Lexical influences on competing speech perception in younger, middle-aged, and older adults

Karen S. Helfer[a)]

Department of Communication Disorders, University of Massachusetts Amherst, 358 North Pleasant Street, Amherst, Massachusetts 01003, USA

Alexandra Jesse

Department of Psychological and Brain Sciences, University of Massachusetts Amherst, 135 Hicks Way, Amherst, Massachusetts 01003, USA

The influence of lexical characteristics of words in to-be-attended and to-be-ignored speech streams was examined in a competing speech task. Older, middle-aged, and younger adults heard pairs of low-cloze probability sentences in which the frequency or neighborhood density of words was manipulated in either the target speech stream or the masking speech stream. All participants also completed a battery of cognitive measures. As expected, for all groups, target words that occur frequently or that are from sparse lexical neighborhoods were easier to recognize than words that are infrequent or from dense neighborhoods. Compared to other groups, these neighborhood density effects were largest for older adults; the frequency effect was largest for middle-aged adults. Lexical characteristics of words in the to-be-ignored speech stream also affected recognition of to-be-attended words, but only when overall performance was relatively good (that is, when younger participants listened to the speech streams at a more advantageous signal-to-noise ratio). For these listeners, to-be-ignored masker words from sparse neighborhoods interfered with recognition of target speech more than masker words from dense neighborhoods. Amount of hearing loss and cognitive abilities relating to attentional control modulated overall performance as well as the strength of lexical influences. © 2015 Acoustical Society of America.
[http://dx.doi.org/10.1121/1.4923155]

[MSS]

Pages: 363–376

I. INTRODUCTION

When we hear a spoken word, words stored in the mental lexicon are considered as viable target words relative to the degree to which they match the incoming information. Models of spoken-word recognition assume this activation of multiple lexical candidates, although the models differ in the exact implementation of the nature of lexical competition (e.g., Gaskell and Marslen-Wilson, 1997; Luce and Pisoni, 1998; McClelland and Elman, 1986; Norris and McQueen, 2008). The ease of recognizing a spoken word thus depends, among other factors, on its similarity to the words stored in the lexicon. A spoken word that is highly similar in its phonological form to many other words (i.e., a word that comes from a dense lexical neighborhood) activates more potential candidates, and is therefore more difficult to recognize (e.g., Dirks et al., 2001; Luce, 1986; Luce and Pisoni, 1998), than a spoken word that is phonologically similar to only a few other words (i.e., a word that comes from a sparse neighborhood). The frequency of occurrence of a spoken word also influences its recognition, as words that people encounter more often in their daily lives are recognized more easily than those that occur less frequently (e.g., Dahan et al., 2001; Goldinger et al., 1989; Howes, 1957).

The purpose of this study was to examine how neighborhood density and word frequency affect spoken-word recognition by younger, middle-aged, and older adults in complex listening situations with two audible speakers. Results of a number of studies of word recognition suggest that lexical factors play a more important role for older individuals than for younger adults (Dirks et al., 2001; Lash et al., 2013; Revill and Spieler, 2012; Spieler and Balota, 2000; Sommers and Danielson, 1999; Taler et al., 2010). Age-related hearing loss likely underlies some of these age differences (Taler et al., 2010), as degradation of sensory input appears to increase lexical effects (Janse and Newman, 2013; Luce and Pisoni, 1998; Norris and McQueen, 2008). However, cognitive functioning (especially inhibitory ability) might also come into play, as performance on cognitive tasks capturing the ability to ignore irrelevant information (as measured by the Stroop task) and short term memory (measured by the digit span task) are correlated with frequency and/or density effects (Sommers and Danielson, 1999; Taler et al., 2010).

Within a theoretical framework of spoken-word recognition in which words are assumed to actively compete for recognition by inhibiting each other, age-related problems identifying words with many lexical neighbors (Sommers and Danielson, 1999) could be explained by assuming that older listeners would be less able to inhibit competing neighbors. However, other studies have not found an association between inhibitory ability and neighborhood effects in older

adults (Ben-David *et al.*, 2011; Revill and Spieler, 2012). In the present study, we investigate the extent to which age-related high-frequency hearing loss and decline in cognitive functions can explain individual differences across younger, middle-aged, and older listeners in the sensitivity to lexical properties of words. Including middle-aged listeners should allow us to determine when and how this increased sensitivity emerges.

Another primary goal was to determine whether lexical factors in *to-be-ignored* speech influence the processing of *to-be-attended* speech, and whether and how these influences vary with aging. Although the influence of lexical characteristics on word recognition is well established, little is known about how neighborhood density and word frequency in to-be-ignored speech streams affect performance. There is evidence that unattended speech is processed at least to some extent linguistically (e.g., Cherry, 1953). This linguistic processing of unattended speech impacts the processing of the to-be-attended speech stream (e.g., Brungart *et al.*, 2001; Carhart *et al.*, 1969; Freyman *et al.*, 1999), as understandable maskers influence the recognition of to-be-attended speech more than speech maskers that are not understandable (e.g., Calandruccio *et al.*, 2010; Freyman *et al.*, 1999; Van Engen and Bradlow, 2007). This raises the possibility that to-be-ignored (i.e., masker) words could compete for recognition with the words in the to-be-attended (i.e., target) speech stream, therefore interfering with their recognition. To our knowledge there has been only one study that considered how competing speech might interfere with lexical activation and competition. Boulenger *et al.* (2010) asked participants for lexical decisions on items in a target speech stream while manipulating the lexical frequency of stimuli in an accompanying masker speech stream (which consisted of lists of words spoken by 2, 4, 6, or 8 talkers). Their results suggest that words in the masker can interfere with the processing of target words, as reaction times to words in the target stream were longer when the to-be-ignored speech stream contained higher frequency words. This masker frequency effect was, however, limited to listening situations with a two-talker masker, suggesting that word frequency is only relevant when maskers are understandable. Hence, the results of Boulenger *et al.* suggest that the presence of comprehensible competing speech messages can alter how a to-be-attended signal is processed.

Here, we investigated the effect of two lexical properties of words from a single-talker masker on target recognition: word frequency and neighborhood density. Neighbors were defined as all words that can be formed by adding, deleting, or substituting one phoneme. We were particularly interested in how these lexical properties of to-be-ignored words affect target word recognition and how this changes with aging. Results of previous studies have demonstrated that understandable competing speech messages are particularly difficult for older adult listeners (e.g., Helfer and Freyman, 2008; Humes and Coughlin, 2009; Rossi-Katz and Arehart, 2009; Tun *et al.*, 2002). This suggests that older individuals might be especially susceptible to lexical properties of words in a to-be-ignored speech stream. If words in the masker are being processed at some level, lexically easy masker words (those from sparse lexical neighborhoods, or those that occur frequently) might be expected to interfere with the lexical processing of words in the target speech stream. If problems ignoring speech maskers are more of a factor for older than younger listeners, we would expect to find greater effects of lexical properties of to-be-ignored words with age.

This study extends previous research in four important ways: by testing whether lexical properties of the target words affect their recognition in a listening situation with a single-talker speech masker; by determining whether lexical properties of the masker influence target recognition; by adding a group of middle-aged participants to determine the onset of aging effects; and by examining what perceptual and cognitive skills predict individual differences in these listening situations. We tested two hypotheses: first, that lexical characteristics of to-be-ignored words influence performance on a competing speech task; and second, that age-related changes in hearing and cognition lead to greater susceptibility to lexical characteristics of words in both speech streams in a competing speech task. We tested these hypotheses by manipulating lexical characteristics of both to-be-attended speech and to-be-ignored speech streams in a single-talker competing speech task. We included groups of younger, middle-aged, and older listeners, from whom both pure-tone thresholds and selected cognitive abilities also were measured. An additional group of younger listeners was tested at a more difficult signal-to-noise ratio (SNR) of -4 dB in order to assess effects of the degradation of sensory input, independent of aging.

II. METHODS

A. Participants

A total of 60 individuals with English as their native language participated in this study: Fifteen older (60–83 yrs, $M = 68$ yrs, 12 women); 15 middle-aged (45–59 yrs, $M = 51$ yrs, 11 women), and 30 younger, normally-hearing adults (19–24 yrs, $M = 21$ yrs, 29 women). All participants reported having normal or corrected-to-normal vision. None of the participants had a self-reported history of neurologic or otologic disorder or had reported experiencing substantial occupational or recreational noise exposure. Degree of hearing loss was restricted in the older and middle-age listeners in order to minimize the influence of audibility: the high-frequency pure-tone average (HFPTA) (average of 2–6 kHz thresholds) in each ear for each participant could be no more than 65 dB hearing level (HL). Younger participants had pure-tone thresholds no higher than 25 dB HL in each ear throughout the standard audiometric range of frequencies (0.25–8 kHz). In order to rule out a conductive component to the hearing loss, all participants were required to have normal tympanograms on the test days. Older and middle-aged participants needed to score at least 26 on the Mini Mental Status Exam (Folstein *et al.*, 1975) in order to participate in this study.

