Age effects on perceptual organization of speech: Contributions of glimpsing, phonemic restoration, and speech segregation

William J. Bologna,^{a)} Kenneth I. Vaden, Jr., Jayne B. Ahlstrom, and Judy R. Dubno

Department of Otolaryngology–Head and Neck Surgery, Medical University of South Carolina,
135 Rutledge Avenue, MSC 550, Charleston, South Carolina 29425, USA

(Received 4 November 2017; revised 6 June 2018; accepted 8 June 2018; published online 16 July 2018)

In realistic listening environments, speech perception requires grouping together audible fragments of speech, filling in missing information, and segregating the glimpsed target from the background. The purpose of this study was to determine the extent to which age-related difficulties with these tasks can be explained by declines in glimpsing, phonemic restoration, and/or speech segregation. Younger and older adults with normal hearing listened to sentences interrupted with silence or envelope-modulated noise, presented either in quiet or with a competing talker. Older adults were poorer than younger adults at recognizing keywords based on short glimpses but benefited more when envelope-modulated noise filled silent intervals. Recognition declined with a competing talker but this effect did not interact with age. Results of cognitive tasks indicated that faster processing speed and better visual-linguistic closure were predictive of better speech understanding. Taken together, these results suggest that age-related declines in speech recognition may be partially explained by difficulty grouping short glimpses of speech into a coherent message.

© 2018 Acoustical Society of America. <https://doi.org/10.1121/1.5044397>

[DB]

Pages: 267–281

I. INTRODUCTION

In daily life, speech communication occurs in environments where background sounds fluctuate in level and mask portions of the intended message. Numerous factors contribute to the difficulty of this perceptual task and many are affected by age. Relative to younger adults, older adults require a more advantageous signal-to-noise ratio (SNR) in a fluctuating background (e.g., Festen and Plomp, 1990), and they are more likely to mistake speech from a competing talker for the intended message (Helfer and Freyman, 2008; Lee and Humes, 2012). Increases in hearing thresholds can contribute to (or compound) the effects of age on speech recognition, particularly in noise, by reducing the audibility of important speech sounds. However, age-related changes in speech recognition are still observed in older adults with normal hearing and under conditions in which differences in audibility are minimized (Dubno *et al.*, 2002, 2003; Eisenberg *et al.*, 1995; Rajan and Cainer, 2008). Declines in temporal processing (Fitzgibbons and Gordon-Salant, 1996; Gordon-Salant and Fitzgibbons, 1993) and certain cognitive abilities, such as working memory and processing speed (Humes *et al.*, 1994; Humes *et al.*, 2013), have been implicated in this residual age-related difficulty. The combined effects of temporal processing, cognitive abilities, and hearing sensitivity have been shown to account for only about 60% of the variance in speech recognition in noise among older adults (Füllgrabe *et al.*, 2015; Humes *et al.*, 2013),

which leaves a considerable amount of residual variance unexplained.

A. Perceptual organization

One explanation that has received less attention is age-related changes in “perceptual organization,” or the process by which the auditory system interprets acoustic input from multiple sources and creates an internal representation of an auditory scene (Bregman, 1990). At least three interrelated auditory processes may contribute to the perceptual organization of speech in noise: glimpsing, phonemic restoration, and speech segregation. “Glimpsing” refers to the process of identifying audible fragments of speech and connecting them together across gaps of missing information to form a single coherent stream (Cooke, 2006; Moore, 2003). “Phonemic restoration” refers to the process of filling in missing information based on the available acoustic cues, knowledge of the language, and semantic context (Bashford and Warren, 1987; Başkent *et al.*, 2010; Warren, 1970). “Speech segregation” refers to the process by which glimpses from the talker of interest are perceptually separated from other sources of speech, such that attention can be focused on a single target when the background includes other talkers (Brungart, 2001; Shinn-Cunningham, 2008). Whereas each of these processes plays a distinct role to facilitate speech recognition in realistic environments, they are inherently interrelated. For example, phonemic restoration depends on accurate glimpsing and segregation of speech glimpses from competing talkers. As such, the same underlying factors may influence in different ways one or more of these components of perceptual organization. The purpose of this study was to systematically investigate all three processes using specific manipulations to determine the extent

^{a)}Current address: National Center for Rehabilitative Auditory Research, Veterans Affairs Portland Health Care System, Department of Veterans Affairs, 3710 Southwest United States Veterans Hospital Road, P5-NCRR, Portland, OR 97239, USA. Electronic mail: William.Bologna@va.gov

to which speech recognition difficulty in noise among older adults may be explained by declines in glimpsing, phonemic restoration, and/or speech segregation.

Glimpsing has been studied in younger and older adults using various experimental methods. One method is to compare speech recognition in steady-state noise to performance in a gated or modulated noise (see Moore, 1990, for a review). Momentary reductions in the level of the noise allow listeners to connect audible glimpses of speech across gaps of missing information, which results in improved speech recognition compared to performance in a steady-state masker. Whereas older adults typically benefit less from glimpsing as compared to younger adults (Eisenberg *et al.*, 1995; Festen and Plomp, 1990; Takahashi and Bacon, 1992), this difference may be explained by greater effects of forward masking for recognition of sentences in interrupted noise among older adults (Dubno *et al.*, 2002, 2003). Glimpsing can be studied without potential effects of forward masking by replacing portions of a sentence or word with silence (Gilbert *et al.*, 2007; Jin and Nelson, 2006, 2010; Miller and Licklider, 1950; Wang and Humes, 2010). Age-related declines in recognition of interrupted speech are typically observed for slower interruption rates, between 2 and 5 Hz (Saija *et al.*, 2014; Shafiro *et al.*, 2016). These rates approximate the syllabic rate of speech and would be representative of the interruptions imposed by a competing talker. A common finding across these studies is that hearing loss contributes strongly to recognition of glimpsed speech, but residual effects of age persist even in the absence of hearing loss (Dubno *et al.*, 2003; Krull *et al.*, 2013; Molis *et al.*, 2015; Shafiro *et al.*, 2016). These residual effects have been linked to age-related declines in temporal processing (Dubno *et al.*, 2003) and cognitive abilities (Krull *et al.*, 2013; Humes *et al.*, 2013, 2016). Taken together, age-related declines in connecting audible portions of speech across time may contribute to speech recognition difficulty of older adults in realistic listening environments.

Glimpsed speech inherently includes missing information, which must be perceptually restored to facilitate speech recognition. This process was initially described by Warren (1970) and subsequently studied with the phonemic restoration paradigm, where noise bursts are inserted into the silent intervals of interrupted speech (Bashford and Warren, 1987; Başkent *et al.*, 2010; Powers and Wilcox, 1977; Verschuure and Brocaar, 1983). Relatively little is known about the effects of age on phonemic restoration. Başkent *et al.* (2010) showed minimal phonemic restoration in older adults with mild-to-moderate hearing loss, although follow-up work suggested this may result from hearing loss rather than age or baseline speech recognition scores (Başkent, 2010). Two studies found an age-related improvement of phonemic restoration in older adults with normal hearing, based on a 2.5-Hz interruption rate in naturally spoken sentences (Saija *et al.*, 2014) or vocoded sentences (Jaekel *et al.*, 2018). These results were interpreted as evidence that older adults may be better able to use lexical factors (i.e., Baltes, 1993) and/or sentence context (i.e., Pichora-Fuller *et al.*, 1995; Pichora-Fuller, 2008) to facilitate top-down restoration of missing information (Jaekel *et al.*, 2018; Saija *et al.*, 2014).

Whereas the importance of these factors for phonemic restoration is well supported (Benard *et al.*, 2014; Verschuure and Brocaar, 1983; Wang and Humes, 2010), there is less evidence that lexical factors and use of context mediate age-related improvements in phonemic restoration (see Jaekel *et al.*, 2018). Thus, the present study aimed to clarify the source of age-related improvements in phonemic restoration.

Segregation of speech from a competing talker may become more difficult with age (e.g., Helfer and Freyman, 2008; Lee and Humes, 2012; Rajan and Cainer, 2008) and may place additional demands on glimpsing and phonemic restoration. Competing talkers result in a combination of masking effects, including energetic masking that breaks up target speech into a series of glimpses and informational masking that requires segregation and focused attention to identify target speech and ignore distracting speech (Brungart, 2001; Durlach *et al.*, 2003). As such, age-related declines in speech recognition with a competing talker may result from slower speech segregation in older adults than younger adults (Ben-David *et al.*, 2012; Ezzatian *et al.*, 2015), and/or declines in cognitive abilities such as working memory and inhibition of competing speech (Helfer and Freyman, 2008; Pichora-Fuller *et al.*, 1995; Tun *et al.*, 2002). Considering that recognition of glimpsed speech is often poorer for older than younger adults, adding demands for perception and focused attention associated with speech segregation may further erode performance by older adults. Potentially separate or interactive declines in speech segregation and glimpsing were tested in the current study using an experimental design in which glimpsing and speech segregation were assessed simultaneously.

