The intelligibility of interrupted and temporally altered speech: Effects of context, age, and hearing loss^{a)}

Valeriy Shafiro,^{b)} Stanley Sheft, and Robert Risley

Department of Communication Disorders and Sciences, Rush University Medical Center, 600 South Paulina Street, Suite 1012 AAC, Chicago, Illinois 60612, USA

(Received 3 June 2015; revised 22 December 2015; accepted 30 December 2015; published online 22 January 2016)

Temporal constraints on the perception of interrupted speech were investigated by comparing the intelligibility of speech that was periodically gated (PG) and subsequently either temporally compressed (PG_{TC}) by concatenating remaining speech fragments or temporally expanded (PG_{TE}) by doubling the silent intervals between speech fragments. Experiment 1 examined the effects of PG_{TC} and PG_{TE} at different gating rates (0.5–16 Hz) on the intelligibility of words and sentences for young normal-hearing adults. In experiment 2, older normal-hearing (ONH) and older hearing-impaired (OHI) adults were tested with sentences only. The results of experiment 1 indicated that sentences were more intelligible than words. In both experiments, PG_{TC} sentences were less intelligible than either PG or PG_{TE} sentences. Compared with PG sentences, the intelligibility of PG_{TE} sentences was significantly reduced by the same amount for ONH and OHI groups. Temporal alterations tended to produce a U-shaped rate-intelligibility function with a dip at 2–4 Hz, indicating that temporal alterations interacted with the duration of speech fragments. The present findings demonstrate that both aging and hearing loss negatively affect the overall intelligibility of interrupted and temporally altered speech. However, a mild-to-moderate hearing loss did not exacerbate the negative effects of temporal alterations associated with aging.

© 2016 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4939891>]

[VB]

Pages: 455–465

I. INTRODUCTION

Speech intelligibility involves assignment of temporally distributed acoustic cues to existing perceptual categories. Previous investigations of the speech signal characteristics that are important for intelligibility have frequently used a well-established experimental paradigm, pioneered by [Miller and Licklider \(1950\)](#), that consists of periodic interruption of speech by silence or modulated noise. The rate of interruption and the duty cycle (i.e., the percentage of speech-on time during each interruption cycle) can be gradually varied to control access to the information in the speech signal, thus revealing its perceptually salient aspects. The general finding has been that for many interruption rates, high intelligibility can be maintained with as little as 25%–50% of the original signal remaining ([Miller and Licklider, 1950](#); [Nelson and Jin, 2004](#); [Shafiro et al., 2011a](#); [Saija et al., 2014](#); [Shafiro et al., 2015](#)). Some investigators have also adopted this experimental approach to study contributions of temporal constraints on the perceptual processing of speech ([Garvey, 1953](#); [Huggins, 1964](#); [Heiman et al., 1986](#); [Shafiro et al., 2011b](#); [Saija et al., 2014](#); [Ghitza, 2014](#)). In a previous study, [Shafiro et al. \(2011b\)](#) examined the intelligibility of sentences that were either gated at different rates (0.5–16 Hz) and a 50% duty cycle or similarly gated and then temporally compressed by concatenating the

consecutive speech fragments retained after gating. The findings indicated a highly rate-dependent effect of concatenation on intelligibility. Compared with regularly gated sentences, intelligibility declined significantly for interruption rates of 2–4 Hz, while at lower or higher rates, both gated and gated-concatenated sentences had the same intelligibility. The present study extended this research to determine whether the perceptual costs associated with limiting the processing time in an interrupted speech signal can be offset by increasing the duration of silent intervals produced during gating. The effects of temporal alterations on the intelligibility of interrupted speech were further investigated to determine the effects of context, age, and hearing status.