B. Stimulus materials

Stimuli for the present study were pairs of target and masker sentences that were modifications of the TVM

sentence corpus (Helfer and Freyman, 2009). Target sentences always began with the cue name Theo, whereas half of all masker sentences began with Victor and half began with Michael. All sentences had the same syntactic structure: "*Cue name* discussed the ___ and the ___ today," where underlines represent one-syllable key words used for scoring. The frequency and neighborhood density of both key words in the sentences were manipulated to create four types of sentences: both key words occurring rarely (according to the Hoosier Mental Lexicon; Nusbaum *et al.*, 1984), both key words occurring frequently, both coming from dense neighborhoods, or both coming from sparse neighborhoods. To manipulate frequency and neighborhood density in both the target and masker streams, half of the sentences of each type were used as target sentences and half as masker sentences. Sentences for the frequency manipulation were paired with sentences containing words from a mid-frequency range (see Table I). Sentences for neighborhood manipulation were paired with sentences having key words from the mid-range of density. Hence 8 sets of sentence pairs (with 26 pairs per set) were created: frequent targets and rare targets with mid-frequency maskers; dense targets and sparse targets with mid-density maskers; mid-frequency targets with frequent and rare maskers; and mid-density targets with dense and sparse maskers. Two additional sets of target sentences were developed for use with a steady-state noise masker: one list of sentences contained key words that were lexically easy (being both frequent and from sparse neighborhoods) and one list of sentences had lexically difficult key words (being both rare and from dense neighborhoods).

Table I shows the characteristics of the stimuli used in this study. Care was taken to assure that all four key words within a pair of sentences were not semantically related and did not rhyme or begin with the same phoneme. Neighborhood density and frequency values of monosyllabic words were taken from the on-line Washington University Neighborhood Activation Model database (Sommers, 2000). Lists were equated for density and frequency within and across type: sentences in each of the "dense" lists had approximately the same density and frequency among lists, as did all "sparse" lists, "rare" lists, "frequent" lists, etc. All mid-density and mid-frequency sentences were matched on both density and frequency. One-tailed *t*-tests (with corrections for unequal variances) comparing pairs of lists on the critical measure (either neighborhood density or frequency) indicated statistically significant comparisons in all cases ($p < 0.001$). Moreover, pairs of lists did not differ significantly from each other on the non-critical lexical property ($p > 0.08$ in all cases).

Each sentence was audio recorded from two female talkers who were instructed to speak in a conversational manner but to attempt to put equal emphasis on each of the two scoring words. Recordings were made in a sound-treated audiometric chamber. Individual sentences were excised from the original recordings and then were equalized for root-mean-square amplitude. The speech-shaped noise masker was created by concatenating sentences from a different female talker and creating a custom fast-Fourier-transform filter from the envelope of these sentences. This filter was used to shape wide-band steady-state noise.

C. Procedure

Each participant completed standard pure-tone audiometric testing, tympanometry, and a battery of tests that measured cognitive skills that may play a role in speech understanding (e.g., Akeroyd, 2008; Anderson *et al.*, 2013; Helfer *et al.*, 2013; Helfer and Freyman, 2014; Humes *et al.*, 2006; Tun and Wingfield, 1999; Tun *et al.*, 2002; Desjardins and Dougherty 2013; Woods *et al.*, 2013). The older and middle-aged participants completed these initial tests in one visit and the speech recognition assessment in a second visit. Younger participants completed all tests in one session.

1. Cognitive tests

a. Letter-number sequencing. The control portion of the Letter-Number Sequence (LNS) test (Gold *et al.*, 1997) was administered to measure auditory short-term memory load. In this "LNS Forward" condition, sequences of letters and numbers (e.g., "3C9B4") were read to participants at a rate of approximately one item/s. The participant had to repeat back the sequence in the order heard. Sequence length ranged from two to seven items, with four trials per sequence length. Starting with two-item sequences, all four sequences of the same length were presented before increasing the number of items. Testing ended when a participant missed all four trials at a given sequence length. The score on this test was the total number of sequences recalled correctly.

TABLE I. Mean word frequency and neighborhood density (*N*) for key words in target-masker sentence pairs within each data set.

Manipulations in data sets	Target category	Target frequency	Target *N*	Masker category	Masker frequency	Masker *N*
Target *N*	Sparse	15.92	5.21	Mid-density	17.31	12.94
	Dense	14.50	24.33	Mid-density	12.08	13.21
Target frequency	Rare	2.54	14.21	Mid-frequency	14.21	13.06
	Frequent	181.67	15.02	Mid-frequency	14.79	13.10
Masker *N*	Mid-density	14.56	13.15	Sparse	18.33	5.42
	Mid-density	14.27	12.60	Dense	11.67	24.48
Masker frequency	Mid-frequency	13.85	12.37	Rare	2.44	13.85
	Mid-frequency	13.92	12.42	Frequent	171.82	14.90
Lexical difficulty	Hard	21.17	23.38	Steady-state noise	—	—
	Easy	64.75	11.38	Steady-state noise	—	—

b. SICSPAN. Working memory was measured using the size-comparison span task (SICSPAN; Sorqvist *et al.*, 2010). Participants were shown size-comparison sentences on a computer monitor (e.g., "Is a cow larger than an elephant?") and clicked on "yes" or "no" in response. After each sentence, participants were shown a word that was in the same semantic category as those in the size comparison sentences and were instructed to remember these words. Words from the size comparison sentences can hence lead to possible intrusion errors. Participants completed ten lists, with between two and six words per list. Each list consisted of words from a different semantic category. At the end of each list of sentences participants were prompted to recall the to-be-remembered words in the order they were shown. For the current study, the metric for analyzing the SICSPAN was the number of words recalled in the correct order.

c. Stroop task. A computerized version of a Stroop task (Jesse and Janse, 2012) was used to measure inhibitory ability. Colored rectangles displayed on a computer monitor contained, in a neutral condition, the symbol string "###" and in the incongruent condition, a color name (red, blue, or green) that was not the color of the rectangle. Participants named the color of the rectangle as quickly as possible. Verbal response times (the interval between the onset of presentation of the rectangle and the onset of the subject's verbal response) were analyzed off-line for each participant's 50 neutral and 50 incongruent trials. Responses that were incorrect or that contained non-word utterances at the onset (e.g., "uh...") were not scored. A normalized Stroop metric was used in the present study, which was the difference in mean response time for each subject's neutral vs incongruent trials, divided by the mean response time in the neutral condition.

d. Connections test. Executive function and cognitive processing speed were measured using The Connections Test (Salthouse *et al.*, 2000; Salthouse, 2011), a modification of a trail-making task. Participants were given forms that contained numbers and/or letters encased in circles. In the simple version (Test A) participants connected either letters or numbers in sequence (e.g., A-B-C-D... or 1-2-3-4...). In the alternating version (Test B), participants connected letters and numbers in an alternating sequence (e.g., A-1-B-2-C-3...). Participants had 20 s to work on each of 4 simple forms and 4 alternating forms and were told to work as quickly as possible. The score for this task was mean performance in the alternating version (Test B) divided by the mean performance in the simple version (Test A).

e. Visual elevator task. The final cognitive test administered was the Visual Elevator task from the Test of Everyday Attention (Robertson *et al.*, 1996). On each trial of this self-paced test of attention switching, participants were shown figures containing a series of elevator doors, representing an elevator that moves from floor to floor. In between these pictures of elevators, a large vertical arrow was sometimes shown to indicate that the elevator switched direction. Participants were instructed to keep track of the

floor of the elevator as it "moved." At the end of each trial, participants reported the floor on which this imaginary elevator stopped. Ten trials were completed, each having between two and six direction switches. The metric used for the current study was the time needed to complete each trial divided by the number of direction switches, averaged across all correct trials.

2. Speech recognition

Each participant heard a total of 260 target sentences in the presence of a masker. Target sentences were either presented with speech from a single competing talker or with steady-state noise. Both targets and maskers were routed to a single front loudspeaker located 1.3 m from the listener at a height approximating that of a seated adult (1.2 m). Stimuli were presented at -1 dB SNR for the older and middle-aged participants. Half of the younger participants heard the stimuli at this SNR; the SNR for the other 15 younger participants was -4 dB. This allowed for age group comparisons with younger adults tested at the same SNR as the other participants, as well as with younger adults tested in a more difficult listening condition.

For target sentences presented with a speech masker, two cues were available to help listeners distinguish the to-be-attended sentence from the to-be-ignored sentence: voice information and cue name. All target sentences were recorded by the same female talker and all masking sentences were utterances from another female talker. Hence, participants could use this consistent voice information to help delineate the to-be-attended speech stream from the to-be-ignored speech stream. Listeners also could use the first word of each sentence: target utterances always began with "Theo" while the masker began with "Victor" or "Michael."

Participants were instructed to repeat back the two key words in the sentence beginning with "Theo" as quickly as possible. An experimenter (who was blind to condition) scored these verbal responses in real-time and recorded participants' actual incorrect responses, which allowed for off-line analysis of error patterns. Participants were given a 15-item practice set of trials before data collection began.