The three interrelated components of perceptual organization (glimpsing, phonemic restoration, and speech segregation) have previously been studied independently. The present study represents a comprehensive assessment of these processes to determine if a shared underlying factor contributes to the effects of age observed in each separate domain. One common interpretation of these aging effects implicates the decline in cognitive abilities, such as processing speed (Salthouse, 2000), working memory (Pichora-Fuller *et al.*, 1995; Pichora-Fuller, 2008), inhibitory control (Tun *et al.*, 2002), and/or linguistic closure (George *et al.*, 2007; Humes *et al.*, 2013, 2016; Krull *et al.*, 2013). Another interpretation (not mutually exclusive) suggests that aging effects may be mediated by lexical factors (Benard *et al.*, 2014; Jaekel *et al.*, 2018; Saija *et al.*, 2014).

Here, we pose an additional interpretation that these components of perceptual organization are more susceptible to envelope distortions for older adults compared to younger adults. Silence-interrupted speech is characterized by discontinuities in the envelope, which may disrupt older adults' ability to use context and other higher-level cues to facilitate speech recognition (e.g., Dubno *et al.*, 2003; Krull *et al.*, 2013; Molis *et al.*, 2015; Shafiro *et al.*, 2016). When silent intervals are filled with noise, the illusory perception of continuity (i.e., Bashford *et al.*, 1988) may enable older adults to more effectively use higher level contextual cues to fill in missing information. In contrast, younger adults may be better able to restore this missing information without the need

for illusory continuity, and thus the addition of noise to silent intervals benefits younger adults less than older adults (e.g., Jaekel *et al.*, 2018; Saija *et al.*, 2014). Thus, the central hypothesis of this study was that perceptual organization of speech among older adults is facilitated by continuous envelope cues, and that effects of cognitive and lexical factors are mediated by continuity of the envelope.

This hypothesis was tested by assessing phonemic restoration in younger and older adults using intervening noise that was modulated by the envelope of the missing speech segment. Previous studies have demonstrated that providing continuous envelope cues over the course of a sentence results in enhanced phonemic restoration relative to unmodulated noise (Bashford *et al.*, 1996; Shinn-Cunningham and Wang, 2008). If the age-related improvement in phonemic restoration is a consequence of greater susceptibility to envelope distortions, then robust phonemic restoration should be observed among older adults when intervening noise provides these missing envelope cues. Alternative hypotheses that cognitive and/or lexical factors contribute to those age effects were explored by performing a cognitive test battery and examining lexical characteristics of keywords from the sentence stimuli.

B. Study design and hypotheses

The current study focused on the extent to which poorer speech recognition in noise reflected age-related declines in (1) glimpsing, (2) phonemic restoration, and (3) speech segregation. Younger and older adults with normal hearing listened to sentences interrupted with either silence or envelope-modulated noise in quiet or with a competing talker. The proportion speech remaining after interruption was manipulated to assess glimpsing. The improvement in recognition with envelope-modulated noise relative to silence was phonemic restoration. The decline in recognition with a competing talker indicated the cost of speech segregation.

Hypotheses predicted several age interactions, such that for older adults compared to younger adults: (1) age $\times$ proportion of speech would indicate poorer glimpsing, (2) age $\times$ interruption type would indicate enhanced phonemic

restoration, (3) age $\times$ background type would indicate poorer speech segregation. The hypothesis that age-related declines in speech recognition are mediated by cognitive abilities predicted interactions between cognitive measures (processing speed, working memory, inhibitory control, and/or linguistic closure) and the design-level factors associated with glimpsing, phonemic restoration, and/or speech segregation. The hypothesis that age-related changes in phonemic restoration are mediated by lexical factors predicted interactions between lexical characteristics of keywords (word frequency, neighborhood density, and phonotactic probability) and interruption type. Thus, the study characterized the relative contributions of these factors and interactions with age to recognition of interrupted speech.

II. METHODS

A. Participants

Two groups were tested, 20 younger adults ranging in age from 18 to 29 yr [mean, 24.7; standard deviation (SD), 2.8] and 20 older adults ranging in age from 63 to 84 yr (mean, 69.9; SD, 5.7). Normal cognitive functioning of older participants was confirmed by a score of 25 or greater on the Mini Mental Status Exam (MMSE; Folstein *et al.*, 1975; Tombaugh and McIntyre, 1992). All participants were native speakers of American English and reported normal or corrected-to-normal vision.

Hearing sensitivity was assessed in all participants by measuring air-conduction thresholds at audiometric frequencies (ANSI, 2010). Hearing threshold criteria for younger participants were defined as thresholds ≤ 25 dB hearing level (HL) for 250–8000 Hz. For older participants, threshold criteria were ≤ 30 dB HL for 250–6000 Hz to ensure that an adequate number of participants could be recruited for testing. Although all participants met hearing threshold criteria, differences in mean thresholds were noted between younger and older participant groups, particularly at higher frequencies (Fig. 1). Subtle declines in hearing sensitivity can result

FIG. 1. (Left) Mean air conduction audiograms for younger adults (open) and older adults (gray), with symbols offset for clarity. Error bars represent ± 1 standard deviation. Frequency ranges of four pure-tone averages (PTAs) are displayed at the bottom. (Right) Mean PTA measures for younger and older adults. Asterisks indicate significant age-group differences based on independent samples *t*-tests with Bonferroni correction (** = $p < 0.001$).

in poorer speech recognition for temporally distorted speech and speech in noise (e.g., [Dubno et al., 2002](#); [Festen and Plomp, 1990](#); [Krull et al., 2013](#)). To address this potential confound several pure-tone averages (PTAs) were calculated to quantify mean hearing sensitivity across different frequency ranges ([Simpson et al., 2013](#)), low-frequency PTA (0.25, 0.5, 1.0 kHz), narrow PTA (0.5, 1.0, 2.0, 4.0 kHz), broad PTA (0.5, 1.0, 2.0, 3.0, 4.0, 6.0, 8.0 kHz), and high-frequency PTA (2.0, 3.0, 4.0, 6.0, 8.0 kHz). Several PTA measures were used so that analyses would be sensitive to potential effects in multiple frequency regions that may not otherwise be captured by a single PTA. Independent samples *t*-tests were performed on PTA measures for younger and older adults using Bonferroni correction for multiple comparisons. Results indicated that all PTA measures were significantly lower for younger adults than older adults [$t_{\text{LFPTA}}(33.18) = 4.11$, $p < 0.001$; $t_{\text{NPTA}}(37.53) = 6.71$, $p < 0.001$; $t_{\text{BPTA}}(35.06) = 8.68$, $p < 0.001$; $t_{\text{HFPTA}}(33.06) = 8.10$, $p < 0.001$]. Potential effects of hearing sensitivity on speech recognition were examined statistically using these PTA measures as predictors of performance in a generalized linear mixed model.

B. Stimuli and apparatus

Speech stimuli were sentences spoken by male and female talkers from the TIMIT corpus ([Garofolo et al., 1993](#)) that were compiled into phonetically balanced lists for the Perceptually Robust English Sentence Test Open-set (PRESTO; [Gilbert et al., 2013](#)). Silence-interrupted and noise-filled versions of the sentences were generated by multiplying the original waveforms by a rectangular gating function and leaving interruptions silent or filling them with envelope-modulated noise. Five gating functions with different duty cycles were selected based on pilot testing to avoid floor and ceiling effects with speech proportions of 0.40, 0.50, 0.55, 0.60, and 0.70. The interruption rate was adjusted as speech proportion varied to maintain fixed 200-ms interruptions [Fig. 2 (left)], which were less than the average word length and have been shown to produce robust phonemic restoration ([Bashford and Warren, 1987](#); [Bashford et al.,](#)

[1988](#)). For noise-filled sentences, envelope-modulated noise was generated by filtering white noise by the long-term average speech spectrum (LTASS) and then modulating the noise by the envelope of the original sentence. The LTASS was measured using an add-overlap method (e.g., [Versfeld et al., 2000](#)) with 16-ms Hanning windows and 50% overlap, measured over 18 concatenated sentences from PRESTO list 1. The envelope was obtained via full-wave rectification, followed by low-pass filtering at 50 Hz with a fifth-order Butterworth filter. Speech and noise were multiplied by inverse gating functions and combined to create envelope-filled sentences [Fig. 2 (right)]. As in previous studies using similar stimuli, speech and noise were combined at 0 dB SNR to maintain an equally scaled modulation of the envelope across speech and noise segments that matched the original uninterrupted sentence ([Bashford et al., 1996](#); [Shinn-Cunningham and Wang, 2008](#)).

Each competing talker sentence was selected from unused opposite-sex sentences in the TIMIT corpus to match the duration of its corresponding target sentence (± 0.5 s). Unprocessed competing talker sentences were mixed with interrupted or envelope-filled target sentences at +3 dB SNR. Pilot testing indicated a floor effect for the most difficult speech proportion (0.40) with a competing talker, so this speech proportion was only tested in quiet. To facilitate identification of the target talker, each competing-talker sentence mixture was preceded by an uninterrupted cue phrase spoken by the target talker in quiet, which was extracted from a standard sentence recorded by each talker in the TIMIT corpus. Similar voice cues have been used to identify the target talker in other multi-talker speech tasks, such as the Coordinate Response Measure ([Brungart, 2001](#)) and the Theo-Victor-Michael sentences ([Helfer and Freyman, 2009](#)). Stimuli were generated with 16-bit resolution at a sampling rate of 16000 Hz prior to data collection. A calibration noise was set to 70 dB sound pressure level (SPL) through the headphones using an acoustic coupler with a Larson Davis model 2559 $\frac{1}{2}$ -inch microphone and a Larson Davis Model 824 sound level meter (Larson Davis, Provo, UT). The level of the

FIG. 2. (Left) Schematic of 5 gating functions used to create interrupted sentences. The initial portion of the sentence was always gated on and 10-ms ramps were used at speech onsets/offsets to avoid transients. (Right) Speech (black) and envelope-modulated noise (gray) were multiplied by inverse gating functions, so the temporal envelope of noise-filled sentences matched the original recording.

sentences was set prior to interruption, such that the overall level of the silence-interrupted sentences was less than 70 dB SPL and varied across sentences based on the interruptions.

