heard in the competing stream.

Together, the results from the behavioral measures and the eye movement measures may be summarized as follows. The presence of competing speech resulted in less accurate behavioral responses. Accurate responses were also slower when there was a competing speech stream, and subjects were more likely to look at the foil word, looked longer at the foil word, and took longer to fixate the target word. When the visually-presented foil was presented as a distractor in the to-be-ignored speech stream, accuracy was further diminished, though RT was faster for correct trials (we discuss this tradeoff below). Subjects also spent more time fixating the foil in this condition. Older subjects generally behaved very much like younger subjects in the quiet and noise conditions but differed from younger subjects in the competing speech conditions, where their accuracy was lower and their RTs for correct trials were slower. The clearest effect of competing speech on the eye movements of older subjects was to slow their fixations to the target even more than for younger subjects.

E. Correlations with hearing loss

We were interested in identifying the degree to which hearing loss in our older participants could account for the results. Pearson r correlations were run with the following variables: bilateral high-frequency average (average of thresholds for 2kHz, 4kHz, and 6kHz pure tones), percent-correct scores in the two competing speech conditions, reaction time in the two speech condition, and proportion of trials with looks to the non-target word in each of the two speech conditions. Degree of high-frequency hearing loss was not significantly associated with any of the measured variables, with r values ranging from .008 ($p = .973$) to .386 ($p = .076$).

IV. DISCUSSION

Results of the present study support previous work demonstrating that older adults are more greatly affected by competing speech than are younger listeners. Each of the metrics used to measure performance in the present study (percent-correct performance, response time, and eye movement patterns) demonstrated that older adults are more susceptible to the negative effects of speech presented in a to-be-ignored ear. In the present study, age-related differences in performance cannot be attributed to increased susceptibility to energetic masking caused by a speech masker, as dichotic presentation likely eliminated energetic masking (this contention is supported by the fact that there was no evidence of masking, for either younger or older subjects, when steady-state noise was presented in the contralateral ear). Further, degree of high-frequency hearing loss was not associated with any of the performance metrics. This pattern of results suggests that the problems experienced by older adults in complex listening situations are at least in part due to difficulty ignoring competing speech.

Our results confirm a body of previous work demonstrating reduced dichotic performance by older adults (relative to younger listeners) (e.g., Bouma and Gootjes, 2011; Clark and Knowles, 1973; Hallgren et al., 2001; Hugdahl, Carlsson, and Eichele, 2001; Jerger et al., 1994; Strouse et al., 2000). Results of previous studies using dichotic presentation of target and masking speech suggest that performance of young adults should be at or near ceiling (Brungart and Simpson, 2002; Kidd, Mason, and Gallun, 2005). This was true in the present study when the word presented visually was random, but even younger adults showed a modest decline in performance when they saw a word corresponding to what was heard in the to-be-ignored stream. Recall that what the subject heard was identical in this condition and when the word presented on the screen was random—the difference was in what the subjects saw.

These results suggest that the presence of the visual form of a word that is to be ignored makes it more difficult to disregard that word when it is presented to the ears. Interestingly, hearing the visually-presented foil did not have much effect on eye movement patterns, and it actually decreased response time; both younger and older subjects responded faster in this condition than when the visually-displayed foil was random. In sum, a speed-accuracy tradeoff was in evidence, in which auditory presentation of the visually-presented foil made subjects less accurate, but faster. This pattern is clearly in need of further investigation, as it suggests that hearing the visually presented foil may cause subjects to adopt a more liberal criterion for responding.

We had predicted that older adults would have been more affected by seeing the distractor word, but there was no significant interaction between group and whether or not the foil word was heard in the to-be-ignored stream. Previous research on attending to auditory information in the presence of a visual distractor suggests that older adults may be less able to inhibit a visual distractor presented during auditory tasks (see Guerreiro, Murphy, and Van Gerven, 2010). Results of the present study do not support age-related problems inhibiting a visually-presented distractor, at least not at a level that reached statistical significance. However, it should be recognized that the relatively small number of subjects in the present study might have led to an inability to find significant results.