Ten types of 26-item target-masker pairs were presented: frequent or rare target sentences with mid-frequency masking sentences; dense or sparse target sentences with mid-density masking sentences; mid-frequency target sentences with rare or frequent masking sentences; mid-density target sentences with dense or sparse masking sentences; and hard or easy target sentences presented in steady-state speech-shaped noise. Half of the target-masker pairs for each of these types of stimuli were presented with an unlimited response time; for the other half, participants needed to complete their response by 4 s after the end of the stimulus (this value was based on pilot testing that suggested a reduction in performance with this time limitation). Sentence pair type and limited/unlimited response time were randomized from trial to trial. Response-time limited trials were signaled with a blinking icon on the computer screen during the trial; responses made after 4 s were scored as missing. This manipulation was included to examine the effect of forcing

Karen S. Helfer and Alexandra Jesse

individuals to respond in a limited amount of time, as occurs during conversation. However, since there was no significant effect of response time limitation, data from the time-limited and unlimited trials were pooled for all analyses.

III. RESULTS

A. Audiometric and cognitive tests

Table II displays descriptive statistics for cognitive test performance and high-frequency pure-tone thresholds for participants in this study. Analyses of variance (ANOVA) with subject group as a between-subjects variable showed significant differences for the following measures: HFPTA [$F(3,56) = 21.73$, $p < 0.001$]; Stroop interference [$F(3,56) = 4.48$, $p = 0.007$]; Connections [$F(3,56) = 4.11$, $p = 0.011$]; and SICSPAN [$F(3,56) = 9.51$, $p < 0.001$]. *Post hoc* Bonferroni-corrected *t*-tests demonstrated significant differences between the older and younger groups for all of these measures. In addition, the middle-aged participants had significantly higher pure-tone thresholds, as compared to the younger listeners, and SICSPAN (working memory) scores that were significantly better than those of the older participants. Group differences for the LNS (short-term memory) and Visual Elevator (attention switching) tasks did not reach statistical significance [LNS: $F(3,56) = 2.19$, $p = 0.099$; Visual elevator: $F(3,56) = 1.43$, $p = 0.245$].

B. Speech recognition

Responses given for each trial were scored as correct if they matched one of the two presented key words in the target stream. Trials with no response given (3% of all trials) were excluded. The proportion of the number of correct answers (0, 1, or 2) for each trial was calculated and transformed into empirical logits. Similarly, we calculated the logit-transformed proportion of masker responses out of all responses given for each trial. Masker responses were those incorrect responses that matched a key word from the masker speech stream. All statistical analyses on these two dependent variables used logit mixed effect modeling

TABLE II. Participant characteristics for each listener group. Hearing acuity = average of thresholds at 2, 3, 4, and 6 kHz in the better ear (in dB HL); Stroop = normalized Stroop effect, defined as the difference in mean response time in the incongruent and neutral condition divided by the mean response time in the neutral condition; SICSPAN = number of items recalled in the correct order; LNS = number of correct trials; Connections = ratio of mean performance (number of correct connections on the alternating trials divided by mean performance on simple trials); Visual Elevator = average time per direction switch (in seconds). Standard deviations are given in parentheses.

Characteristic	Younger (−1 SNR)	Middle-aged	Older
Age	20.87 (0.92)	50.93 (4.59)	67.60 (6.06)
Hearing acuity	3.93 (3.20)	17.27 (7.65)	24.13 (14.34)
Stroop	−0.14 (0.12)	−0.19 (0.12)	−0.29 (0.17)
SICSPAN	30.67 (4.51)	30.40 (4.15)	21.93 (7.51)
LNS	20.40 (2.30)	20.33 (2.97)	18.27 (3.61)
Connections	0.55 (0.05)	0.57 (0.07)	0.64 (0.10)
Visual elevator	3.84 (0.86)	3.64 (0.80)	4.27 (1.39)

(Dixon, 2008; Jaeger, 2008), as implemented in the lmer functions (lme4 package, Bates and Sarkar, 2009) of the R statistical program (Version 2.8.1; R Development Core Team, 2007). Analyses were conducted separately on these two dependent variables for data sets manipulating frequency and neighborhood density, respectively, in the target and in the speech masker streams. Another set of analyses evaluated the effect of lexical difficulty (low-frequency words from dense neighborhoods vs high-frequency words from sparse neighborhoods) on target recognition in the presence of the steady-state noise masker. In all of these analyses, the listener group as well as frequency, neighborhood density, and lexical difficulty were evaluated as fixed categorical factors. Regression weights of categorical factors reflect the adjustment to the intercept across conditions. For example, the estimate for frequency indicates how the intercept of a model fit to data from trials with lower frequency words is adjusted to fit data from trials with higher frequency words. Frequency (−0.5 for rare and +0.5 for frequent), neighborhood density (−0.5 for sparse and +0.5 for dense), and lexical difficulty (−0.5 for easy and +0.5 for hard) were contrast coded. Group (older adults, middle-aged adults, younger adults tested at −1 dB SNR, younger adults tested at −4 SNR) was a treatment-coded fixed factor, with the middle-aged group as the reference group mapped onto the intercept. Further planned tests were used to compare other listener groups. In particular, one planned comparison was between younger listeners tested at −4 dB SNR and those tested at the more favorable SNR of −1 dB. All models contained both subjects and items as random effects. *P*-values were estimated with Markov chain Monte Carlo simulations ($n = 10\,000$).

1. Target word recognition accuracy

Figure 1 shows the mean proportion of correct target recognition in listening situations with a single-talker speech masker as a function of listener group, separately for the manipulation of frequency and neighborhood density of target and masker words. In general, performance decreased as listener age increased. The recognition of target words was modulated by both their lexical characteristics as well as (to a smaller extent) by the neighborhood density of the accompanying masker words. These effects seem to vary with age, such that older listeners were, compared to middle-aged listeners, less sensitive to the target word frequency but more sensitive to target neighborhood density. The neighborhood density of the masker words also affected performance for some listeners, in that target recognition suffered when maskers came from sparse rather than dense neighborhoods.

A first set of statistical analyses examined, separately, the effect of frequency and neighborhood density of words in the target stream on their recognition by different listener groups. A second set of analyses examined how frequency and neighborhood density of masker words affected recognition of words in the target stream by different listener groups. Table III shows the results of these statistical analyses. Only the results of additional planned tests are reported in the text. In general, older listeners performed overall

J. Acoust. Soc. Am. **138** (1), July 2015

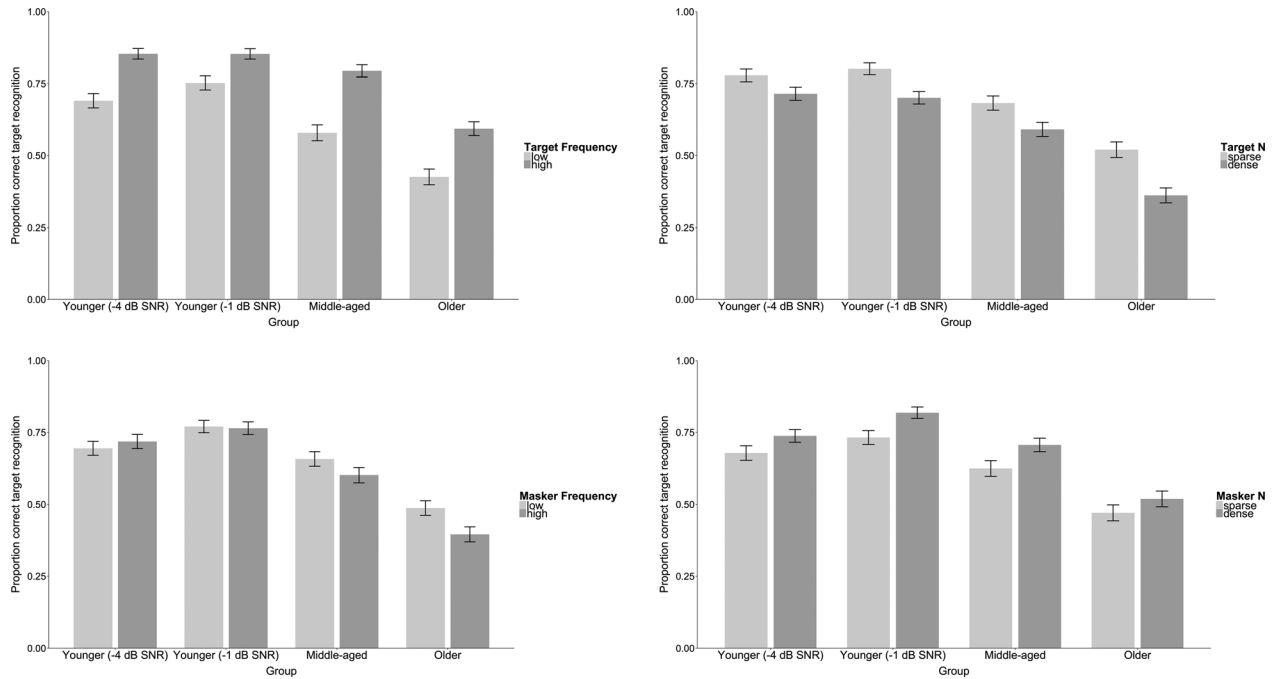Karen S. Helfer and Alexandra Jesse   367

FIG. 1. Proportion of correct target recognition across listener groups as a function of manipulating frequency and neighborhood density (N) in target and masker words in listening situations with a single-talker masker.

worse than middle-aged listeners, who, in turn, performed worse than the younger listener group tested at −1 dB SNR. Middle-aged adults performed similarly to younger adults tested at −4 dB SNR, with the exception of the masker frequency data set.