C. Procedures

1. Speech measures

Measurement of speech recognition was completed in a sound-attenuating booth. Presentation and keyword scoring was controlled by Token software (Kwon, 2012) with a Lynx Two multichannel audio interface (Lynx, Costa Mesa, CA), Tucker-Davis Technologies programmable attenuator and headphone buffer (PA4 and HB6; Tucker-Davis, Alachua, FL), and Sennheiser HDA 200 headphones (Wedemark, Germany). Stimuli were presented monaurally to the right ear, except for four older participants who only met the hearing criteria in their left ear.

Testing was completed within a single two-hour session, blocked by background (quiet/competing talker) and counterbalanced across participants. Data collection consisted of 1 list of 18 sentences per interruption type (silence-interrupted or envelope-filled) at each speech proportion in a random order, for a total of 10 lists in the quiet background (5 speech proportions) and 8 lists in the competing talker background (4 speech proportions). Participants completed two practice lists designed for familiarization with the stimuli and response paradigm, based on pilot data that suggested two lists were sufficient to produce stable performance.

For testing in quiet, participants were instructed to repeat each sentence, guessing if they were not able to understand all the words in the sentence. For testing with the competing talker, participants were instructed to use the cue phrase to identify the target talker and repeat the interrupted sentence spoken by that talker. Participant responses were scored online by the experimenter using a strict scoring rule (i.e., no additional or missing suffixes) and recorded using a Realistic Highball Dynamic 33-984C microphone and Adobe Audition for offline review and confirmation of scoring, as needed.

2. Item-level factors

Lexical characteristics for the keywords were obtained using The Irvine Phonotactic Online Dictionary (IPhOD version 2.0; Vaden *et al.*, 2009). The selected measures included context-weighted word frequency, neighborhood density, and biphone probability. Word frequency measures the count of each keyword in a film subtitle database, adjusted for the number of different films in which the word occurs (i.e., context-weighted, Brysbaert and New, 2009). Neighborhood density measures were based on the number of words in the database that differed from the keyword by only one phoneme (i.e., Cluff and Luce, 1990). Biphone probability was calculated based on the likelihood of phoneme sequences within the keyword (i.e., Vitevitch *et al.*, 1999). For keywords with multiple pronunciation entries in the IPhOD, the entry was selected that most closely matched the pronunciation in the recording.

Three additional item-level factors were evaluated for each keyword; talker sex, sentence length, and keyword position. Each PRESTO list contains nine male talkers and nine female talkers. Sentence length was defined as the total number of keywords in the sentence, which ranged from three to six across sentences. Keyword position was calculated as the keyword’s serial order position among keywords in the sentence, divided by the total number of keywords in the sentence. Quantified this way, keyword position represents the proportional position of the keyword within the sentence, so that position effects could be evaluated across sentences with different lengths.

3. Cognitive measures

In a separate two-hour test session, participants completed a battery of cognitive measures that assessed processing speed (“Connections,” Salthouse *et al.*, 2000; “Purdue Pegboard,” Tiffin and Asher, 1948), working memory capacity (“Reading Span,” Daneman and Carpenter, 1980; Rönnerberg, 1990), inhibitory control (“Stroop,” Stroop, 1935; Trenerry *et al.*, 1989), and visual linguistic closure (text reception threshold, “TRT,” Zekveld *et al.*, 2007); descriptions of these measures and scoring can be found in Table I. These tests represented a broad selection of cognitive

TABLE I. Task and scoring descriptions for five measures included in the cognitive test battery.

Cognitive measure	Domain	Task	Scoring
Connections (Salthouse <i>et al.</i> , 2000)	Processing speed	Trace a line through a series of numbered and lettered circles in an alternating order	Mean number of connected circles across four trials
Purdue Pegboard (Tiffin and Asher, 1948)	Processing speed	Insert pieces into a board using both hands in an alternating fashion	Mean number of pieces inserted into the board across two trials
Reading Span (Daneman and Carpenter, 1980; Rönnerberg, 1990)	Working memory	Recall either the first or last word in sets of visually presented sentences	Percentage of correctly recalled words across 54 sentences
Stroop (Stroop, 1935; Trenerry <i>et al.</i> , 1989)	Inhibitory control	Read color words printed in incongruent colored ink, or name the color of the ink and ignore the word itself	Difference in reaction time (seconds/word) for word reading and color naming, normalized for the individual’s word reading reaction time
TRT (Zekveld <i>et al.</i> , 2007)	Visual linguistic closure	Read sentences presented behind an array of vertical black bars; the width of the bars varies adaptively based on performance	The average proportion of visible text on the last eight sentences in a run; final threshold was a mean of four runs

measures that have been implicated in the effects of age on speech recognition or phonemic restoration in previous studies (Benard *et al.*, 2014; Füllgrabe *et al.*, 2015; Humes *et al.*, 2013; Jaekel *et al.*, 2018). All participants completed the cognitive tests, except for one younger participant who could not perform the Stroop test due to a reported blue-green colorblindness.

D. Statistical approach

A logistic regression analysis of keyword recognition was implemented in R (R Core Team, 2016) using a generalized linear mixed model (GLMM; R-package: lme4; Bates *et al.*, 2015). The GLMM specified recognition of an individual keyword as the dependent variable ($W=0$ for incorrect keywords, 1 for correct keywords) and estimated separate β coefficients for each independent variable included in the model. The GLMM analysis was performed across all keywords independently (i.e., without nesting at the sentence level) for all participants with the following design-level factors: speech proportion (Prop), background (Bg), interruption type (Int), age group (Age), and random subject effects (Subj). All factors were scaled with mean=0 and SD=1 to allow for direct comparison of standard estimates. The random subject effects term included Prop, Bg, and Int, such that the model could adjust for additive subject-level differences in those effects (i.e., Clark, 1973). This GLMM can be expressed as $W = Prop + Bg + Int + Age + (Prop + Bg + Int | Subj) + error$.

Several additional factors were evaluated using model testing (Hofmann, 1997) to determine whether the fit of the GLMM significantly improved with the addition of a given factor. The tested factors included interactions, item-level factors such as talker sex,¹ sentence length, keyword position, and lexical characteristics of the keywords, and subject-level factors such as years of education, PTA measures, and cognitive test scores. A combination of stepwise factor addition and elimination was performed to optimize model fit with all significant factors and interactions.

Correlations between fixed effects that exceeded 0.3 (i.e., greater than negligible correlation; Hinkle *et al.*, 2003) were addressed by residualizing and rescaling predictor variables to reduce collinearity. Although a single model was used to test all factors, the results for design-level, item-level, and subject-level factors are reported in separate sections for clarity. Significant interactions were interpreted using separate *post hoc* models with split factors to describe the effect of a given factor across the two levels of its interacting factor. Standard estimates, standard error, and z -statistics for each factor, including split factors from *post hoc* models, can be found in Table II.

III. RESULTS

A. Design-level factors

Keyword recognition, transformed into rationalized acrisine units (rau; Studebaker, 1985), is plotted in the top panels of Fig. 3, with phonemic restoration plotted in the middle

panels, and the effect of age plotted in the bottom panels. Model testing confirmed significant improvements in the fit of the model with the addition of each design-level factor and several two-way interactions: speech proportion, background, interruption type, age group, proportion $\times$ background, proportion $\times$ age, proportion $\times$ interruption, interruption $\times$ age, and background $\times$ interruption (all $\chi^2 \geq 9.18$, $p < 0.01$). None of the additional two-, three-, or four-way interaction terms significantly affected model fit [all $\chi^2 < 3.80$, non-significant (ns) in all cases]. Modeling results indicated that younger adults significantly outperformed older adults (Fig. 3, black vs gray in top panels, bottom panels; $\beta_{Age} = -0.29$; $z = -5.23$; $p < 0.001$). Keyword recognition improved significantly with increasing proportion of speech (Fig. 3, top panels, slope of lines; $\beta_{Prop} = 1.03$; $z = 58.41$; $p < 0.001$), consistent with the assumption that the amount of glimpsed speech was a strong predictor of performance. Keyword recognition also improved significantly with the addition of envelope-modulated noise to silent intervals (Fig. 3, middle panels; $\beta_{Int} = 0.23$; $z = 14.51$; $p < 0.001$), indicating that phonemic restoration facilitated recognition of envelope-filled sentences. Keyword recognition declined significantly with the addition of a competing talker (Fig. 3, right vs left panels; $\beta_{Bg} = -0.75$; $z = -39.11$; $p < 0.001$), indicating an increase in task difficulty associated with segregating speech glimpses from a single talker background. *Post hoc* models with split factors were constructed to test hypotheses and explore significant interaction terms.