Analysis of the eye movement patterns provided additional evidence that older adults were more affected than younger listeners by the presence of competing speech, as older subjects were especially slowed in fixating the target word in this condition. We offer two possible explanations to account for this finding. First, the results could be interpreted as showing a problem at the level of selection. A factor that could have influenced the results was the

uncertainty involved in determining which ear was to be attended at the beginning of each trial. Since the to-be-attended ear was not pre-cued, subjects had to wait to hear the cue word to determine which ear contained the to-be-attended stream. If older adults were slower or less efficient at this selection process, this could account (at least partially) for differences found between groups. This idea is bolstered by results of a recent study (Passow et al., 2011) suggesting that older adults are less flexible in allocating attention between the ears during dichotic listening. If that is the case, then the lack of cueing the to-be-attended ear would be expected to have a greater negative effect on our older (vs. younger) listeners. However, giving participants a cue before stimulus presentation might not have substantially affected our results, as one previous study found that older and younger listeners obtain equivalent benefit in competing speech tasks from a cue that signals the to-be-attended location (Singh, Pichora-Fuller, and Schneider, 2008). Second, it is possible that even if older adults are able to rapidly determine to which ear the to-be-attended stream is being presented, they may be less effective than younger adults in following the target message over time and/or maintaining attention exclusively to this stream (e.g., Alain, Ogawa, and Woods, 1996; Mackersie, Prida, and Stiles, 2001). Follow-up work should be conducted to differentiate between these hypotheses.

One other possible contributor to the present results is the fact that we presented a visual form of auditorily-presented words—that is, the subjects were primed. Previous research (Freyman, Balakrishnan, and Helfer, 2004; Jones and Freyman, 2012) has established that younger adults benefit from a visually-presented preview of target speech, particularly in the presence of competing speech. Although there is evidence to suggest little difference between older and younger adults on the beneficial effects of providing a prime of a to-be-attended speech signal (Ezzatian et al., 2011), research has yet to determine the extent to which priming a to-be-ignored speech message (as was the case in the present study when the foil was from the to-be-ignored stream) might affect performance.

How do the results of the present study apply to real-life competing speech situations, in which the to-be-attended and to-be-ignored sounds are not separated by ear? Humes et al. (2006) found strong correlations between performance on monaural and dichotic presentation conditions in a competing speech task, suggesting that they tap into at least some of the same processes. This is not surprising, as it is assumed that information from an unattended ear is, at some level, combined with information from the to-be-attended stream (e.g., Brungart and Simpson, 2007; Triesman, 1964a). Competing speech in actual listening situations, however, also produces energetic masking, which was essentially eliminated in the dichotic paradigm used in this study. Moreover, there are substantial differences in the processing of dichotic stimuli and that of stimuli that are spatially separated in real rooms. Regardless of the face validity of the present task, the results suggest that older listeners experience problems in competing speech situations not only because they have difficulty understanding the target message, but also because they may be less able to inhibit attention to background conversations—even in the extreme example here of to-be-ignored and to-be-attended messages that are presented in opposite ears.

Acknowledgments

The authors would like to thank Lori Astheimer, Matt Abbott, Caitlin Cotter, and Daniel Petty for their work on this project. Portions of these data were presented at the American Auditory Society Conference, Scottsdale, AZ, March 2011. This work was supported by NIDCD R01 DC01623.