Lexical characteristics of the target words affected their recognition in situations where a second speaker was also audible. Words in the target stream were more often accurately recognized when they were frequent rather than rare and when they came from sparse rather than dense neighborhoods. Planned further tests showed these effects for all listener groups, except for a lack of a neighborhood density effect for younger listeners tested at −4 dB SNR ($p = 0.3$). The size of these effects varied, however, across groups. In terms of the frequency effect, middle-aged adults and younger adults tested at −4 dB SNR showed a similar-sized effect, as did older adults and younger adults tested at −1 dB SNR [$\beta = 0.10$, standard error (SE) = 0.06, $p = 0.20$]. The latter two groups showed a smaller frequency effect compared to the middle-aged group. A different pattern emerged for the effect of target words' neighborhood density. The size of the neighborhood density effect for any listener group

TABLE III. Linear mixed-effect models for accuracy of target word recognition in the speech-masker condition.

	Target frequency			Target density		
	Estimate	SE	p	Estimate	SE	p
Lexical property	0.436	0.09	<0.0001	−0.207	0.09	0.02
Older adults	−0.317	0.1	0.002	−0.395	0.09	<0.0001
Younger adults (−1 dB SNR)	0.246	0.1	0.02	0.221	0.09	0.02
Younger adults (−4 dB SNR)	0.172	0.1	0.09	0.202	0.09	0.03
Lexical property * older adults	−0.165	0.06	0.007	−0.112	0.06	0.08
Lexical property * younger adults (−1 dB SNR)	−0.242	0.06	<0.0001	−0.012	0.06	0.8
Lexical property * younger adults (−4 dB SNR)	−0.065	0.06	0.28	0.1	0.06	0.1
	Masker frequency			Masker density		
	Estimate	SE	p	Estimate	SE	p
Lexical property	−0.154	0.1	0.14	0.189	0.1	0.07
Older adults	−0.332	0.1	0.0004	−0.331	0.1	0.001
Younger adults (−1 dB SNR)	0.252	0.09	0.008	0.206	0.1	0.04
Younger adults (−4 dB SNR)	0.159	0.09	0.09	0.089	0.1	0.37
Lexical property * older adults	0.012	0.06	0.19	−0.087	0.06	0.16
Lexical property * younger adults (−1 dB SNR)	0.166	0.06	0.007	0.04	0.06	0.51
Lexical property * younger adults (−4 dB SNR)	0.219	0.06	0.0004	−0.055	0.06	0.37

Karen S. Helfer and Alexandra Jesse

was the same as for the middle-aged adults. Older listeners did not differ in the size of the neighborhood density effect from that of younger listeners tested at -1 dB SNR ($\beta = 0.1$, SE $= 0.06$, $p = 0.1$), but showed a significantly larger neighborhood density effect when compared to the younger listener group tested at -4 dB SNR ($\beta = 0.213$, SE $= 0.06$, $p < 0.000\,01$), as the latter group was the only one that did not show a significant neighborhood density effect. Within the younger listeners, SNR did not modulate the size of the neighborhood density effect ($\beta = 0.11$, SE $= 0.06$, $p = 0.057$), but the size of the frequency effect was larger at the less favorable SNR of -4 dB than at -1 dB ($\beta = 0.179$, SE $= 0.06$, $p = 0.002$).

Lexical characteristics of the masker words also affected target word recognition. Target words were somewhat more reliably recognized in the target stream when the masker word was more difficult to recognize due to the density of its neighborhood. However, this effect was only significant for younger listeners when tested at -1 dB SNR ($\beta = 0.227$, SE $= 0.12$, $p = 0.03$). Overall, the neighborhood density of the masker words had only a marginally significant effect on the recognition of words in the target stream ($p = 0.07$). In contrast, the recognition of target words was not affected by the frequency of the masker words. Even though this non-significant trend was larger for middle-aged adults than for younger adults, planned tests showed that the trend never became significant for any listener group ($p > 0.05$).

A third set of analyses examined the joint effect of neighborhood density and frequency of target words in a listening situation with a steady-state noise masker. These results are provided in Table IV. Figure 2 shows the data for this condition as a function of lexical difficulty and listener group. In steady-state noise, older listeners and younger listeners tested at -4 dB SNR recognized fewer target words than middle-aged adults. Younger listeners tested at -1 dB SNR and middle-aged adults performed similarly overall. The effect of target word difficulty was not significant overall. Compared to the middle-aged group, older adults and younger adults tested at -4 dB SNR were, however, more affected by lexical difficulty. Indeed, planned tests showed that this effect was only significant for these two listener groups (older adults; $\beta = -0.188$, SE $= 0.09$, $p = 0.04$; younger adults tested at -4 dB SNR: $\beta = -0.251$, SE $= 0.12$, $p = 0.04$). That is, listeners who had more overall difficulty recognizing words in this listening situation were sensitive to target word difficulty.

In summary, older participants performed, as expected, generally poorer than middle-aged listeners, who, in turn,
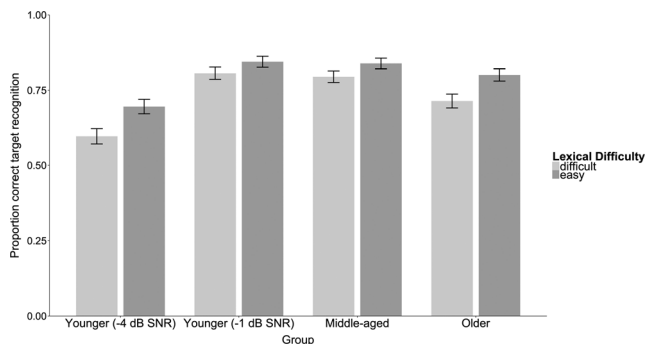

FIG. 2. Proportion of correct target recognition across listener groups as a function of lexical difficulty of target words in listening situations with a steady-state noise masker.

performed worse than younger adults tested at the same SNR. It should be noted that in some listening conditions, speech recognition ability by middle-aged adults was even poorer than that of younger participants listening at a 3 dB more disadvantageous SNR. In the presence of a single competing speech message, frequency and neighborhood density of target words affected word recognition by younger, middle-aged, and older listeners. The importance of neighborhood density of target words was modulated by participant age, as older listeners were more affected by this factor than younger adults tested in the same listening condition. This was probably not due to reduced audibility of the stimuli, since only younger adults who listened at the more favorable -1 dB SNR but not those who listened in the more difficult -4 dB SNR condition showed a neighborhood density effect. The effect of target word frequency was most pronounced in middle-aged adults, compared to older and younger listeners tested at -1 dB SNR. Younger adults tested at the less favorable SNR showed larger frequency effects than younger listeners tested at a better SNR. Most interestingly, when the masker was speech from a second talker, lexical properties of masker words, more precisely their neighborhood density, affected the recognition of words in the target stream. Recognition of the target words was poorer when words in the masker were easier to understand (that is, when they were from sparse lexical neighborhoods), but only for listeners with better overall speech recognition ability; that is, younger listeners tested at a more favorable SNR. In the steady-state noise masker, an effect of lexical difficulty of the target word was found for listeners with overall poorer performance in this condition (older listeners and younger listeners tested at the more difficult SNR of -4 dB).

2. Masker word response errors

When listeners made mistakes, they often responded with a word from the to-be-ignored masker stream (23.94% of all errors, 7.47% of all responses). We analyzed how the probability of responding with a masker word varied as a function of listener group and lexical properties of the target and masker words. Figure 3 shows the data for these analyses. Older listeners gave more masker responses than middle-aged and younger listeners. Lexical properties of the target words

TABLE IV. Linear mixed-effect models for accuracy of target word recognition in the steady-state-noise masker condition.