A. Intelligibility of interrupted and temporally altered speech

Initial methods of systematically interrupting speech pioneered by [Miller and Licklider \(1950\)](#) were later combined with temporal alteration, leading to the development of time-compression techniques ([Garvey, 1953](#); [Fairbanks and Kodman, 1957](#)). In initial work, speech transmission rates were increased by concatenating the speech fragments that remained after periodic gating ([Garvey, 1953](#); [Fairbanks and Kodman, 1957](#); [Lee, 1971](#)). At sufficiently high interruption rates that preserved speech fragments of 40 ms or less, high speech intelligibility could be maintained even when up to 50% of the speech was deleted. Although this method of speech compression through concatenation has been supplanted by more advanced spectrally based techniques, its simplicity continues to make it an attractive paradigm for

^{a)}Portions of the data were presented at the 165th Meeting of the Acoustical Society of America

^{b)}Electronic mail: Valeriy_Shafiro@rush.edu

speech perception research. That is, periodically gated speech differs from compressed gated-concatenated speech by only a single stimulus parameter—the time between the retained speech fragments. Variation in intelligibility associated with changes in the duration of time intervals between successive speech fragments can therefore indicate temporal constraints on perceptual processing of the same degraded speech input across interruption rates.

A major focus of research inspired by the Miller and Licklider (1950) study has been on determining factors that affect the detection of audible speech fragments, often referred to as glimpses. Such glimpses of the original speech are thought to be integrated into higher order perceptual categories using some kind of “an intelligent warping process” (Moore, 2003). Past research has indicated that intelligibility of interrupted speech is strongly affected by the proportion of audible glimpses, i.e., the proportion of the original speech signal that can be detected (Cooke, 2006; Buss *et al.*, 2009; Wang and Humes, 2010; Kidd and Humes, 2010). However, other studies have also revealed a highly nonlinear interaction pattern between the proportion of the original speech retained and intelligibility (Gustafsson and Arlinger, 1994; Wang and Humes, 2010; Shafiro *et al.*, 2011a; Shafiro *et al.*, 2015). In that context, combining periodic interruptions with temporal alterations of speech has been a useful approach for identifying the temporal constraints on the perceptual integration of speech fragments of different duration.

In an early study that employed tape-splicing techniques to compare the intelligibility of gated vs gated-concatenated word stimuli, Garvey (1953) reported differences between the two types of interruption only after the proportion of deleted speech exceeded 50%. Later Huggins (1964) compared the intelligibility of interrupted speech passages presented at either a regular speaking rate or slightly speeded up by a factor of 1.19, observing some changes in intelligibility across interruption rates depending on speech presentation speed. Further developing this paradigm, Heiman *et al.* (1986) evaluated differences in the comprehension of spoken passages that were either gated or gated and subsequently concatenated. Listeners were read two 500-word passages and answered ten questions about each one. The researchers found that performance differences between the regularly gated vs gated-concatenated passages consistently increased as the total proportion of the deleted signal increased. When 33% of the speech was deleted, the differences were less than six percentage points, but they reached over 58 percentage points when 60% of speech was deleted. The researchers suggested that these decrements in performance arose due to insufficient time available for the processing of individual words in the passages when sensory input was severely limited. They concluded that as the perceptual redundancy of the speech cues is decreased due to gating, the amount of time necessary to process such incomplete signals must increase, an observation consistent with other findings in both auditory and visual modalities (Tulving and Gow, 1963; Shinn-Cunningham and Best, 2008; Ahissar *et al.*, 2008; Nahum *et al.*, 2008; Rönnberg *et al.*, 2013).

In more recent work, Shafiro *et al.* (2011b) used a similar approach to investigate the role of interval timing in the

perception of periodically gated and gated-concatenated speech. Unlike Heiman's *et al.* study in which the proportion of deleted speech was confounded with both interruption rate and duty cycle (because silent intervals were always 30 ms long for all conditions), Shafiro *et al.* (2011b) interrupted sentences at different rates from 0.5 to 16 Hz. In addition, the same proportion of speech in each interruption interval was maintained by keeping the same duty cycle of 50% across all rates. Similar to Heiman *et al.* (1986), the investigators found overall lower performance for gated-concatenated speech than for gated speech without concatenation. The results suggested that limiting processing time through removal of the silent intervals hindered perceptual integration of the information contained in adjacent speech fragments. The detrimental effect of concatenation was most apparent at interruption rates of 2–4 Hz, the rates for which word information is most ambiguous subsequent to gating.