References

- Agus TR, Akeroyd MA, Gatehouse S, Warden D. 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]
- Ahmed L, de Fockert JW. Focusing on attention: The effects of working memory capacity and load on selective attention. *PLoS ONE*. 2012; 7(8):e43101. [PubMed: 22952636]
- Akeroyd MA. Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int J Audiol*. 2008; 47(Suppl.2):S53–S71. [PubMed: 19012113]
- Alain C, Ogawa KH, Woods DL. Aging and the segregation of auditory stimulus sequences. *J Geron: Psych Sci*. 1996; 51B:P91–P93.
- Allopenna P, Magnuson JS, Tanenhaus MK. Tracking the time course of spoken word recognition using eye-movements: evidence for continuous mapping models. *J Mem Lang*. 1998; 38:419–439.
- Arbogast TL, Mason CR, Kidd G. The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners. *J Acoust Soc Am*. 2005; 117:2169–2180. [PubMed: 15898658]
- Ben-David BM, Chambers CG, Daneman M, Pichora-Fuller MK, Reingold EM, Schneider BA. Effects of aging and noise on real-time spoken word recognition: evidence from eye movements. *J Speech Lang Hear Res*. 2011; 54:243–262. 2011. [PubMed: 20689026]
- Bopp KL, Verhaeghen P. Aging and verbal memory span: A meta-analysis. *J Ger: Psychol Sci Soc Sci*. 2005; 60:223–233.
- Bouma A, Gootjes L. Effects of attention on dichotic listening in elderly patients with dementia of the Alzheimers type. *Brain Cog*. 2011; 76:286–293.
- Brungart DS, Simpson BD. Within-ear and across-ear interference in a cocktail-party listening task. *J Acoust Soc Am*. 2002; 112:2985–2995. [PubMed: 12509020]
- Brungart DS, Simpson BD. Effect of target-masker similarity on across-ear interference in a dichotic cocktail-party listening task. *J Acoust Soc Am*. 2007; 122:1724. [PubMed: 17927432]
- Carhart R, Tillman TW. Interaction of competing speech signals with hearing losses. *Arch Otolaryngol*. 1970; 91:273–279. [PubMed: 5414080]
- Clark L, Knowles J. Age differences in dichotic listening. *Q J Exp Psychol*. 1973; 17:227–240.
- Cooper RM. The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cog Psychol*. 1974; 6:84–107.
- Dubno JR, Horwitz AR, Ahlstrom JB. Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing levels. *J Acoust Soc Am*. 2002; 111:2897–2907. [PubMed: 12083223]
- 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, Schneider BA, Pichora-Fuller MK, Li L. The effect of priming on Release from informational masking is equivalent for younger and older adults. *Ear Hear*. 2011; 32:84–96. [PubMed: 21178568]
- Freyman RL, Balakrishnan U, Helfer KS. Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J Acoust Soc Am*. 2004; 115:2246–2256. [PubMed: 15139635]
- Guerreiro MJS, Murphy DR, Van Gerven PWM. The role of sensory modality in age-related distraction: A critical review and a renewed view. *Psychol Bull*. 2010; 136:975–1022. [PubMed: 21038938]
- Hallgren M, Larsby B, Lyxell B, Arlinger A. Cognitive effects in dichotic speech testing in elderly persons. *Ear Hear*. 2001; 22:120–129. [PubMed: 11324841]
- 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, Freyman RL. Lexical and indexical cues in masking by competing speech. *J Acoust Soc Am*. 2009; 125:447–456. [PubMed: 19173430]
- Helfer KS, Mason C, Marino C. Aging and the perception of temporally-interleaved words. *Ear Hear*. 2013; 34:160–167. [PubMed: 23135617]
- Helfer KS, Vargo M. Speech recognition and temporal processing in middle-aged women. *J Am Acad Audiol*. 2009; 20:264–271. [PubMed: 19927698]
- Huetting F, McQueen JM. The tug of war between phonological, semantic and shape information in language-mediated visual search. *J Mem Lang*. 2011; 57:460–482.
- Hugdahl K, Carlsson G, Eichele T. Age effects on dichotic listening to consonant-vowel syllables: Interactions with attention. *Dev Neuropsychol*. 2001; 20:445–457. [PubMed: 11827098]
- 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; 120:2926–2937. [PubMed: 17139749]
- Jerger J, Chmiel R, Allen J, Wilson A. Effects of age and gender on dichotic sentence identification. *Ear Hear*. 1994; 15:274–286. [PubMed: 7958527]
- Jones JA, Freyman RL. Effect of priming on energetic and informational masking in a same-different task. *Ear Hear*. 2012; 33:124–133. [PubMed: 21841488]
- Kane MJ, Engle RW. Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *J Exp Psychol: Gen*. 2003; 132:47–70. [PubMed: 12656297]
- Kidd G Jr, Mason CR, Gallun FJ. Combining energetic and informational masking for speech identification. *J Acoust Soc Am*. 2005; 118:982–992. [PubMed: 16158654]
- Li L, Daneman M, Qi JG, Schneider BA. Does the information content of an irrelevant sound source differentially affect spoken word recognition in younger and older adults? *J Exp Psychol Hum Percept Perform*. 2004; 30:1077–1091. [PubMed: 15584816]
- Lunner T, Rudner M, Ronnberg J. Cognition and hearing aids. *Scand J Psychol*. 2009; 50:395–403. [PubMed: 19778387]
- Mackersie CL, Prida TL, Stiles D. The role of sequential stream segregation and frequency selectivity in the perception of simultaneous sentences by listeners with sensorineural hearing loss. *J Speech Lang Hear Res*. 2001; 44:19–28. [PubMed: 11218102]
- Mattys SL, Davis MH, Bradlow AR, Scott SK. Speech recognition in adverse conditions : A review. *Lang Cog Proc*. 2012; 27:953–978.
- McQueen JM, Viebahn MC. Tracking recognition of spoken words by tracking looks to printed words. *Q J Exp Psychol*. 2007; 60:661–671.
- Murphy DR, Daneman M, Schneider BA. Why do older adults have difficulty following conversations? *Psychol Aging*. 2006; 21:49–61. [PubMed: 16594791]
- Neher T, Behrens T, Carlile S, Jin C, Kragelund L, Petersen AS, vanSchaik A. Benefit from spatial separation of multiple talkers in bilateral hearing-aid users: Effects of hearing loss, age, and cognition. *Int J Audiol*. 2009; 48:758–774. [PubMed: 19951144]
- Passow S, Westerhausen R, Wartenburger I, Hugdahl K, Heekeren HR, Lindenberger U, Li S-C. Human aging compromises attentional control of auditory perception. *Psychol Aging*. 2012; 27:99–105. [PubMed: 21988156]
- Pollack I. Auditory informational masking. *J Acoust Soc Am*. 1975; 57:S5.
- Poole BJ, Kane MJ. Working-memory capacity predicts the executive control of visual search among distractors: The influences of sustained and selective attention. *Q J Exp Psychol*. 2009; 62:1430–1454.
- Rayner K. Eye movements and attention in reading, scene perception, and visual search. *Q J Exp Psychol*. 2009; 62:1457–1506.
- Revill KP, Spieler DH. The effect of lexical frequency on spoken word recognition in young and older listeners. *Psychol Aging*. 2012; 27:80–87. [PubMed: 21707175]