The hypothesis of an age-related decline in glimpsing was supported by an interaction between age group and proportion of speech ($\beta_{Prop \times Age} = 0.07$; $z = 4.43$; $p < 0.001$). *Post hoc* modeling revealed that speech proportion was a stronger predictor of keyword recognition among older adults ($\beta_{Prop_O} = 1.06$) than younger adults ($\beta_{Prop_Y} = 0.92$). This finding indicates that older adults relied more on the total amount of glimpsed speech information, whereas younger adults were better able to connect short glimpses of speech across time. As a result, age effects were largest for sentences with the smallest speech proportion (Fig. 3, bottom panels).

The hypothesis of an age-related improvement in phonemic restoration was supported by an interaction between age group and interruption type ($\beta_{Int \times Age} = 0.06$; $z = 3.61$; $p < 0.001$). *Post hoc* modeling indicated that the effect of interruption type was larger among older adults ($\beta_{Int_O} = 0.26$) than younger adults ($\beta_{Int_Y} = 0.16$). The interaction between age and interruption type indicated a larger phonemic restoration effect among older adults compared to younger adults (Fig. 3, middle panels), consistent with an age-related increase in susceptibility to envelope distortions in silence-interrupted speech.

The hypothesis of an age-related decline in speech segregation was not supported, as the interaction between age and background did not significantly improve model fit ($\chi^2 = 1.73$, ns). This suggests that the addition of a competing talker increased the overall difficulty of the task, but the decrease in scores was similar for younger and older adults.

TABLE II. GLMM standard estimates, standard error, and z-statistics are displayed for each significant fixed effect and interaction term in the order they appear in the text. Factors included in the random effects term are marked with “†.” Each Interaction was explored with a separate *post hoc* model with split factors, and statistical results are indented below interactions. Asterisks indicate significance levels for z-statistics (*** <0.001; ** <0.01; * <0.05).

Factor	Coding scheme	Standard estimate (β)	Standard error	z-statistic
Age group	Younger = -1; Older = 1	-0.29	0.06	-5.23***
Proportion [†]	Normalized for mean = 0, SD = 1	-1.03	0.02	58.41***
Interruption type [†]	Silence = -1; Noise = 1	-0.23	0.02	14.51***
Background [†]	Quiet = -1; Competing Talker = 1	-0.75	0.02	-39.11***
Proportion × background		-0.10	0.01	6.62***
Proportion, competing talker		-1.12	0.03	38.57***
Proportion, quiet		-0.93	0.02	42.33***
Proportion × age group		-0.07	0.02	4.43***
Proportion, older		-1.06	0.02	42.79***
Proportion, younger		-0.92	0.02	39.13***
Proportion × interruption type		-0.07	0.01	-5.42***
Proportion, silence interrupted		-1.06	0.03	42.14***
Proportion, envelope filled		-0.92	0.02	38.44***
Interruption type × age group		-0.06	0.02	3.61***
Interruption type, older		-0.26	0.02	11.95***
Interruption type, younger		-0.16	0.02	7.59***
Background × interruption type		-0.03	0.01	3.03***
Background, silence interrupted		-0.73	0.02	-34.79***
Background, envelope filled		-0.68	0.02	-33.58***
Word frequency		-0.28	0.01	27.72***
Biphone probability		-0.03	0.01	2.90***
Neighborhood density		-0.15	0.01	-14.00***
Word frequency × interruption type		-0.02	0.01	-2.32***
Interruption type, uncommon words		-0.25	0.01	18.64***
Interruption type, common words		-0.20	0.02	11.28***
Sentence length		-0.13	0.01	-12.45***
Keyword position		-0.17	0.01	16.38***
Keyword position × age group		-0.04	0.01	-3.75***
Keyword position, younger		-0.19	0.01	13.64***
Keyword position, older		-0.11	0.01	7.09***
Sentence length × background		-0.09	0.01	-7.99***
Sentence length, competing talker		-0.22	0.02	-12.45***
Sentence length, quiet		-0.05	0.01	-3.89***
Keyword position × background		-0.11	0.01	10.12***
Keyword position, competing talker		-0.28	0.02	17.10***
Keyword position, quiet		-0.07	0.01	5.28***
Talker sex × age group		-0.08	0.01	-7.95***
Talker sex, younger	Female = 1; Male = -1	-0.07	0.01	5.29***
Talker sex, older		-0.07	0.02	-4.70***
Talker sex × background		-0.03	0.01	3.03***
Talker sex, quiet		-0.03	0.01	-2.54***
Talker sex, competing talker		-0.04	0.02	2.18***
Talker sex × background × age group		-0.04	0.01	-4.00***
Talker sex, quiet, younger		-0.01	0.02	0.53 ns
Talker sex, quiet, older		-0.07	0.02	-3.91***
Talker sex, competing talker, younger		-0.15	0.02	7.02***
Talker sex, competing talker, older		-0.09	0.02	-3.84***
Education level		-0.10	0.04	2.35***
Connections		-0.13	0.05	2.80***
TRT ^a		-0.13	0.05	2.60***
TRT × interruption type		-0.06	0.01	-4.19***
TRT, silence interrupted		-0.19	0.05	3.40***
TRT, envelope filled		-0.07	0.05	1.44 ns
Interruption type, poorer TRT		-0.27	0.02	12.40***
Interruption type, better TRT		-0.21	0.02	9.53***

^aInitial modeling results revealed significant collinearity between TRT and years of education, and so TRT was residualized for both age and years of education.

FIG. 3. (Top row) Keyword recognition (rau) plotted as a function of proportion of speech with interruption type and age as parameters. Sentences in quiet are plotted on the left and with a competing talker on the right. Error bars represent ± 1 standard error. (Middle row) Phonemic restoration, quantified as the difference in keyword recognition between envelope-filled and silence-interrupted sentences, is plotted as a function of proportion of speech with age as the parameter. (Bottom row) Effect of age, quantified as the difference in mean keyword recognition between younger and older groups, is plotted as a function of proportion of speech with interruption type as the parameter.

Several other interactions significantly improved the fit of the model, and these interactions were also explored with *post hoc* models. Proportion of speech significantly interacted with background ($\beta_{\text{Prop} \times \text{Bg}} = 0.10$; $z = 6.62$; $p < 0.001$), and this interaction remained significant after removing the quiet data at 0.40 proportion of speech to account for unequal testing conditions. *Post hoc* modeling revealed that speech proportion was a stronger predictor of keyword recognition with the competing talker ($\beta_{\text{Prop-CT}} = 1.12$) than in quiet ($\beta_{\text{Prop-Q}} = 0.93$). This result suggests that connecting glimpses of speech across time was more difficult with the competing talker than in quiet. Proportion of speech also significantly interacted with interruption type ($\beta_{\text{Prop} \times \text{Int}} = -0.07$; $z = -5.42$; $p < 0.001$). *Post hoc* modeling revealed that proportion was a stronger predictor of keyword recognition for silence-interrupted sentences ($\beta_{\text{Prop-SInt}} = 1.06$) than envelope-filled sentences ($\beta_{\text{Prop-Env}} = 0.92$). This result suggests that, with the addition of envelope-modulated noise, listeners were better able to use the available glimpses of speech to recognize the sentence. In the absence of envelope cues, keyword recognition depended more on the total amount of glimpsed speech information. Finally, background significantly interacted with interruption type ($\beta_{\text{Bg} \times \text{Int}} = 0.03$; $z = 3.03$; $p < 0.01$). *Post hoc* modeling indicated that the competing talker had a more detrimental effect on keyword recognition for silence-interrupted ($\beta_{\text{Bg-SInt}} = -0.73$) than envelope-filled sentences ($\beta_{\text{Bg-Env}} = -0.68$). This suggests that the addition of envelope-modulated noise to silent intervals may have helped listeners segregate the glimpses of target speech from the competing talker.

B. Item-level factors

1. Lexical characteristics

Initial modeling results indicated that the lexical characteristics tested [word frequency (WF), neighborhood density (ND), and biphone probability (BP)] were correlated and the strongest predictor among the three factors was word frequency. To address collinearity, neighborhood density and biphone probability were residualized with respect to word frequency. Model testing confirmed all three lexical characteristics and an interaction between word frequency and interruption type significantly improved model fit (all $\chi^2 \geq 5.42$, $p < 0.05$). Correct recognition was significantly more likely for more commonly used keywords ($\beta_{\text{WF}} = 0.28$; $z = 27.72$; $p < 0.001$), keywords from sparser neighborhoods ($\beta_{\text{ND}} = -0.15$; $z = -14.00$; $p < 0.001$), and keywords with more common phoneme sequences ($\beta_{\text{BP}} = 0.03$; $z = 2.90$; $p < 0.01$). The significant interaction between word frequency and interruption type ($\beta_{\text{Int} \times \text{WF}} = -0.02$; $z = -2.32$; $p < 0.05$) was explored with a *post hoc* model, which showed larger phonemic restoration effects for less common keywords ($\beta_{\text{Int-loWF}} = 0.25$) than for more common keywords ($\beta_{\text{Int-hiWF}} = 0.20$). This suggests that phonemic restoration was most beneficial for the harder words in the sentence, and recognition of these difficult keywords may have been more disrupted by envelope discontinuities than the easier keywords. However, the hypothesis that this lexical effect on phonemic restoration would mediate the observed effects of age was not supported, as additional interactions between lexical characteristics and age were tested but did not

significantly improve the fit of the model ($\chi^2 \leq 1.67$, ns in all cases).