Similarly to earlier work (Heiman *et al.*, 1986), these results were interpreted in terms of processing time constraints on the perception of sensory degraded speech. This explanation fits well with several contemporary accounts of speech perception that distinguish between fast and automatic vs slow and effortful perceptual processing of speech dependent on the level of sensory distortion. These accounts emphasize the role of trade-offs between the redundancy of the sensory input of speech and involvement of temporally constrained processing resources such as working memory and attention (Mattys, 1997; Ahissar *et al.*, 2008; Nahum *et al.*, 2008; Shinn-Cunningham and Best, 2008; Gygi and Shafiro, 2014; Jerger *et al.*, 1991; Wingfield and Tun, 2007; Miller and Wingfield, 2010). One way to further investigate the effect of processing time for lexically uncertain speech fragments is to further manipulate the duration of silent intervals while retaining the same amount of speech as produced by gating. An increased duration of silences in each interruption cycle can potentially provide listeners with more time to process the fragmented speech input and benefit intelligibility.

Furthermore, the rate-dependent effect of concatenation on the intelligibility of gated speech reported by Shafiro *et al.* (2011b), with a local minima at 2 Hz, is consistent with the previous findings of a U-shaped performance function across interruption rates (Miller and Licklider, 1950; Huggins, 1975; Nelson and Jin, 2004; Jin and Nelson, 2010; Shafiro *et al.*, 2011a; Shafiro 2011b; Saija *et al.*, 2014; Shafiro *et al.*, 2015). This characteristic function shape has been attributed to differences in the perceptual processes that support intelligibility on the low- and high-frequency sides of the rate-intelligibility function (Huggins, 1975; Ghitza and Greenberg, 2009; Ghitza, 2011; Shafiro *et al.*, 2015). On the low side, remaining speech fragments are, on average, long enough to support the percept of full words, while intervals between the fragments may be too long to effectively “bridge” the gaps. The opposite is the case on the high-frequency side. Although smaller individual speech fragments may not be long enough to effectively support single word identification, they are sufficiently frequent to allow for perceptual integration across the temporal gaps separating the smaller fragments. Consequently, if additional

processing time can be used to augment perceptual processing, it will be most beneficial at these rates at interruption rates with the lowest intelligibility (2–4 Hz).

B. Effects of age and hearing loss

Older adults, and especially those with a hearing loss, tend to experience a general decline in their speech perception abilities (Humes and Dubno, 2010). Compared with young normal-hearing (YNH) adults, they are more affected by signal distortions, such as the interruptions produced by gating or masking by modulated noise (Wilson *et al.*, 2010). Similarly, in older normal-hearing (ONH) and older hearing-impaired (OHI) adults, intelligibility declines with an increase in speech rate whether through natural speaker adjustments or signal processing (Janse, 2009). A number of studies have consistently demonstrated that the difficulty of ONH and OHI adults in the perception of fast natural or time-compressed speech is associated with higher-order cognitive factors such as working-memory capacity, processing speed and attention, factors that are also known to decline with age (Pichora-Fuller and Souza, 2003; Schneider *et al.*, 2005; Gordon-Salant *et al.*, 2007; Rönnberg *et al.*, 2013; Janse and Jesse, 2014).

The general consensus of these studies appears to be that when the rate of speech information transmission is increased, older adults do not have sufficient time to complete the lexical, syntactic, and semantic processing required to determine accurate word responses. However, studies that have examined the effects of temporally expanded speech have produced mixed results. Some have reported a positive effect of slower speech compared to the original natural tempo (Schmitt and McCroskey, 1981; Schmitt, 1983; Kupryjanow and Chyzewski, 2012; Piquado *et al.*, 2012; Gygi and Shafiro, 2012), while others have reported either no benefit or a decrement in speech perception (Gordon-Salant and Fitzgibbons, 1997; Nejime and Moore, 1998; Vaughan *et al.*, 2002). Gygi and Shafiro (2014) pointed out that the positive effects of slower speech were found mostly in studies that used cognitively demanding materials or tasks with increased attentional and working-memory load such as when listeners had to pay attention to the speech of two concurrent talkers. This is consistent with the notion that additional processing time available to listeners in slower speech may allow for more extensive cognitive and linguistic information processing.