- Rossi-Katz J, Arehart KH. Message and talker identification in older adults: Effects of task, distinctiveness of the talker's voices, and meaningfulness of the competing message. *J Speech Lang Hear Res.* 2009; 52:435–453. [PubMed: 19064902]
- Russo FA, Pichora-Fuller MK. Tune in or tune out: Age-related differences in listening to speech in music. *Ear Hear.* 2008; 29:746–760. [PubMed: 18596643]
- Schneider BA, Daneman M, Murphy DR, Kwong See S. Listening to discourse in distracting settings: The effects of aging. *Psychol Aging.* 2000; 15:110–125. [PubMed: 10755294]
- Singh G, Pichora-Fuller MK, Schneider BA. The effect of age on auditory 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 young and older adults: The interaction of lexical competition and semantic context. *Psychol Aging.* 1999; 14:458–472. [PubMed: 10509700]
- Staub, A.; Rayner, K. Eye movements and on-line comprehension processes.. In: Gaskell, G., editor. *The Oxford handbook of psycholinguistics.* Oxford University Press; Oxford, UK: 2007. p. 327-342.
- Strouse A, Wilson RH, Brush N. Effect of order bias on the recognition of dichotic digits in young and elderly listeners. *Audiol.* 2000; 39:93–101.
- Studebaker GA. A “rationalized” arcsine transform. *J Speech Hear Res.* 1985; 28:455–462. [PubMed: 4046587]
- Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. Integration of visual and linguistic information in spoken language comprehension. *Science.* 1995; 268:1632–1634. [PubMed: 7777863]
- Treisman AM. The effect of irrelevant material on the efficiency of selective listening. *Am J Psychol.* 1964; 77:533–546. [PubMed: 14251963]
- Tun PA, O’Kane G, Wingfield A. Distraction by competing speech in young and older listeners. *Psychol Aging.* 2002; 17:453–457. [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; 54:P317–P327. [PubMed: 10542824]
- Verhaeghen P, Salthouse TA. Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychol Bull.* 1997; 122:231–249. [PubMed: 9354147]
- Wightman FL, Kistler DJ, O’Bryan A. Individual differences and age effects in a dichotic informational masking paradigm. *J Acoust Soc Am.* 2010; 128:270–279. [PubMed: 20649222]
- Yee E, Sedivy JC. Eye movements to pictures reveal transient semantic activation during spoken word recognition. *J Exp Psychol Learn Mem Cogn J Exp Psychol Learn Mem Cogn.* 2006; 32:1–14.

Figure 1. Composite audiogram for the older subjects in the present study. The dotted lines represent the range of pure-tone thresholds.

Figure 2. Mean word identification accuracy for each condition. Error bars represent the standard error.

Figure. 3.
Mean response time (in msec) for correct trials. Error bars represent the standard error.

Figure 4. Fixation proportion over time to target and foil words, by age group. Individual panels display data in the four experimental conditions.

Figure 5. Probability of fixating on the foil word at any time between onset of the to-be-attended word and the subject's response. Error bars represent the standard error.

Figure 6.
Total time fixating on the foil (in msec). Error bars represent the standard error.

Figure 7. Latency between the onset of the to-be-attended word and the first fixation on that word (in msec). Error bars represent the standard error.