2. Sentence length and keyword position

Model testing confirmed significant improvements in model fit with the addition of sentence length (SL), keyword position (Pos), and three interaction terms (SL $\times$ Bg, Pos $\times$ Age, Pos $\times$ Bg; all $\chi^2 \geq 14.15$, $p < 0.001$). As illustrated in the left panels of Fig. 4, keyword recognition was poorer in sentences with more than three keywords ($\beta_{SL} = -0.13$; $z = -12.45$; $p < 0.001$). The significant interaction between sentence length and background (Fig. 4, bottom-left panel; $\beta_{SL \times Bg} = -0.09$; $z = -7.99$; $p < 0.001$) indicated more precipitous declines in recognition with increasing sentence length with a competing talker ($\beta_{SL_CT} = -0.22$) than in quiet ($\beta_{SL_Q} = -0.05$). This finding suggests that the additional demands associated with speech segregation compounded the difficulty associated with recognition of keywords in longer sentences. The right panels of Fig. 4 plot keyword recognition as a function of keyword position, which was scaled for the total number of keywords in the sentence prior to model testing. Trendlines in Fig. 4 are based on a quadratic fit of the rau data and reveal U-shaped curves, characterized by better keyword recognition for initial and final keywords relative to middle keywords (i.e., recency and primacy effects; Murdock, 1968). Model testing indicated that the overall linear effect of keyword recognition was better recognition for keywords later in the sentence

($\beta_{Pos} = 0.17$; $z = 16.38$; $p < 0.001$), indicating effects of recency were stronger than primacy. The significant interaction between keyword position and age (Fig. 4, top-right panel; $\beta_{Pos \times Age} = -0.04$; $z = -3.75$; $p < 0.001$) indicated that keyword recognition improved more over the course of the sentence for younger adults ($\beta_{Pos_Y} = 0.19$) than for older adults ($\beta_{Pos_O} = 0.11$). The significant interaction between keyword position and background (Fig. 4, bottom-right panel; $\beta_{Pos \times Bg} = 0.11$; $z = 10.12$; $p < 0.001$) indicated greater improvements in recognition over the course of the sentence with a competing talker ($\beta_{Pos_CT} = 0.28$) than in quiet ($\beta_{Pos_Q} = 0.07$). This finding is consistent with the hypothesis that speech segregation develops slowly over time (Ezzatian *et al.*, 2015), resulting in better recognition of keywords occurring later in the sentence. However, the three-way interaction between keyword position, background, and age did not significantly improve the model ($\chi^2 = 0.45$, ns). This finding is inconsistent with results from Ezzatian *et al.* (2015) and suggests that the time course of speech segregation was similar for younger and older adults.

C. Subject-level factors

1. Years of education

Younger and older participants had similar years of education: 16.8 yr (SD = 2.3) for younger participants and 16.4 yr (SD = 2.4) for older participants [$t(37.88) = 0.54$, ns]. Model testing confirmed a significant improvement in model fit related to years of education ($\chi^2 = 5.16$; $p < 0.05$),

FIG. 4. Keyword recognition (rau) plotted as a function of sentence length (left panels) and keyword position (right panels), with age as the parameter (top panels) and background as the parameter (bottom panels). (Left panels) Sentence length was defined as the number of keywords in the sentence. Error bars represent ± 1 standard error. (Right panels) Keyword position was defined as the keyword's serial order position among keywords in the sentence divided by the total number of keywords in the sentence. Trendlines are based on a quadratic fit of rau data.

such that keyword recognition increased significantly with increasing years of education ($\beta_{\text{Edu}} = 0.10$; $z = 2.35$; $p < 0.05$).

2. Hearing sensitivity

Effects of hearing sensitivity were evaluated based on four PTA measures. Group differences resulted in collinearity of effects when PTA and age group were included in the GLMM. To determine if differences in hearing sensitivity were driving the observed effects of age, models were generated in which the age factor was replaced with each of the four PTA measures. Comparisons of β values across these models indicated that age was a better predictor of keyword recognition than any measure of PTA ($\beta_{\text{Age}} = -0.29$; $\beta_{\text{NPTA}} = -0.24$; $\beta_{\text{BPTA}} = -0.26$; $\beta_{\text{LFPTA}} = -0.23$; $\beta_{\text{HFPTA}} = -0.26$). Next, each PTA measure was residualized for age group differences and added to the original GLMM. None of the residualized factors significantly improved model fit ($\chi^2_{\text{NPTA}} = 1.12$; $\chi^2_{\text{BPTA}} = 1.58$; $\chi^2_{\text{LFPTA}} = 1.68$; $\chi^2_{\text{HFPTA}} = 1.83$; ns), indicating that differences in hearing sensitivity among younger and older adults were not predictive of keyword recognition, beyond effects that could be attributed to age.

3. Cognitive measures

Cognitive test results for younger and older adults were transformed into z -scores and are displayed in Fig. 5. Independent samples t -tests were performed on each cognitive measure for younger and older adults using Bonferroni correction for multiple comparisons. Results indicated younger adults outperformed older adults for all cognitive measures [$t_{\text{Connections}}(37.10) = 5.21$, $p < 0.001$; $t_{\text{Pegboard}}(37.82) = 4.91$, $p < 0.001$; $t_{\text{RSpan}}(35.26) = 3.26$, $p < 0.01$; $t_{\text{Stroop}}(32.47) = 3.60$, $p < 0.01$; $t_{\text{TRT}}(26.88) = 4.41$, $p < 0.001$].

Group differences resulted in collinearity of effects when cognitive factors and age group were included in the GLMM. To address this issue, each cognitive measure was residualized for age group differences and added to the

GLMM. Significant improvements in model fit were obtained with the addition of a measure of processing speed (Connections), a measure of visual linguistic closure (TRT), and an interaction between TRT and interruption type (all $\chi^2 \geq 6.31$; $p < 0.05$). Results indicated that keyword recognition improved with faster speed of processing ($\beta_{\text{Connections}} = 0.13$; $z = 2.80$; $p < 0.01$) and better linguistic closure ($\beta_{\text{TRT}} = 0.13$; $z = 2.60$; $p < 0.01$). The significant interaction between TRT and interruption type ($\beta_{\text{TRT} \times \text{Int}} = -0.06$; $z = -4.19$; $p < 0.001$) was explored with a *post hoc* model, which indicated that TRT significantly predicted keyword recognition for silence-interrupted sentences ($\beta_{\text{TRT_SInt}} = 0.19$; $z = 3.40$; $p < 0.001$) but not for envelope-filled sentences ($\beta_{\text{TRT_Env}} = 0.07$; $z = 1.44$; ns). Other cognitive factors (Purdue Pegboard, Reading Span, Stroop) and interactions were explored, but model testing revealed that these additional factors did not improve the fit of the model ($\chi^2 < 1.40$; ns in all cases). Thus, the hypothesis that cognitive abilities contribute to age-related declines in speech recognition was only partially supported by the data.

IV. DISCUSSION

This study investigated the contributions of glimpsing, phonemic restoration, and speech segregation to recognition of interrupted speech in younger and older adults with normal hearing. As expected, older adults performed poorer than younger adults across all speech conditions. Similar age-related differences in recognition of interrupted speech, even in the absence of significant threshold elevation, have been observed in several other studies (Gordon-Salant and Fitzgibbons, 1993; Krull *et al.*, 2013; Shafiro *et al.*, 2015). Older adults in the current study had only slightly elevated mean thresholds as compared to the younger adults and variance in hearing sensitivity was less predictive of keyword recognition than age. Several factors unrelated to pure-tone thresholds, namely glimpsing, speech segregation, and phonemic restoration, were found to significantly contribute to keyword recognition and to the effects of age.

A. Glimpse duration and proportion of speech

The ability to use glimpsed speech information to recognize the sentence was operationalized here as speech proportion. Sentences were processed such that equal duration segments were removed, and the duration of glimpses increased with the proportion of speech. Increasing speech proportion also affected the interruption rate, resulting in fewer overall interruptions, which may have contributed to the improvement in speech recognition with increasing proportion of speech. As such, the increase in speech recognition with increasing proportion of speech reflects the extent to which longer glimpses with fewer interruptions facilitated better keyword recognition. Although these factors covaried in the present investigation, previous studies in which interruption rate, glimpse duration, and speech proportion were varied orthogonally indicated that overall speech proportion is the best predictor of recognition of interrupted speech

FIG. 5. Cognitive test results, transformed to z -scores for younger (white) and older (gray) participants. Asterisks indicate significant age-group differences based on independent samples t -tests with Bonferroni correction (* = $p < 0.01$; ** = $p < 0.001$).

(Gilbert *et al.*, 2007; Kidd and Humes, 2012; Wang and Humes, 2010).

The ability to use glimpsed speech information is hypothesized to decline with age, which is supported by age-related declines in overall performance (Gordon-Salant and Fitzgibbons, 1993; Krull *et al.*, 2013; Shafiro *et al.*, 2015). Further support can be drawn from the observed interaction between age and speech proportion. This interaction indicated that older adults were less able to make efficient use of limited speech information to recognize keywords than younger adults. The discontinuities of the envelope imposed by interruptions were more disruptive to performance for older adults than younger adults, and sentences with more interruptions demonstrated greater effects of age. These results are consistent with the hypothesis of an age-related decline in glimpsing, such that older adults are more susceptible to envelope distortions and less adept at connecting short glimpses of speech together across time into a single coherent message.