Recently the effect of the temporal constraints associated with time compression on the perception of periodically interrupted speech in older listeners was examined by Saija *et al.* (2014). The researchers presented YNH and ONH listeners with sentences at three different speech rates: 0.5, 1, and 2, which, respectively, increased the speech rate by two, preserved the original, or slowed the rate by half using the pitch synchronous overlap and add algorithm (Moulines and Charpentier, 1990). Subsequent to compression, the sentences were interrupted by either gating with silence or with noise at six interruption rates: 0.625, 1.25, 2.5, 5, 10, and 20 Hz. As in other research with interrupted speech (Kidd and Humes, 2010; Molis *et al.*, 2015; Shafiro *et al.*, 2015;

Fogerty *et al.*, 2015), young adults outperformed older listeners overall. Sentences interrupted with noise were considerably more intelligible than those interrupted with silence. The researchers interpreted the differences in accuracy between interruptions with noise and silence as evidence for a top-down repair mechanism (Miller and Licklider, 1950; Verschuure and Brocaar, 1983). They suggested that the presence of noise during the interruptions provided greater sensory activation than silent interruptions, inducing more rigorous lexical activation and linguistic processing of the input speech. In that context, the researchers also observed that for older adults the positive effect of noise interruption compared to silence was greater for slower speech, suggesting that slower speech provided older listeners with more time to complete the perceptual processing during interruption intervals. On the other hand, for temporally compressed faster sentences, especially at higher interruption rates, performance in noise conditions was inferior as compared to gating with silence, consistent with energetic masking of the speech by the noise.

Overall, Saija's *et al.* (2014) findings are consistent with the view that successful perception of either compressed or expanded speech critically depends on temporally constrained cognitive processing resources. However, because the interruptions were applied to speech after it was already temporally altered, the amount of linguistic information within each retained speech fragment varied across interruption rates. In addition, the presence of energetic masking introduced by noise might have obscured the role of processing time in some conditions. The approach taken in the current study provides a complementary way to investigate the effect of temporal constraints on speech perception. Here, after gating the original speech at a given rate, the retained speech fragments remained the same, while the duration of silence between fragments varied. Differences in performance can thus indicate the effect of time between speech fragments of the same duration on intelligibility.

C. Context effects

Previous work consistently demonstrates that listeners tend to compensate for degraded sensory information by relying on higher-order contextual cues that can restrict the number of response alternatives in speech perception tasks (Miller *et al.*, 1951; Verschuure and Brocaar, 1983; Bronkhorst *et al.*, 1993; Pichora-Fuller *et al.*, 1995; Benichov *et al.*, 2012; Bernstein *et al.*, 2012; Kidd and Humes, 2010; Molis *et al.*, 2015). Syntactic structures and semantic relationships among the words of a sentence, combined with coarticulatory and suprasegmental cues contained in speech signals, can be used to resolve the uncertainty produced by sensory degradation (Pichora-Fuller *et al.*, 1995; Boothroyd, 2010). Compared with sentences, intelligibility generally tends to be lower for semantically unrelated words when speech signals are degraded (Shafiro *et al.*, 2011a; Fogerty *et al.*, 2012). Furthermore, for speech gated at different rates, the intelligibility advantage of sentences, as compared with unrelated words, is present only at some interruption rates. This suggests that the integration of

higher-order contextual information may be influenced by the characteristics of individual speech fragments retained or discarded during interruptions at specific rates.