	Estimate	SE	p
Lexical difficulty	−0.076	0.09	0.42
Older adults	−0.163	0.05	0.003
Younger adults (−1 dB SNR)	−0.019	0.05	0.72
Younger adults (−4 dB SNR)	−0.388	0.05	<0.0001
Lexical property * older adults	−0.11	0.06	0.047
Lexical property * younger adults (−1 dB SNR)	−0.162	0.06	0.77
Lexical property * younger adults (−4 dB SNR)	−0.178	0.06	0.0013

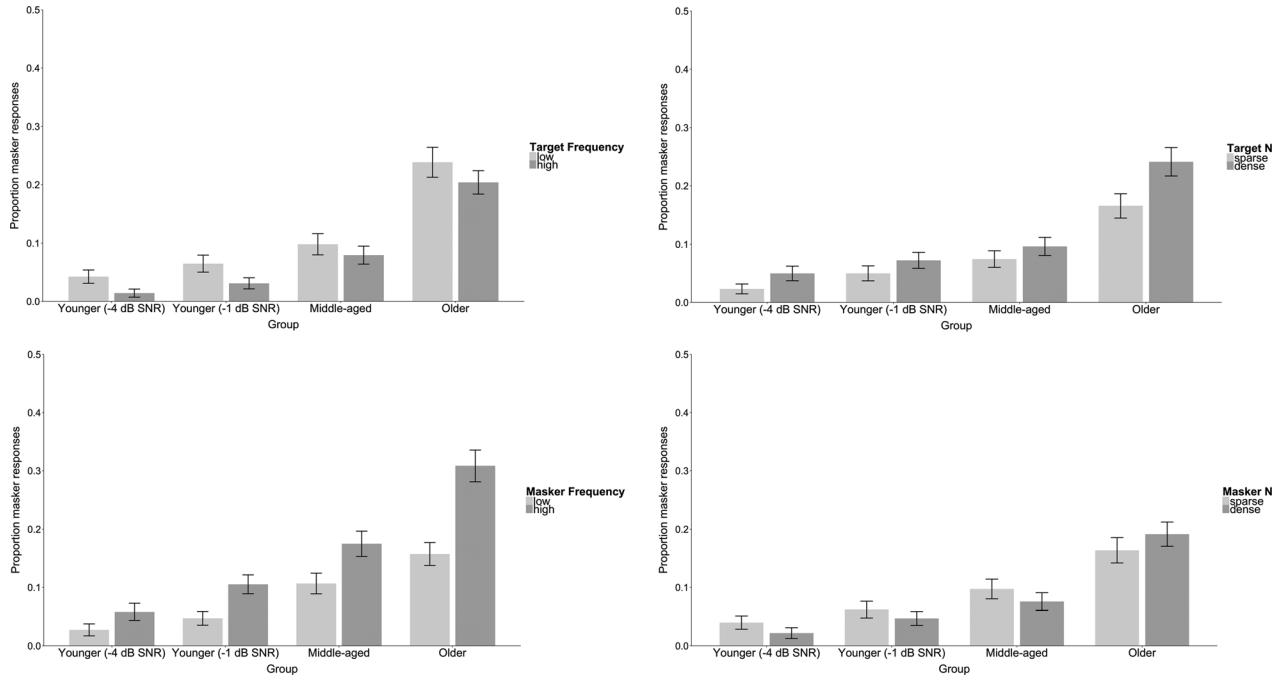Karen S. Helfer and Alexandra Jesse   369

FIG. 3. Proportion of masker responses (calculated out of all responses given) across listener groups as a function of manipulating frequency and neighborhood density (N) in target and masker words, respectively, in listening situations with a single-talker masker.

seemed not to affect the overall probability of responding with a to-be-ignored masker word. Lexical properties of the masker words, however, did modulate the likelihood of giving a masker response. Specifically, masker frequency affected the probability of these responses in older listeners (but less so in middle-aged and younger listeners). Masker density seemed to matter less than masker frequency.

Formal statistical analyses confirmed these observations, as summarized in Table V. In all analyses, older listeners gave more masker responses than middle-aged participants. Middle-aged and younger listener groups performed the same

(all $p > 0.05$), with the exception of younger listeners tested at $-4\,dB$ SNR, who gave fewer masker responses than middle-aged participants for the data sets testing the effect of masker frequency.

The lexical properties of the words in the target streams, i.e., their frequency and neighborhood density, did not determine the probability of erroneously responding with a masker word in the overall analyses (all $p > 0.05$). Planned tests for each listener group individually showed only a significant target frequency effect for younger listeners tested at $-4\,dB$ SNR ($\beta = -0.105$, SE $= 0.03$, $p = 0.002$). These

TABLE V. Linear mixed-effect models for masker word responses in the speech-masker condition.

	Target frequency			Target density		
	Estimate	SE	p	Estimate	SE	p
Lexical property	−0.46	0.05	0.33	0.064	0.05	0.19
Older adults	0.334	0.08	<0.0001	0.326	0.07	<0.0001
Younger adults ($-1\,dB$ SNR)	−0.073	0.08	0.39	−0.042	0.07	0.56
Younger adults ($-4\,dB$ SNR)	−0.123	0.08	0.15	−0.075	0.07	0.29
Lexical property * older adults	−0.036	0.05	0.43	0.049	0.05	0.31
Lexical property * younger adults ($-1\,dB$ SNR)	−0.013	0.05	0.77	−0.008	0.05	0.87
Lexical property * younger adults ($-4\,dB$ SNR)	−0.061	0.05	0.18	−0.018	0.05	0.70
	Masker frequency			Masker density		
	Estimate	SE	p	Estimate	SE	p
Lexical property	0.1667	0.07	0.02	−0.093	0.07	0.08
Older adults	0.265	0.09	0.002	0.272	0.09	0.002
Younger adults ($-1\,dB$ SNR)	−0.119	0.09	0.16	−0.05	0.09	0.57
Younger adults ($-4\,dB$ SNR)	−0.194	0.09	0.03	−0.106	0.09	0.23
Lexical property * older adults	0.052	0.05	0.30	0.122	0.05	0.01
Lexical property * younger adults ($-1\,dB$ SNR)	−0.058	0.05	0.24	−0.01	0.05	0.84
Lexical property * younger adults ($-4\,dB$ SNR)	−0.086	0.05	0.09	0.036	0.05	0.44

 Karen S. Helfer and Alexandra Jesse

participants gave fewer masker responses when trying to recognize more-frequent rather than less-frequent target words. This trend was marginally significant for older listeners ($\beta = -0.08$, SE $= 0.05$, $p = 0.09$).

The lexical properties of the words in the masker stream influenced their likelihood of becoming the response. Masker words were more likely to become the response when they were more frequent rather than less frequent. There was no difference in the size of this effect for any group compared to the middle-aged group (all $p > 0.05$). Further analyses showed a significant masker frequency effect for older adults ($\beta = 0.218$, SE $= 0.09$, $p = 0.01$) and marginally significant effects for middle-aged adults ($\beta = 0.167$, SE $= 0.09$, $p = 0.07$) and for younger adults tested at $-1\,dB$ SNR ($\beta = 0.106$, SE $= 0.06$, $p = 0.07$). The effect was not significant for younger adults tested at $-4\,dB$ SNR ($\beta = 0.078$, SE $= 0.05$, $p = 0.14$). SNR did not modulate any trend of an effect within younger listeners ($p > 0.05$). Masker words were also somewhat more likely to become the response when they came from sparse rather than dense neighborhoods, although this effect failed to reach the significance level ($p = 0.08$). Planned further analyses showed that this effect was only marginally significant for younger listeners tested at $-1\,dB$ SNR ($\beta = -0.103$, SE $= 0.06$, $p = 0.08$) and was not significant for any other group (all $p > 0.05$). Even though the effect was not significant for either group by itself, middle-aged adults were more affected by masker neighborhood density than older adults.

In summary, on average listeners responded about 25% of the time with masker words. These erroneous masker responses were made more often by older participants compared to middle-aged and younger individuals, but were produced to some extent by all listeners. Lexical properties of the masker word modulated the probability of responding with the masker: these errors were made more often when the to-be-ignored word was frequent vs rare. This effect of masker frequency increased with age and reached significance for older adults. Younger adults were affected by the frequency of the target words in their likelihood of responding with a masker, but only when tested at a more difficult SNR. The neighborhood density of target and masker words, however, did not have a significant effect on the likelihood of responding with a word from the masker.

3. Individual differences

We also evaluated to what extent hearing and cognitive abilities explained individual differences in overall

TABLE VI. Pearson coefficients of correlations between participant characteristics.

	Age$_{Resid}$	Hearing acuity	Stroop	Connections
Age$_{Resid}$[a]				
Hearing acuity	0.00			
Stroop	0.16	0.33		
Connections	0.33	0.20	0.01	
Visual elevator	-0.11	0.30	0.22	0.23

[a]Hearing acuity was partialled out of Age.

performance as well as listeners' sensitivity to lexical properties of the words in the streams. Only data on target recognition accuracy from groups tested at $-1\,dB$ SNR in conditions with speech maskers were included. We first checked for intercorrelations between the background measures we evaluated, namely, hearing acuity and performance on the Stroop task, the Connections test, and on the Visual Elevator test. LNS and SICSPAN were not included in the analyses, since both measures were correlated highly with several other cognitive measures upon initial analysis. Given a Bonferroni-corrected alpha level of 0.005, only the correlation between hearing acuity and age was significant ($r = 0.76$, $p = 0.013$). We assigned the shared variance between these two background variables to hearing acuity by residualizing age (Age$_{resid}$). Table VI shows the Pearson correlation coefficients for the participant characteristics. None of these measures correlated with each other (all $p > 0.05$).