Proportion was also observed to interact with several other factors in the current study. These interactions suggested a multiplicative effect of factors that challenge performance (competing talker, older age, silent interruptions), which combine and result in stronger speech proportion effects. That is, factors that increased the difficulty of the task also increased the contributions of glimpse duration and the available speech information to overall sentence recognition. The competing talker likely resulted in some amount of energetic masking that reduced the “effective” proportion of speech. That effect, coupled with the need to segregate glimpses from the competing talker, resulted in greater reliance on available speech information. Likewise, silent interruptions removed a cue provided by envelope-modulated noise, preventing listeners from effectively using speech glimpses and connecting them over time. The improvements in speech recognition that occurred when interruptions were filled with envelope-modulated noise (phonemic restoration) effectively minimized the importance of the available glimpses, allowing both younger and older adults to recognize more keywords with shorter glimpses of speech.

B. Effects of age on phonemic restoration

Older adults demonstrated a larger benefit than younger adults when silent interruptions were filled with envelope-modulated noise. Similar findings have been observed in two recent studies (Jaekel *et al.*, 2018; Saija *et al.*, 2014). In these studies, phonemic restoration was elicited using steady-state noise and results indicated enhanced restoration for older adults in sparse subsets of conditions. In the current study, phonemic restoration was elicited with noise that carried missing envelope cues and restoration was greater for older than younger adults across all test conditions. Taken together, these results suggest that the observed “enhancement” of phonemic restoration with advancing age may reflect greater sensitivity to envelope distortions among older adults.

Under conditions in which the speech envelope is severely distorted (i.e., silence-interrupted speech), effects of

age are robust and consistent across studies (Gordon-Salant and Fitzgibbons, 1993; Jaekel *et al.*, 2018; Krull *et al.*, 2013; Saija *et al.*, 2014; Shafiro *et al.*, 2015). When silent intervals are filled with steady-state noise, envelope modulations are partially restored, leading to improved continuity and better speech recognition (Bashford *et al.*, 1988). These incremental improvements appear to reflect a greater dependence on envelope cues for older adult speech recognition, compared to younger adults (Jaekel *et al.*, 2018; Saija *et al.*, 2014). In the current study, envelope-modulated noise was used to preserve envelope continuity and facilitate speech recognition for the older adults, which sharply contrasted with the discontinuities of silent-interruptions. This pattern of results provides support for the interpretation that a primary effect of aging observed across these studies is a greater reliance on continuous envelope cues and a poorer ability to resolve temporal distortions in the envelope.

An alternative hypothesis suggested that older adults may be better able to compensate for age-related declines in speech recognition by relying on lexical factors and their vocabulary and/or by exerting additional cognitive effort to fill in missing information. To test these hypotheses, the contributions of lexical factors of keywords and cognitive abilities were assessed with item- and subject-level factors. Analyses of lexical factors of keywords and subject-level differences in education did not support the hypothesis that language abilities contribute to enhanced phonemic restoration. Item-level variation in word frequency, biphone probability, and neighborhood density were each predictive of keyword recognition, similar to results of Wang and Humes (2010). However, only word frequency affected phonemic restoration; this effect was relatively small and no associations were observed between any of the lexical factors and age. Whereas vocabulary or language abilities of participants were not directly assessed, participants with more years of education likely had larger vocabularies and better language abilities than participants with fewer years of education (Verhaeghen, 2003). Although performance improved with increasing years of education, younger and older participants did not differ in years of education, and this factor did not influence the improvement associated with phonemic restoration. As such, these results provide little support for the hypothesis that stronger language abilities of older adults contribute to enhanced phonemic restoration.

The contributions of cognitive abilities to phonemic restoration were evaluated with a battery of cognitive tests, including the TRT, which assessed the ability to make use of partial linguistic information in the visual modality. Better performance on the TRT predicted better recognition of silence-interrupted sentences, in agreement with results from Krull *et al.* (2013). The task of reading a visually obscured sentence is intuitively similar to recognition of interrupted sentences and the association between these two measures suggests a common amodal ability to make use of partial linguistic information (Krull *et al.*, 2013). However, older participants performed poorer on the TRT overall and participants with better TRT scores demonstrated *less* phonemic restoration than those with poorer TRT scores. These findings remained consistent even after TRT scores were

residualized for differences in age and years of education. This suggests that greater cognitive effort and use of linguistic abilities to fill in missing information are not likely to explain increased phonemic restoration among older adults. If anything, efforts to fill in missing information are likely to improve recognition of silence-interrupted speech, thereby reducing the amount of improvement that is observed with phonemic restoration. Other cognitive measures were not found to predict keyword recognition, with the exception of the Connections test. Results of the Connections test indicated that faster processing speed contributed to better keyword recognition but did not influence the degree of improvement with phonemic restoration.

The observed age-related increases in phonemic restoration may result from declines in temporal resolution and glimpsing. In this study, older adults were particularly poor at recognizing short glimpses of speech and connecting them across time. Glimpsing, in this context, may involve a form of template matching in which the internal representation of the glimpsed signal is compared to several possible words that may match the stimulus (i.e., [Srinivasan and Wang, 2005](#)). As such, a critical ability for recognition of glimpsed speech may be maintaining an accurate internal representation of the glimpses and their relative position across time. To the extent that the listener retains the precise temporal relationships between glimpses, the ability to match their internal representation of the signal to the appropriate speech template would be optimized. Age-related declines in temporal resolution are known to compromise the ability to maintain temporal relationships by older adults ([Fitzgibbons and Gordon-Salant, 1995, 1996, 2001](#)) and particular difficulty has been noted among older adults for recognition of silent temporal intervals ([Fitzgibbons and Gordon-Salant, 1987, 1994; Gordon-Salant et al., 2006](#)). As such, age-related declines in temporal resolution may be particularly disruptive to recognition of silence-interrupted speech. Inserting noise into silent intervals may provide an auditory scaffold that helps older adults maintain accurate temporal representations of the glimpses across the length of the sentence. Therefore, enhanced phonemic restoration observed for older adults may reflect the combined benefits of phonemic restoration and improved temporal representations of the glimpses relative to silence-interrupted speech.

C. Perceptual “cost” of speech segregation

Speech recognition was substantially poorer with the competing talker compared to quiet, which reflects the combined effects of energetic masking and increased perceptual demand associated with speech segregation. Although the effects of energetic masking cannot be ignored, the masking effects of a single competing talker are thought to predominantly reflect the costs associated with speech segregation and selection of the appropriate target talker (i.e., a form of “informational masking”; [Brungart, 2001; Kidd et al., 2005](#)). In addition, the two item-level factors that interacted with background, keyword position and sentence length, both related to the time course of the sentence. These results

revealed that with the competing talker, longer sentences were generally more difficult, but listeners were more likely to correctly identify keywords in the later part of the sentence. Although effects of keyword position and sentence length were significant in the quiet background, these effects were substantially larger with a competing talker. Because energetic masking would be equivalent for sentences of any length and remain essentially stable throughout the sentence, the observed interactions among keyword position, sentence length, and background support the interpretation that the effect of the competing talker was an increase in task demands associated with speech segregation.

Age-related declines in speech recognition are also hypothesized to result from difficulty segregating target speech from a competing talker, based on poorer speech recognition by older adults compared to younger adults in competing talker backgrounds (e.g., [Helfer and Freyman, 2008; Lee and Humes, 2012](#)) and a prolonged time course of speech segregation for older adults ([Ben-David et al., 2012; Ezzatian et al., 2015](#)). As such, the decline in performance with a competing talker was expected to be greater for older adults than younger adults; similarly, the improvement in performance over the course of the sentence was expected to be greater for older adults than younger adults. Neither of these predictions was supported by the data. Recognition of silence-interrupted speech is a difficult task, particularly for older adults, and the addition of a competing talker may have resulted in floor performance for some older participants. As a result, effects of a competing talker among older adults may have been underestimated in this experiment. Moreover, the effect of keyword position observed here may reflect factors other than speech segregation. For example, the benefits associated with context generally improve over the course of a sentence ([Cole and Perfetti, 1980; Kidd and Humes, 2012](#)) and these context-dependent effects may have been more beneficial in the more difficult competing talker background ([Dirks et al., 1986](#)).

Poorer performance with a competing talker for longer sentences and for keywords occurring earlier in the sentence is suggestive of effects related to working memory capacity. Sentences with more keywords likely placed greater demands on working memory, which may have contributed to the overall negative effect of sentence length on performance. Keywords occurring later in the sentence likely placed less demands on working memory due to recency effects, which improved the likelihood of correct recognition for later occurring keywords ([Surprenant et al., 1993](#)). Larger effects with a competing talker are in line with previous research indicating that poorer intelligibility increases the demands on working memory ([DeCaro et al., 2016](#)). Participants with higher working memory capacity were expected to leverage their strong cognitive abilities to facilitate better speech recognition, particularly in the more difficult competing talker background. However, performance on the Reading Span (a visual measure of working memory) did not predict keyword recognition or influence the effect of sentence length on performance. One possibility is that the visual measure of working memory capacity was not sensitive to the effects of working memory on the auditory speech

recognition task. Stronger associations may have been observed with an auditory measure of working memory capacity (e.g., [Gordon-Salant and Cole, 2016](#)).