Shafiro *et al.* (2011a) compared the effect of gating on the intelligibility of either meaningful sentences or sequences of randomly chosen monosyllabic words (one to three words per sequence). For YNH listeners, with 50% of the original speech present, sentences were always more intelligible than unrelated words. However, when the amount of the original speech in each interruption interval was decreased to 25%, the difference between sentences and words was less pronounced and depended on interruption rate. Overall, the sentences were still perceived more accurately at the lowest gating rate of 0.5–1 Hz and the highest rate of 8 Hz, but they were no longer more intelligible than words with gating rates of 2–4 Hz. This suggests that the effective use of higher-order contextual cues was influenced by both the total amount of sensory evidence and by the duration of individual speech fragments. However, the sentence advantage may be similarly lost even for the sentences gated with a 50% duty cycle by limiting the time between speech fragments that could be used to resolve perceptual ambiguity by utilizing sentence-level cues. In that case, at these middle rates of 2–4 Hz, the intelligibility of concatenated sentences will be similar to that of unrelated words. On the other hand, as the silent-interval duration is doubled, sentences may regain their advantage over words which lack contextual cues to assist higher-order processing.

D. Present study

The present study investigated the effects of temporal alterations, i.e., compression and expansion, on the intelligibility of interrupted speech. Speech stimuli were periodically gated with silence (PG) or periodically gated and either temporally concatenated (PG_{TC}) by concatenating adjacent speech fragments or temporally expanded (PG_{TE}) by doubling the duration of the silent intervals within each gating cycle. Experiment 1 investigated the role of sentence-level context by comparing the intelligibility of temporally altered interrupted sentences to that of unrelated words in YNH adults. Experiment 2 examined the intelligibility of temporally altered interrupted speech for listeners of varying age and hearing status. It was hypothesized that if silent intervals within interruption cycles can be effectively utilized for additional perceptual processing of speech fragments, performance with both PG and PG_{TE} speech should be superior to that with PG_{TC} speech. Furthermore, these positive effects of processing time will result in greater intelligibility of sentences over words in experiment 1 as well as lead to a rate-dependent performance benefit for ONH and OHI listeners in experiment 2.

II. EXPERIMENT 1

Experiment 1 investigated the intelligibility of interrupted and temporally altered speech using sentences and semantically unrelated words. It was designed to compare rate-intelligibility functions for sentences and words to

determine how sentence-level context may influence performance.

A. Method

1. Subjects

Participants were 10 YNH adult native speakers of American English (mean age 23.2 yr, range: 19–28 yr, 9 females) with hearing thresholds between 0.5 and 8 kHz of 15 dB hearing level (HL) or less. Participants received a partial course credit for participation.

2. Stimuli, design, and procedure

Speech stimuli were HINT sentences (Nilsson *et al.*, 1994) with simple syntactic and semantic structure (e.g., “strawberry jam is sweet”) and monosyllabic words, CNC (consonant-vowel-consonant) words (Peterson and Lehiste, 1962), spoken by a male talker of American English (TigerSpeech Technology <http://www.tigerspeech.com/>). For the word stimuli, three randomly selected words were combined to form each stimulus. As in Shafiro *et al.* (2011a), individual words within each word sequence were separated by 80-ms intervals. A 5-s interstimulus interval was used between consecutive stimulus trials.

All words and sentence stimuli were interrupted using one of the three interruption types (PG, PG_{TC}, PG_{TE}). Stimulus interruption was implemented using MATLAB 7.5 software. All stimuli for every interruption type and rate condition were processed prior to presentation during subject testing. For the PG conditions, sentences and words were gated at six rates (0.5, 1, 2, 4, 8, and 16 Hz) with a 50% duty cycle by multiplying the original stimuli with a rectangular gating window and a random starting phase.¹ Stimuli were similarly gated in the PG_{TC} conditions with subsequent concatenation of all retained speech fragments to eliminate the silent intervals of each interruption cycle. This resulted in shortening of the total duration of each stimulus utterance by half. For the final PG_{TE} conditions, subsequent to stimulus gating, the durations of speech-off silent intervals of each interruption cycle were doubled. This resulted in expanding the total duration of each stimulus utterance by 1.5. The three types of interruptions are illustrated in Fig. 1(a) using a sample sentence stimulus shown intact in the top panel. The lower three panels [Figs. 1(b), 1(c), and 1(d)] show the waveform of the same sentence periodically gated using a 1-Hz interruption rate and 50% duty cycle, for the PG, PG_{TC}, and PG_{TE} conditions, respectively.