Systematic stepwise model comparisons based on likelihood ratio tests established the best fitting model for each dataset. The starting models included the background measures described above and their interactions with frequency or neighborhood density. As age was evaluated as a continuous measure, group was no longer included as a factor. From these full models, we removed all effects in a stepwise fashion that did not contribute to a better fit. We first removed any non-significant interactions, before also removing any non-significant main effects. The main effects of factors that were part of a significant higher-order interaction were not removed. We always tested first the effect with the largest p value for possible removal. The best-fitting models are reported in Table VII. Below we only report significant results. All models included subject and items as random factors.

First, we evaluated what factors predicted the overall recognition of words in a target stream in the presence of a speech masker. For this analysis, we pooled the data across all conditions with speech maskers and did not include frequency or neighborhood density as factors. Listeners were less likely to recognize words in the target stream correctly if they were older (Age$_{resid}$; $\beta = -0.122$, SE $= 0.003$, $p = 0.0038$), had more difficulty ignoring irrelevant information (Stroop; $\beta = -0.947$, SE $= 0.3$, $p = 0.0017$), and were worse at attention switching (Visual Elevator; $\beta = -0.092$, SE $= 0.04$, $p = 0.03$).

Next, we evaluated what predicted individual differences in participants' susceptibility to the target words' frequency and neighborhood density. Just as in the overall pooled analyses, overall target word recognition accuracy decreased in these subset analyses with decline in the ability to ignore irrelevant information (Stroop; frequency data sets: $\beta = -0.934$, SE $= 0.32$, $p = 0.0036$; neighborhood density data sets: $\beta = -0.957$, SE $= 0.3$, $p = 0.0014$) and with age (Age$_{resid}$; frequency data sets: $\beta = -0.014$, SE $= 0.004$, $p = 0.001$; neighborhood density data sets: $\beta = -0.012$, SE $= 0.004$, $p = 0.0012$). The effect of attention switching was only marginally significant for the frequency subset of the data (Visual Elevator; $\beta = -0.079$, SE $= 0.05$, $p = 0.09$). Hearing acuity modulated performance only in the neighborhood density data sets ($\beta = -0.008$, SE $= 0.004$, $p = 0.03$).

J. Acoust. Soc. Am. **138** (1), July 2015

Karen S. Helfer and Alexandra Jesse   371

TABLE VII. Estimates of best-fitting models for the individual differences analyses, split by data sets manipulating frequency and neighborhood density of target and masker words.

Effects	Target manipulation		Masker manipulation	
	Frequency	Neighborhood density	Frequency	Neighborhood density
Lexical characteristic	0.301***	−0.246**	−0.093	0.175#
Age$_{Resid}$	−0.014***	−0.012**	−0.012**	−0.014***
Hearing acuity	−0.004	−0.008*	−0.008*	−0.006
Stroop	−0.934**	−0.957**	−0.853**	−0.915**
Connections	0.693	—	0.600	1.162*
Visual elevator	−0.08#	—	−0.067	−0.087*
Lexical characteristic × Hearing acuity	0.007**	—	−0.008***	−0.009***
Lexical characteristic × Stroop	−0.58***	−0.362*	—	—
Lexical characteristic × Connections	0.761*	—	−0.594#	—
Lexical characteristic × Visual elevator	−0.087***	—	0.047#	—

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, # $p < 0.1$.

In line with the results of the main analyses above, words were more reliably recognized when they were frequent rather than rare ($\beta = 0.3$, SE $= 0.09$, $p = 0.0007$) and when they came from sparse rather than dense neighborhoods ($\beta = -0.246$, SE $= 0.08$, $p = 0.0029$). The frequency effect became larger for those listeners with worse hearing ($\beta = 0.007$, SE $= 0.002$, $p = 0.005$) and worse executive function abilities/processing speed, as measured by the Connections test (Connections; $\beta = 0.76$, SE $= 0.32$, $p = 0.017$). The frequency effect was also larger for listeners with better attention switching skills (Visual Elevator; $\beta = -0.087$, SE $= 0.03$, $p = 0.0006$) but disappeared for those participants with reduced ability to ignore irrelevant information (Stroop; $\beta = -0.58$, SE $= 0.18$, $p = 0.001$). Listeners with poorer ability to ignore irrelevant information showed, however, a larger neighborhood density effect (Stroop; $\beta = -0.36$, SE $= 0.17$, $p = 0.04$).

In another set of analyses, we evaluated predictors of susceptibility to masker word properties. Overall performance in these subsets was lower for individuals with poorer ability to ignore irrelevant information (Stroop; frequency data sets: $\beta = -0.853$, SE $= 0.3$, $p = 0.005$; neighborhood density data sets: $\beta = -0.915$, SE $= 0.29$, $p = 0.0015$) and with advanced age (Age$_{resid}$; frequency data sets: $\beta = -0.012$, SE $= 0.004$, $p = 0.0025$; neighborhood density data sets: $\beta = -0.014$, SE $= 0.004$, $p = 0.0002$). In addition, overall performance in the frequency data set was lower for listeners with worse hearing; this was not the case for the neighborhood data set (frequency data sets: $\beta = -0.008$, SE $= 0.004$, $p = 0.045$; neighborhood density data set: $\beta = -0.006$, SE $= 0.004$, $p = 0.11$). Overall performance for the neighborhood density data set was also lower for individuals with worse attention switching skills (Visual Elevator; $\beta = -0.09$, SE $= 0.04$, $p = 0.04$) and with better Connections test performance ($\beta = 1.161$, SE $= 0.55$, $p = 0.04$). As found in the main analyses reported above, masker words' frequency did not exert a significant effect on overall performance ($\beta = -0.093$, SE $= 0.1$, $p = 0.33$), but the trend of poorer target recognition in the presence of higher vs lower frequency masker words became larger for participants with worse hearing ($\beta = -0.0075$, SE $= 0.004$, $p = 0.0009$). There was also marginally significant evidence that the masker frequency effect went away for those with reduced attention switching skills (Visual Elevator; $\beta = 0.0474$, SE $= 0.04$, $p = 0.07$) and was larger for listeners with worse executive function and slower processing speed, as indexed by the Connections test performance (Connections; $\beta = -0.594$, SE $= 0.58$, $p = 0.07$). Target recognition was marginally significantly better when masker words came from dense than from sparse neighborhoods ($\beta = 0.175$, SE $= 0.1$, $p = 0.08$). This effect was stronger for listeners with better hearing ($\beta = -0.009$, SE $= 0.002$, $p = 0.0001$).

In summary, words in a target speech stream presented in a listening situation with a single-talker speech masker were more accurately recognized when listeners were younger, better at ignoring irrelevant information, and could more effectively switch attention. Hearing acuity modulated lexical effects on target word recognition. Specifically, the poorer the hearing acuity of listeners, the more they were influenced by the frequency of the target and masker words and the less they were influenced by neighborhood density of the maskers. Individuals with poorer executive control skills and slower processing speed (as measured jointly by the Connections Test) and participants with better attention-switching ability were more strongly influenced by target word frequency. Moreover, reduced ability to ignore irrelevant information was associated with decreased impact of target word frequency but increased influence of target word neighborhood density.

IV. DISCUSSION

The present study tested two hypotheses: that lexical characteristics of to-be-ignored words influence performance on a competing speech task; and that age-related changes in hearing and cognition lead to greater susceptibility to lexical characteristics of words in both speech streams in this type of paradigm. As we discuss below, data support the first hypothesis as we found a clear influence of lexical characteristics of the to-be-ignored speech stream on recognition of target words. The second hypothesis was partially supported; both hearing loss and selected cognitive abilities were related to the influence of lexical information, but older

adults were not invariably more affected by lexical properties of the stimuli.

Overall, we found, as expected, that target speech recognition declines and responses with words from the masker stream become more likely as listener age increases. Data from the middle-aged participants in this study support the idea that substantial changes in speech perception can occur in this age group. In most competing speech conditions, our middle-aged participants' performance was poorer than that obtained by younger adult participants, even those who listened at a 3 dB poorer SNR. Of note is that speech recognition ability in steady-state noise was comparable between our middle-aged participants and younger individuals tested at the same SNR.

A. Lexical properties of to-be-attended words

A novel finding in the present study is that lexical properties of the target words affected their recognition in a listening situation with a single-talker speech masker. These effects were found for all listener groups, but to different extents. Middle-aged listeners appeared to be particularly affected by manipulations involving target word frequency, compared to other listener groups. However, it should be noted that younger adults tested at an SNR that caused them to perform at a similar overall level as the middle-aged participants showed frequency effects of the same size, suggesting that these effects are the largest for a middle range of overall performance. Thus, the advantage of using frequently-occurring words appears to be most substantial when the words are sufficiently (but not necessarily entirely) audible.

Older adults (as compared to younger participants tested at a poorer SNR) were more influenced by the density of a target word's neighborhood, that is, by the amount of lexical competition that the target words received. This is in line with the previously proposed idea of reduced inhibitory control in older adults (e.g., Hasher *et al.*, 1991; Lash *et al.*, 2013; Sommers and Danielson, 1999; Taler *et al.*, 2010). Our results further show that this increased sensitivity to lexical competition seems to become significant some time beyond mid-life, as middle-aged listeners were no more sensitive to lexical competition than were younger adults. The present results support previous reports of an increased sensitivity of older adults compared to younger adults to lexical difficulty of target words (Sommers, 1996), in that they suggest that these effects are primarily driven by neighborhood density. Our results are at odds with those of others who have found larger age-related differences for frequency than for neighborhood density (e.g., Dirks *et al.*, 2001; Revill and Spieler, 2012; Spieler and Balota, 2000), as lexical frequency in the to-be-attended stream was equally important for the older listeners in the present study as for younger listeners. This apparent disconnect in findings is likely due, at least in part, to differences in stimuli and procedures.