V. CONCLUSIONS

This study used an interrupted speech paradigm to determine the extent to which age-related declines in speech recognition could be explained by difficulty with glimpsing, phonemic restoration, and/or speech segregation. Poorer speech recognition was observed for older adults compared to younger adults and this decline was better predicted by age than by variance in hearing sensitivity. Poorer speech recognition among older adults was associated with declines in glimpsing, particularly for silence-interrupted speech. Phonemic restoration was greater in older adults than younger adults, but this finding may be attributable to age-related declines in glimpsing and greater sensitivity to envelope distortions among older adults. The additional cues available in envelope-filled sentences compared to silence-interruptions may have provided an auditory scaffold that helped older adults address these difficulties. Subject-level differences in processing speed and linguistic closure were predictive of better keyword recognition, indicating that recognition of interrupted speech requires fast and efficient use of partial linguistic information. Taken together, these results suggest that multiple factors involved in perceptual organization of speech are affected by age and declines in the ability to resolve envelope distortions may result in poorer speech recognition in realistic listening environments among older adults.

ACKNOWLEDGMENTS

This work was supported (in part) by research Grant Nos. R01 DC00184 and P50 DC00422 from the National Institutes of Health/National Institute on Deafness and Other Communication Disorders (NIH/NIDCD), Grant No. 017727 from the American Academy of Audiology Foundation, and the South Carolina Clinical and Translational Research (SCTR) Institute, with an academic home at the Medical University of South Carolina, through NIH Grant No. UL1 TR000062. This investigation was conducted in a facility constructed with support from Research Facilities Improvement Program Grant No. C06 RR14516 from the National Center for Research Resources, NIH. This work was completed as part of the dissertation by W.J.B. to fulfill the requirement for Ph.D. at the University of Maryland. The authors would like to thank University of Maryland dissertation committee members Monita Chatterjee, Sandra Gordon-Salant, Rochelle Newman, and Jonathan Simon for their advice and support on this work, Lois Matthews for her assistance with participant recruitment, Larry Humes for providing the English version of the TRT, and the Associate Editor, two anonymous reviewers, and Michelle Molis for comments on an earlier version of this manuscript.

¹The fixed effect of talker sex did not significantly improve model fit, but significant improvements in model fit were observed with interactions between talker sex and age, talker sex and background, and a three-way

interaction between talker sex, background, and age. A detailed discussion of these complex interactions is beyond the scope of this report, but statistical results of *post hoc* models involving talker sex are included in Table II.

- ANSI (2010). ANSI/ASA S3.6-2010, *American National Standard Specification for Audiometers* (American National Standards Institute, New York).
- Baltes, P. B. (1993). "The aging mind: Potential and limits," *Gerontologist* **33**, 580–594.
- Bashford, J. A., Jr., Meyers, M. D., Brubaker, B. S., and Warren, R. M. (1988). "Illusory continuity of interrupted speech: Speech rate determines durational limits," *J. Acoust. Soc. Am.* **84**, 1635–1638.
- Bashford, J. A., Jr., and Warren, R. M. (1987). "Multiple phonemic restorations follow the rules for auditory induction," *Percept. Psychophys.* **42**, 114–121.
- Bashford, J. A., Jr., Warren, R. M., and Brown, C. A. (1996). "Use of speech-modulated noise adds strong 'bottom-up' cues for phonemic restoration," *Percept. Psychophys.* **58**, 342–350.
- Başkent, D. (2010). "Phonemic restoration in sensorineural hearing loss does not depend on baseline speech perception scores," *J. Acoust. Soc. Am.* **128**, EL169–EL174.
- Başkent, D., Eiler, C. L., and Edwards, B. (2010). "Phonemic restoration by hearing-impaired listeners with mild to moderate sensorineural hearing loss," *Hear. Res.* **260**, 54–62.
- Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). "Fitting linear mixed-effects models using lme4," *J. Stat. Softw.* **67**, 1–48.
- Benard, M. R., Mensink, J. S., and Başkent, D. (2014). "Individual differences in top-down restoration of interrupted speech: Links to linguistic and cognitive abilities," *J. Acoust. Soc. Am.* **135**, EL88–EL94.
- Ben-David, B. M., Tse, V. Y., and Schneider, B. A. (2012). "Does it take older adults longer than younger adults to perceptually segregate a speech target from a background masker?," *Hear. Res.* **290**, 55–63.
- Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* (MIT Press, Cambridge, MA), pp. 1–705.
- Brungart, D. S. (2001). "Informational and energetic masking effects in the perception of two simultaneous talkers," *J. Acoust. Soc. Am.* **109**, 1101–1109.
- Brysbaert, M., and New, B. (2009). "Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English," *Behav. Res. Methods*, **41**, 977–990.
- Clark, H. H. (1973). "The language-as-fixed-effect fallacy: A critique of language statistics in psychological research," *J. Verb. Learn. Verb. Behav.* **12**, 335–359.
- Cluff, M. S., and Luce, P. A. (1990). "Similarity neighborhoods of spoken two-syllable words: Retroactive effects on multiple activation," *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 551–563.
- Cole, R. A., and Perfetti, C. A. (1980). "Listening for mispronunciations in a children's story: The use of context by children and adults," *J. Verbal Learn. Verbal Behav.* **19**, 297–315.
- Cooke, M. (2006). "A glimpsing model of speech perception in noise," *J. Acoust. Soc. Am.* **119**, 1562–1573.
- Daneman, M., and Carpenter, P. A. (1980). "Individual differences in working memory and reading," *J. Verbal Learn. Verbal Behav.* **19**, 450–466.
- DeCaro, R., Peelle, J. E., Grossman, M., and Wingfield, A. (2016). "The two sides of sensory-cognitive interactions: Effects of age, hearing acuity, and working memory span on sentence comprehension," *Front. Psychol.* **7**, 236.
- Dirks, D. D., Bell, T. S., Rossman, R. N., and Kincaid, G. E. (1986). "Articulation index predictions of contextually dependent words," *J. Acoust. Soc. Am.* **80**, 82–92.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2002). "Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing," *J. Acoust. Soc. Am.* **111**, 2897–2907.
- Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (2003). "Recovery from prior stimulation: Masking of speech by interrupted noise for younger and older adults with normal hearing," *J. Acoust. Soc. Am.* **113**, 2084–2094.
- Durlach, N. I., Mason, C. R., Kidd, G., Arbogast, T. L., Colburn, S., and Shinn-Cunningham, B. G. (2003). "Note on informational masking (L)," *J. Acoust. Soc. Am.* **113**, 2984–2987.
- Eisenberg, L. S., Dirks, D. D., and Bell, T. S. (1995). "Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing," *J. Speech Hear. Res.* **38**, 222–233.