Sentence and word stimuli were presented in two different sessions, each lasting about 1 h. Interruption rates (0.5–16 Hz) were blocked by interruption types (PG, PG_{TC}, PG_{TE}), with both interruption rates and interruption types randomized for each subject. All stimuli were presented in a double-walled sound-attenuated room, through an Edirol UA25 24-bit soundcard with on-board anti-alias filtering. Stimuli were presented diotically in quiet through Sennheiser 250 II headphones at 70 dB sound pressure level (SPL). Every ten-sentence HINT list was preceded by a short practice period with five IEEE sentences interrupted in the

same manner as the test list. These IEEE practice sentences were not scored, and no feedback was provided for either practice or test sentences. For conditions with CNC words as stimuli, subjects were first presented with 8 three-word sequences to practice, followed by a list of 17 three-word test stimuli. For both sentences and words, subjects were asked to repeat what they heard after every trial.

B. Results and discussion

Percent correct intelligibility scores, converted to rationalized arcsine-transformed units (RAUs; [Studebaker, 1985](#)) are shown in Fig. 2 for sentences and words. A $2 \times 3 \times 6$ analysis of variance (ANOVA) indicated significant main effects of all three factors: speech materials [$F(1, 8) = 521.49$, $p < 0.001$, $\eta^2 = 0.98$], interruption type [$F(2, 16) = 26$, $p < 0.001$, $\eta^2 = 0.76$], and rate [$F(5, 40) = 426.11$, $p < 0.001$, $\eta^2 = 0.98$]. The significant main effect of speech materials confirmed that sentences were significantly more intelligible than words. The significant effect of interruption rate similarly followed expectations based on past studies reviewed in the preceding text. Planned comparisons demonstrated that periodically gated (PG) or temporally expanded (PG_{TE}) conditions were not significantly different from each other ($p > 0.61$), while both were significantly more intelligible than temporally compressed (PG_{TC}) speech stimuli ($p < 0.001$). There were three two-way interactions, rate $\times$ interruption type [$F(10, 80) = 7.68$, $p < 0.001$, $\eta^2 = 0.49$], rate $\times$ material [$F(5, 40) = 6.28$, $p < 0.001$, $\eta^2 = 0.44$], and material $\times$ interruption type [$F(2, 16) = 8.28$, $p < 0.003$, $\eta^2 = 0.51$]. A three-way interaction of rate $\times$ interruption type $\times$ material was also significant [$F(10, 80) = 2.88$, $p < 0.004$, $\eta^2 = 0.26$]. *Post hoc* pairwise comparisons across specific interruption rates for the three interruption types

were conducted to further investigate possible causes of these interactions. A Bonferroni correction was used to control the familywise error rate with multiple comparisons. *Post hoc* results indicated that sentences were perceived more accurately than words at all interruption rates, in the PG_{TE} conditions and at all but 2 Hz in the PG conditions. However, in the PG_{TC} conditions, sentences were perceived as accurately as words in the lower-mid rates (0.5, 1, 2, 4 Hz), with better intelligibility for sentences than words only at the higher rates (8, 16 Hz). This suggests that the lack of time intervals between speech fragments in the PG_{TE} conditions prevented listeners from utilizing higher-order contextual sentence-level cues with low-to-mid rate gating, thus eliminating the intelligibility advantage of sentences over words in a rate-dependent manner. In contrast, at the two highest rates, with speech fragments of much smaller size ($\sim 62, 31$ ms) and more frequent sampling of speech, listeners were able to regain access to sentence-level cues.