B. Lexical properties of to-be-ignored words and masker error responses

Perhaps the most compelling finding in the current study is that manipulating neighborhood density and word frequency in a to-be-ignored speech stream affected identification of the to-be-attended words. However, although we had anticipated that our older participants might be most heavily affected by characteristics of the masking words due to a decline in inhibitory ability, this was not the case. Instead, only listeners who performed at relatively high levels of accuracy (i.e., younger adults tested at the more favorable −1 dB SNR) were influenced by the ease at which the masker word was recognized. The finding that only young participants tested in a relatively easy listening condition were affected by lexical properties of the masker could potentially be explained by cognitive load. According to the load theory of attention (e.g., Francis, 2010; Lavie *et al.*, 2004) if one task is very difficult, there will be little interference from a second task because all of the available resources are directed toward resolving the first task. In the present study, individuals who had to devote more of their available resources to segregating the to-be-attended from the to-be-ignored messages and focusing attention on the target stream (whether from increased age, reduced audibility, or poorer cognitive abilities) would have fewer resources left to process the to-be-ignored speech stream. Therefore, manipulations in the to-be-ignored stream would have less of an effect for these individuals, as compared to the younger participants tested in the easier listening condition, who needed to devote fewer resources to segregation and attentional control. Consistent with load theory, the effect of manipulating lexical factors in the masker went away for younger adults when target processing was made more difficult by changing the SNR.

The nature of the masker effects was such that the more similar-sounding words that were activated by a masker word (i.e., the more global activation there was in the lexical neighborhood of the masker word), the greater the likelihood of correctly recognizing the target word. It seems, however, that this effect was driven by the additional activation of more of the surrounding words, rather than by the activation level of the masker word itself, since frequency of the masker had no effect on target recognition. Masker words and their neighbors should receive greater activation when the masker is more rather than less frequent. Yet, the lexical frequency of masker words did not affect target recognition. It seems that the number of strongly competing lexical candidates matters more than the activation level.

The frequency of words in the masker, more so than the density of their neighborhoods, modulated the probability of erroneously responding with a masker word. Specifically, masker words that were frequent were more likely to be reported erroneously than were masker words that occurred infrequently. Only older participants, however, were influenced by the frequency of masker words in giving masker responses, even though all listener groups produced masker errors. Younger participants were influenced by the frequency of the target words in their likelihood of erroneously responding with the masker word, but only when tested in the more difficult SNR. Overall, our results support the proposition that difficulty suppressing lexical competitors could contribute to age-related changes in speech perception (Lash *et al.*, 2013; Sommers, 1996; Sommers and Danielson, 1999)

and could even lead to a greater incidence of "false hearing" in older adults, in which they are highly confident that their incorrect responses are, in fact, correct (Rogers *et al.*, 2012). It should be noted, however, that older adults' propensity for reporting masker words could also reflect a guessing strategy—when faced with the task of repeating back what they heard, it is possible that older adults, when uncertain of whether a word was from the to-be-attended or to-be-ignored stream, responded with the more frequent item. This explanation could account for the fact that although older adults produced more masker errors, it was the younger adults listening at a more favorable SNR for whom lexical factors in the masker influenced target word recognition. The effect of lexical factors on word recognition thus may emerge at different points during recognition for these two age groups, that is, during perception for younger adults and at a later, post-perceptual stage for older adults.

In contrast to our results, Boulenger *et al.* (2010) found that lexical frequency of masker words influenced target word recognition. In that study, young adults performed an auditory lexical decision task on words and non-words in a target stream presented in multitalker babble noise. Listeners were slower at correctly recognizing tokens in the target stream as words when two-talker babble contained frequent rather than less-frequent words. Numerous differences between the two studies exist that make direct comparison difficult [e.g., SNR, type of task (identification vs lexical decision, and accuracy vs reaction times), and nature of the masker (single- vs two-talker competing speech)]. The results of these two studies converge, however, on what they tell us about spoken word recognition in the presence of competing speech: The easier it is to recognize the masker word (either because it has less competition from lexical neighbors or because it is more frequent), the more target recognition suffers, both in terms of accuracy (as shown here) and in terms of speed of recognition (as shown by Boulenger *et al.*, 2010).

C. Accounting for individual differences

Last, we asked what perceptual and cognitive abilities of the listeners might explain their susceptibility to the lexical properties of target and masker words. As expected, high-frequency hearing was one factor that explained individual differences in the influence of lexical properties on speech recognition in our listeners, who ranged in age from young adults to older adults. Target and masker frequency effects in the present study were greater for those with more hearing loss, supporting the idea that sensory degradation increases the effects of word frequency (e.g., Norris and McQueen, 2008). Our results further specified that this increased reliance on word frequency with sensory degradation holds for words in both to-be-attended and to-be-ignored streams. The influence of neighborhood density effects in the masking stream was in the opposite direction: the worse the hearing, the less the influence of the masker's neighbors. Hence, poorer hearing acuity actually provided, in one sense, an advantage in this listening task. If listeners have difficulty accessing acoustic information in the masker, then there is

also less global activation in the neighborhood of the masker. In the present study, only young listeners run at a more advantageous SNR (for whom the masker words were more audible than was the case for the other participants) were affected by masker neighborhood density. So although, as discussed above, cognitive load might explain younger participants' susceptibility to masker characteristics, audibility of the masker also could contribute to this finding.

Our results clearly support a view that attentional control skills also matter in the recognition of speech in the presence of a single competing message. In particular, target recognition suffered for participants with reduced ability to ignore irrelevant information (as measured with the Stroop task) and with poorer attention-switching capability (as measured with the Visual Elevator task). Hence, listeners who could ignore the masker speech stream or re-direct their attention back to the target speech if distracted by the masker coped better in this single-talker competing speech situation. This strongly supports the contention that while age-related hearing loss undoubtedly limits speech recognition, changes in higher-level abilities also contribute, as found in a growing body of previous work (e.g., Anderson *et al.*, 2013; Desjardins and Dougherty 2013; Helfer *et al.*, 2013; Humes *et al.*, 2006; Jesse and Janse, 2012; Tun and Wingfield, 1999; Tun *et al.*, 2002).

Cognitive abilities also mediated susceptibility to lexical effects in target speech streams. Reduced inhibitory ability was associated with less susceptibility to target word frequency, but with higher susceptibility to target word neighborhood density. This supports the idea that the ability to inhibit modulates the influence of lexical competitors (Lash *et al.*, 2013; Sommers, 1996; Sommers and Danielson, 1999). In line with this, participants who were most influenced by frequency of the to-be-attended words also had reduced executive control/processing speed (as measured with the Connections test). Somewhat surprising is the finding that participants with better ability to switch attention were actually more influenced by the word frequency. This association was significant for frequency effects elicited from target words but there was only a marginally significant trend for frequency effects elicited from masker words. One possible explanation for this result could be that listeners who are better able to switch attention used this ability to alternatively sample the target and masking streams, and thus were better able to separately activate target and masker words. This benefitted overall target recognition, but also made these individuals more susceptible to the frequency information stored in the representations of target and masker words. Future research should be directed toward identifying how precisely these cognitive abilities mediate performance in competing speech situations.

D. Summary

In summary, the results of this study suggest that a single competing speech message may be especially disruptive because it activates lexical competitors, which interferes with the phonological processing of words in the to-be-attended message. Hence, listeners process to-be-ignored

speech streams to an extent that lexical characteristics of words in those streams can influence performance. Our results also support the idea that the relative importance of word frequency and neighborhood density changes across the adult lifespan, which likely is due to both age-related hearing loss and degradation of abilities that mediate attentional control.

Akeroyd, M. A. (**2008**). "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. Aud. **47**(2), S53–S71.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (**2013**). "A dynamic auditory-cognitive system supports speech-in-noise perception in older adults," Hear. Res. **300**, 18–32.

Bates, D. M., and Sarkar, D. (**2009**). lme4: Linear mixed-effects models using s4 classes (Version R package version 0.999375-27).

Ben-David, B. M., Chambers, C. G., Daneman, M., Pichora-Fuller, M. K., Reingold, E. M., and Schneider, B. A. (**2011**). "Effects of aging and noise on real-time spoken word recognition: Evidence from eye movements," J. Speech Lang. Hear. Res. **54**, 243–262.

Boulenger, V., Hoen, M., Ferragne, E., Pellegrino, F., and Meunier, F. (**2010**). "Real-time lexical competitions during speech-in-speech comprehension," Speech Comm. **52**, 246–253.

Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (**2001**). "Informational and energetic masking effects in the perception of multiple simultaneous talkers," J. Acoust. Soc. Am. **110**, 2527–2538.