- Ezzatian, P., Li, L., Pichora-Fuller, K., and Schneider, B. A. (2015). "Delayed stream segregation in older adults: More than just informational masking," *Ear Hear.* **36**, 482–484.
- Festen, J. M., and Plomp, R. (1990). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," *J. Acoust. Soc. Am.* **88**, 1725–1736.
- Fitzgibbons, P. J., and Gordon-Salant, S. (1987). "Temporal gap resolution in listeners with high-frequency sensorineural hearing loss," *J. Acoust. Soc. Am.* **81**, 133–137.
- Fitzgibbons, P. J., and Gordon-Salant, S. (1994). "Age effects on measures of auditory duration discrimination," *J. Speech Hear. Res.* **37**, 662–670.
- Fitzgibbons, P. J., and Gordon-Salant, S. (1995). "Age effects on duration discrimination with simple and complex stimuli," *J. Acoust. Soc. Am.* **98**, 3140–3145.
- Fitzgibbons, P. J., and Gordon-Salant, S. (1996). "Auditory temporal processing in elderly listeners," *J. Am. Acad. Audiol.* **7**, 183–189.
- Fitzgibbons, P. J., and Gordon-Salant, S. (2001). "Aging and temporal discrimination in auditory sequences," *J. Acoust. Soc. Am.* **109**, 2955–2963.
- 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. Psychiatr. Res.* **12**, 189–198.
- Füllgrabe, C., Moore, B. C., and Stone, M. A. (2015). "Age-group differences in speech identification despite matched audiometrically normal hearing: Contributions from auditory temporal processing and cognition," *Front. Aging. Neurosci.* **6**, 347.
- Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. S., and Dahlgren, N. L. (1993). "DARPA TIMIT acoustic-phonetic continuous speech corpus CDROM," Linguistic Data Consortium, Philadelphia.
- George, E. L. H., Zekveld, A. A., Kramer, S. E., Goverts, S. T., Festen, J. M., and Houtgast, T. (2007). "Auditory and nonauditory factors affecting speech reception in noise by older listeners," *J. Acoust. Soc. Am.* **121**, 2362–2375.
- Gilbert, G., Bergeras, I., Voillery, D., and Lorenzi, C. (2007). "Effects of periodic interruptions on the intelligibility of speech based on temporal fine-structure or envelope cues," *J. Acoust. Soc. Am.* **122**, 1336–1339.
- Gilbert, J. L., Tamati, T. N., and Pisoni, D. B. (2013). "Development, reliability, and validity of PRESTO: A new high-variability sentence recognition test," *J. Am. Acad. Audiol.* **24**, 26–36.
- Gordon-Salant, S., and Cole, S. S. (2016). "Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing," *Ear Hear.* **37**, 593–602.
- Gordon-Salant, S., and Fitzgibbons, P. J. (1993). "Temporal factors and speech recognition performance in young and elderly listeners," *J. Speech Hear. Res.* **36**, 1276–1285.
- Gordon-Salant, S., Yeni-Komshian, G. H., Fitzgibbons, P. J., and Barrett, J. (2006). "Age-related differences in identification and discrimination of temporal cues in speech segments," *J. Acoust. Soc. Am.* **119**, 2455–2466.
- 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.
- Hinkle, D. E., Wiersma, W., and Jurs, S. G. (2003). *Applied Statistics for the Behavioral Sciences*, 5th ed. (Houghton Mifflin, Boston).
- Hofmann, D. A. (1997). "An overview of the logic and rationale of hierarchical linear models," *J. Manage.* **23**, 723–744.
- Humes, L. E., Kidd, G. R., and Lentz, J. J. (2013). "Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults," *Front. Syst. Neurosci.* **7**, 1–16.
- Humes, L. E., Kidd, G. R., and Lentz, J. J. (2016). "Corrigendum: Auditory and cognitive factors underlying individual differences in aided speech-understanding among older adults," *Front. Syst. Neurosci.* **10**, 1.
- Humes, L. E., Watson, B. U., Christensen, L. A., Cokely, C. G., Halling, D. C., and Lee, L. (1994). "Factors associated with individual differences in clinical measures of speech recognition among the elderly," *J. Speech Hear. Res.* **37**, 465–474.
- Jaekel, B. N., Newman, R. S., and Groupell, M. J. (2018). "Age effects on perceptual restoration of degraded interrupted sentences," *J. Acoust. Soc. Am.* **143**, 84–97.
- Jin, S. H., and Nelson, P. B. (2006). "Speech perception in gated noise: The effects of temporal resolution," *J. Acoust. Soc. Am.* **119**, 3097–3108.
- Jin, S. H., and Nelson, P. B. (2010). "Interrupted speech perception: The effects of hearing sensitivity and frequency resolution," *J. Acoust. Soc. Am.* **128**, 881–889.
- Kidd, G., Jr., Arbogast, T. L., Mason, C. R., and Gallun, F. J. (2005). "The advantage of knowing where to listen," *J. Acoust. Soc. Am.* **118**, 3804–3815.
- Kidd, G. R., and Humes, L. E. (2012). "Effects of age and hearing loss on the recognition of interrupted words in isolation and in sentences," *J. Acoust. Soc. Am.* **131**, 1434–1448.
- Krull, V., Humes, L. E., and Kidd, G. R. (2013). "Reconstructing wholes from parts: Effects of modality, age, and hearing loss on word recognition," *Ear Hear.* **34**, e14–e23.
- Kwon, B. J. (2012). "Token (version 1.36) [computer program]," available at auditorypro.com/aux/ (Last viewed 5 September 2012).
- Lee, J. H., and Humes, L. E. (2012). "Effect of fundamental-frequency and sentence-onset differences on speech-identification performance of young and older adults in a competing-talker background," *J. Acoust. Soc. Am.* **132**, 1700–1717.
- Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech," *J. Acoust. Soc. Am.* **22**, 167–173.
- Molis, M. R., Kampel, S. D., McMillan, G. P., Gallun, F. J., Dann, S. M., and Konrad-Martin, D. (2015). "Effects of hearing and aging on sentence-level time-gated word recognition," *J. Speech Lang. Hear. Res.* **52**, 481–496.
- Moore, B. C. (1990). "Co-modulation masking release: Spectro-temporal pattern analysis in hearing," *Br. J. Audiol.* **24**, 131–137.
- Moore, B. C. (2003). "Temporal integration and context effects in hearing," *J. Phon.* **31**, 563–574.
- Murdock, B. B., Jr. (1968). "Serial order effects in short-term memory," *J. Exp. Psychol.* **76**, 1–15.
- Pichora-Fuller, M. K. (2008). "Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing," *Int. J. Audiol.* **47**, S72–S82.
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). "How young and old adults listen to and remember speech in noise," *J. Acoust. Soc. Am.* **97**, 1197–1205.
- Powers, G. L., and Wilcox, J. C. (1977). "Intelligibility of temporally interrupted speech with and without intervening noise," *J. Acoust. Soc. Am.* **61**, 195–199.
- Rajan, R., and Cainer, K. E. (2008). "Ageing without hearing loss or cognitive impairment causes a decrease in speech intelligibility only in informational maskers," *Neuroscience* **154**, 784–795.
- R Development Core Team (2016). "R: A language and environment for statistical computing," R Foundation for Statistical Computing, Vienna, Austria, available at <http://www.R-project.org> (Last viewed 21 June 2018).
- Rönnerberg, J. (1990). "Cognitive and communicative function: The effects of chronological age and 'handicap age,'" *Eur. J. Cogn. Psychol.* **2**, 253–273.
- Saija, J. D., Akyurek, E. G., Andringa, T. C., and Başkent, D. (2014). "Perceptual restoration of degraded speech is preserved with advancing age," *J. Assoc. Res. Otolaryngol.* **15**, 139–148.
- Salthouse, T. A. (2000). "Aging and measures of processing speed," *Biol. Psychol.* **54**, 35–54.
- Shafiro, V., Sheft, S., and Risley, R. (2016). "The intelligibility of interrupted and temporally altered speech: Effects of context, age, and hearing loss," *J. Acoust. Soc. Am.* **139**, 455–465.
- Shafiro, V., Sheft, S., Risley, R., and Gygi, B. (2015). "Effects of age and hearing loss on the intelligibility of interrupted speech," *J. Acoust. Soc. Am.* **137**, 745–756.
- Shinn-Cunningham, B. G. (2008). "Object-based auditory and visual attention," *Trends Cogn. Sci.* **12**, 182–186.
- Shinn-Cunningham, B. G., and Wang, D. (2008). "Influences of auditory object formation on phonemic restoration," *J. Acoust. Soc. Am.* **123**, 295–301.
- Simpson, A. N., Matthews, L. J., and Dubno, J. R. (2013). "Lipid and C-reactive protein levels as risk factors for hearing loss in older adults," *Otolaryngol. Head Neck Surg.* **148**, 664–670.
- Srinivasan, S., and Wang, D. (2005). "A schema-based model for phonemic restoration," *Speech Commun.* **45**, 63–87.
- Stroop, J. R. (1935). "Studies of interference in serial verbal reactions," *J. Exp. Psychol.* **18**, 643–662.
- Studebaker, G. A. (1985). "A 'rationalized' arcsine transform," *J. Speech Hear. Res.* **28**, 455–462.
- Surprenant, A. M., Pitt, M. A., and Crowder, R. G. (1993). "Auditory recency in immediate memory," *Q. J. Psychol. A* **46**, 193–223.
- Takahashi, G. A., and Bacon, S. P. (1992). "Modulation detection, modulation masking, and speech understanding in noise in the elderly," *J. Speech Hear. Res.* **35**, 1410–1421.

- Tiffin, J., and Asher, E. J. (1948). "The Purdue Pegboard: Norms and studies of reliability and validity," *J. Appl. Psychol.* **32**, 234–247.
- Tombaugh, T. N., and McIntyre, N. J. (1992). "The mini-mental state examination: A comprehensive review," *J. Am. Geriatr. Soc.* **40**, 922–935.
- Trener, M. R., Crosson, B., DeBoe, J., and Leber, W. R. (1989). *The Stroop Neuropsychological Screening Test* (Psychological Assessment Resources, Odessa, FL).
- Tun, P. A., O’Kane, G., and Wingfield, A. (2002). "Distraction by competing speech in young and older adult listeners," *Psychol. Aging* **17**, 453–467.
- Vaden, K. I., Halpin, H. R., and Hickok, G. S. (2009). "Irvine Phonotactic Online Dictionary, version 2.0," available at <http://www.iphod.com> (Last viewed 27 September 2017).
- Verhaeghen, P. (2003). "Aging and vocabulary scores: A meta-analysis," *Psychol. Aging* **18**, 332–339.
- Verschuure, J., and Brocaar, M. P. (1983). "Intelligibility of interrupted meaningful and nonsense speech with and without intervening noise," *Percept. Psychophys.* **33**, 232–240.
- Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000). "Method for the selection of sentence materials for efficient measurement of the speech reception threshold," *J. Acoust. Soc. Am.* **107**, 1671–1684.
- Vitevitch, M. S., Luce, P. A., Pisoni, D. B., and Auer, E. T. (1999). "Phonotactics, neighborhood activation, and lexical access for spoken words," *Brain Lang.* **68**, 306–311.
- Wang, X., and Humes, L. E. (2010). "Factors influencing recognition of interrupted speech," *J. Acoust. Soc. Am.* **128**, 2100–2111.
- Warren, R. M. (1970). "Perceptual restoration of missing speech sounds," *Science* **167**, 392–393.
- Zekveld, A. A., George, E. L., Kramer, S. E., Goverts, S. T., and Houtgast, T. (2007). "The development of the text reception threshold test: A visual analogue of the speech reception threshold test," *J. Speech Lang. Hear. Res.* **50**, 576–584.