On the other hand, lack of significant differences between PG and PG_{TE} conditions for both words and sentences indicates that the additional time in each cycle that could potentially be used for more exhaustive perceptual processing of ambiguous speech fragments did not benefit intelligibility. Nor did it lead to any intelligibility decrement as could have resulted from additional demands on memory storage. However, it is possible that as working-memory capacity and perceptual processing speed decrease with age and as the sensory input of the speech signal deteriorates due to hearing loss, the effect of additional processing time for ONH and OHI listeners may differ from that for YNH listeners. If additional time can be used to support more exhaustive perceptual processing of higher-order contextual cues, the performance in the PG_{TE} conditions, compared with PG performance, will improve. Alternatively, memory constraints may play a dominant role, limiting the amount of

FIG. 1. Illustrations of temporally altered speech. Top (a) shows the waveform of an original intact sentence. (b) shows the same sentence gated using a 1-Hz interruption rate and a 50% duty cycle (PG). (c) shows the corresponding time-compressed sentence after concatenation of the remaining speech within each interruption cycle (PG_{TC}), while panel (d) shows the temporally expanded version of the sentence with each silent interval obtained through gating and doubled in duration.

FIG. 2. Intelligibility of HINT sentences and CNC words across interruption rates in young normal hearing listeners shown for three interruption types of the study. Left panel shows results for periodic gating (PG). Middle panel show results for temporal compression in which periodic gating was followed by concatenation of adjacent fragments (PG_{TC}). Right panel shows results for temporal expansion in which periodic gating was followed by doubling silent intervals between adjacent speech fragments (PG_{TE}). The error bars represent 95% confidence intervals.

time sensory input can be stored. Performance in PG_{TE} conditions would then decline in comparison to PG conditions. These competing hypotheses were tested in experiment 2.

III. EXPERIMENT 2

A. Method

Experiment 2 examined the intelligibility of interrupted and temporally altered speech in ONH and OHI listeners. In conjunction with the YNH results from experiment 1, it was designed to examine the effect of aging and hearing loss.

1. Subjects

Participants in experiment 2 were separated into two groups based on their age and hearing status (Fig. 3). Participants in the ONH group were 17 adults (mean age 68.2 yr, range: 61–87 yr, 12 females) with a mean pure-tone average (PTA) measured across 0.5, 1, and 2 kHz for their better ear of 15 dB HL (SD = 3.3 dB, range: 10–22 dB HL). The hearing status criteria for this group were adopted from previous work (Shafiro *et al.*, 2015; Sheft *et al.*, 2012). Nevertheless, the ONH participants, despite being labeled normal-hearing based on the better-ear PTA, had at least mild hearing losses in at least one ear at 4 kHz (12 participants) and 8 kHz (14 participants). Participants in the OHI group were 14 adults (mean age 70.6 yr, range: 60–85 yr, 6 females) with a mean PTA for their better ear of 31.5 dB HL (SD = 6.1 dB, range: 23–42 dB HL). All participants in experiment 2 were native speakers of American English and received a small financial compensation for their participation in the study.

2. Stimuli, design, and procedure

The sentences of experiment 1 were used as stimuli in experiment 2. They were interrupted and temporally altered in the same way as in experiment 1 but did not include stimuli

interrupted at the rate of 1 Hz to compensate for the longer testing time taken by ONH and OHI listeners.² As in experiment 1, HINT lists were randomly assigned to interruption rates which were in turn blocked by interruption types and randomized for each subject. Prior to each ten-sentence test condition, listeners practiced with a list of five IEEE sentences that were interrupted and temporally altered in the same way as the test sentences. In all other respects, the procedures of experiment 2 followed those of experiment 1.