Calandruccio, L., Dhar, S., and Bradlow, A. R. (**2010**). "Speech-on-speech masking with variable access to the linguistic content of the masker speech," J. Acoust. Soc. Am. **128**, 860–869.

Carhart, R., Tillman, T. W., and Greetis, E. S. (**1969**). "Perceptual masking in multiple sound backgrounds," J. Acoust. Soc. Am. **45**, 694–703.

Cherry, E. C. (**1953**). "Some experiments on the recognition of speech with one and two ears," J. Acoust. Soc. Am. **25**, 975–979.

Dahan, D., Magnuson, J. S., and Tanenhaus, M. K. (**2001**). "Time course of frequency effects in spoken-word recognition: Evidence from eye movements," Cognit. Psychol. **42**, 317–367.

Desjardins, J. L., and Doherty, K. A. (**2013**). "Age-related changes in listening effort for various types of masker noises," Ear Hear. **34**, 261–272.

Dirks, D. D., Takayanagi, S., Moshfegh, A., Noffsinger, P. D., and Fausti, S. A. (**2001**). "Examination of the neighborhood activation theory in normal and hearing-impaired listeners," Ear Hear. **22**, 1–13.

Dixon, P. (**2008**). "Models of accuracy in repeated-measures design," J. Mem. Lang. **59**, 447–456.

Folstein, M. F., Folstein, S. E., and McHugh, P. R. (**1975**). "Mini-mental state: A practical method for grading the cognitive state of patients for the clinician," J. Psychiat. Res. **12**, 189–198.

Francis, A. L. (**2010**). "Improved segregation of simultaneous talkers differentially affects perceptual and cognitive capacity demands for recognizing speech in competing speech," Percep. Psychophys. **72**, 501–516.

Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (**1999**). "The role of perceived spatial separation in the unmasking of speech," J. Acoust. Soc. Am. **106**, 3578–3588.

Gaskell, M. G., and Marslen-Wilson, W. D. (**1997**). "Integrating form and meaning: A distributed model of speech perception," Lang. Cognit. Processes **12**, 613–656.

Gold, J. M., Carpenter, C., Randolph, C., Goldberg, T. E., and Weinberger, D. R. (**1997**). "Auditory working memory and Wisconsin Card Sorting Test performance in schizophrenia," Arch. Gen. Psychiat. **54**, 159–165.

Goldinger, S. D., Luce, P. A., and Pisoni, D. B. (**1989**). "Priming lexical neighbors of spoken words: Effects of competition and inhibition," J. Mem. Lang. **28**, 501–518.

Hasher, L., Stoltzfus, E. R., Zacks, R. T., and Rypma, B. (**1991**). "Age and inhibition," J. Exp. Psychol. Learn. Mem. Cog. **17**, 163–169.

Helfer, K. S., and Freyman, R. L. (**2008**). "Aging and speech-on-speech masking," Ear Hear. **29**, 87–98.

Helfer, K. S., and Freyman, R. L. (**2009**). "Lexical and indexical cues in masking by competing speech," J. Acoust. Soc. Am. **125**, 447–456.

Helfer, K. S., and Freyman, R. L. (**2014**). "Stimulus and listener factors affecting age-related changes in competing speech perception," J. Acoust. Soc. Am. **136**, 748–759.

Helfer, K. S., Mason, C., and Marino, C. (**2013**). "Aging and the perception of temporally-interleaved words," Ear Hear. **34**, 160–167.

Howes, D. (**1957**). "On the relation between probability of a word as an Association and in general linguistic usage," J. Abnorm. Psychol. **54**, 75–85.

Humes, L. E., and Coughlin, M. P. (**2009**). "Aided speech-identification performance in single-talker competition by older adults with impaired hearing," Scand. J. Psychol. **50**, 485–494.

Humes, L. E., Lee, J. H., and Coughlin, M. P. (**2006**). "Auditory measures of selective and divided attention in young and older adults using single-talker competition," J. Acoust. Soc. Am. **120**, 2926–2937.

Jaeger, T. F. (**2008**). "Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models," J. Mem. Lang. **59**, 434–446.

Janse, E., and Newman, R. S. (**2013**). "Identifying nonwords: Effects of lexical neighborhoods, phonotactic probability, and listener characteristics," Lang. Speech **56**, 421–441.

Jesse, A., and Janse, E. (**2012**). "Audiovisual benefit for recognition of speech presented with single-talker noise in older listeners," Lang. Cognit. Processes **27**, 1167–1191.

Lash, A., Rogers, C. S., Zoller, A., and Wingfield, A. (**2013**). "Expectation and entropy in spoken word recognition: Effects of age and hearing acuity," Exp. Aging Res. **39**, 235–253.

Lavie, N., Hirst, A., de Fockert, J. W., and Viding, E. (**2004**). "Load theory of selective attention and cognitive control," J. Exp. Psychol. Gen. **133**, 339–354.

Luce, P. A. (**1986**). "Neighborhoods of words in the mental lexicon," Ph.D. dissertation, Indiana University, in Research on Speech Perception, Technical Report No. 6, Speech Research Laboratory, Department of Psychology, Indiana University.

Luce, P. A., and Pisoni, D. B. (**1998**). "Recognizing spoken words: The neighborhood activation model," Ear Hear. **19**, 1–36.

McClelland, J. L., and Elman, J. L. (**1986**). "The TRACE model of speech perception," Cognit. Psychol. **18**, 1–86.

Norris, D., and McQueen, J. M. (**2008**). "Shortlist B: A Bayesian model of continuous speech recognition," Psychol. Rev. **115**, 357–395.

Nusbaum, H. C., Pisoni, D. B., and Davis, C. K. (**1984**). "Sizing up the Hoosier Mental Lexicon: Measuring the familiarity of 20,000 words," Research on Speech Perception Progress Report #10, Indiana University.

R Development Core Team (**2007**). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria, URL http://www.R-project.org (Last viewed October 16, 2014).

Revill, K. P., and Spieler, D. H. (**2012**). "The effect of lexical frequency on spoken word recognition in young and older listeners," Psych. Aging **27**, 80–87.

Robertson, I. H., Ward, T., Ridgeway, V., and Nimmo-Smith, I. (**1996**). "The structure of normal human intelligence: The test of everyday attention," J. Int. Neuropsych. Soc. **6**, 525–534.

Rogers, C. S., Jacoby, L. L., and Sommers, M. S. (**2012**). "Frequent false hearing by older adults: The role of age differences in metacognition," Psych. Aging **27**, 33–45.

Rossi-Katz, J., and Arehart, K. H. (**2009**). "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. **52**, 435–453.

Salthouse, T. A. (**2011**). "Cognitive correlates of cross-sectional differences and longitudinal changes in trail making performance," J. Clin. Exp. Neuropsych. **33**, 242–248.

Salthouse, T. A., Toth, T., Daniels, K., Parks, C., Pak, R., Wolbrette, M., and Hocking, K. J. (**2000**). "Effects of aging on efficiency of task switching in a variant of the Trail Making Test," Neuropsych. **14**, 102–111.

Sommers, M. S. (**1996**). "The structural organization of the mental lexicon and its contribution to age-related declines in spoken-word recognition," Psychol. Aging **11**, 333–341.

Sommers, M. S. (**2000**). Washington University Neighborhood Database. Retrieved from http://neighborhoodsearch.wustl.edu/neighborhood/Home.asp (Last viewed October 16, 2014).

Sommers, M. S., and Danielson, S. M. (**1999**). "Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context," Psych. Aging **14**, 458–472.

Sorqvist, P., Ljungberg, J. K., and Ljung, R. (**2010**). "A sub-process view of working memory capacity: Evidence from effects of speech on prose memory," Memory **18**, 310–326.

Spieler, D. H., and Balota, D. A. (**2000**). "Factors influencing word naming in younger and older adults," Psychol. Aging **15**, 225–231.

Taler, V., Aaron, G. P., Steinmetz, L. G., and Pisoni, D. B. (**2010**). "Lexical neighborhood density effects on spoken word recognition and production in healthy aging," J. Gerontol. B Psychol. Sci. Soc. Sci. **65**, 551–560.

Tun, P. A., O'Kane, G., and Wingfield, A. (**2002**). "Distraction by competing speech in young and older listeners," Psychol. Aging **17**, 453–457.

Tun, P. A., and Wingfield, A. (**1999**). "One voice too many: Adult age differences in language processing with different types of distracting sounds," J. Gerontol. **54B**, P317–P327.

Van Engen, K. J., and Bradlow, A. R. (**2007**). "Sentence recognition in native- and foreign-language multi-talker background noise," J. Acoust. Soc. Am. **121**, 519–526.

Woods, W. S., Kalluri, S., Pentony, S., and Nooraei, N. (**2013**). "Predicting the effect of hearing loss and audibility on amplified speech perception in a multi-talker listening scenario," J. Acoust. Soc. Am. **133**, 4268–4278.

376   J. Acoust. Soc. Am. **138** (1), July 2015

Karen S. Helfer and Alexandra Jesse