B. Results and discussion

Percent correct intelligibility scores of the ONH and OHI participants, converted to RAU, were combined with YNH participant scores from experiment 1 to examine the

FIG. 3. Better-ear audiometric thresholds for older-normal hearing (ONH) and older hearing-impaired (OHI) groups. The error bars represent 1 standard deviation, shown on one side of each curve for better visibility.

effect of age. A $3 \times 3 \times 5$ ANOVA was performed across three listener groups (YNH, ONH, OHI), three interruption types (PG_{TC} , PG, PG_{TE}), and five rates (0.5–16 Hz). The ANOVA revealed significant main effects of listener group [$F(2, 38) = 31.26$, $p < 0.001$, $\eta^2 = 0.62$], interruption type [$F(2, 76) = 54.47$, $p < 0.001$, $\eta^2 = 0.59$], and rate [$F(4, 152) = 343.04$, $p < 0.001$, $\eta^2 = 0.9$]. Planned comparisons indicated that overall intelligibility was highest for YNH, lower for ONH, and lower still for OHI listeners ($p < 0.01$). PG speech was significantly more intelligible than PG_{TE} speech, which in turn was significantly more intelligible than PG_{TC} speech. All two-way and one three-way interactions among group, interruption type and rate were also significant at $p < 0.05$.

The differences in group performance across interruption type and rate are illustrated in Fig. 4. For YNH listeners in PG conditions, performance continuously increased with gating rate. For ONH listeners, PG performance remained asymptotic between 0.5 and 2 Hz and then increased, while for OHI, PG performance decreased between 0.5 and 4 Hz before rising again for 8 and 16 Hz. In the PG_{TC} conditions, all three groups demonstrated a rate-dependent variation in performance with local minima at 2 Hz. In contrast, for PG_{TE} performance of the three groups diverged. For YNH listeners, performance was similar to PG performance, but for ONH and OHI listeners it was significantly ($p < 0.05$) lower than in the PG condition.

Overall the smallest group differences for the three interruption types (PG, PG_{TC} , PG_{TE}) were found at the lowest interruption rate of 0.5 Hz. These group differences, ranging between 1.6 and 10.8 RAU, were nevertheless significant ($p < 0.05$) for PG and PG_{TE} interruption conditions. The largest overall intelligibility differences between YNH and the two older groups were found in the PG_{TE} conditions. The highest magnitude of group difference was 73.6 RAU at 4 Hz for YNH and OHI groups, while it was 41.0 RAU for YNH and ONH groups at the same rate. On the other hand,

in contrast with absolute intelligibility scores, the overall magnitude of intelligibility decrements between PG and PG_{TE} conditions was the same for ONH and OHI groups. The maximum decrement due to temporal expansion (i.e., PG score - PG_{TE} score) was at 8 Hz where it reached 30 RAU for both groups. This indicates that for ONH and OHI listeners, increasing the duration of silent intervals between speech fragments did not lead to improved intelligibility compared with regularly gated speech. Instead it had a detrimental effect. However, the magnitude of the negative effect of temporal alterations, for either PG_{TC} or PG_{TE} , did not differ between ONH and OHI groups.

To further delineate contributions of hearing status and aging to the intelligibility of interrupted and temporally altered speech, a partial correlation analysis was conducted. It revealed that when better-ear PTA was controlled, there were no significant correlations of ONH and OHI listener age with intelligibility scores for any interruption type at any interruption rate. In contrast, when age was controlled, ONH and OHI better-ear PTA was significantly correlated with intelligibility ($p < 0.05$, with Bonferroni correction) at rates of 4–16 Hz for the three interruption types, with Pearson correlation magnitudes in the moderate range: -0.49 to -0.69 . The prevalence of significant higher-magnitude correlations on the higher-frequency side of the rate-intelligibility function suggests that the decrement associated with hearing loss is enhanced for speech fragments of small (sub-syllabic) duration. This is consistent with the present finding that the differences between ONH and OHI group performance, described in the preceding text, are also larger at interruption rates of 4–16 Hz. At these rates, information contained in individual speech fragments appears insufficient to support lexical access without temporal integration across speech fragments (Molis *et al.*, 2015). Reduced audibility may thus obscure perceptually salient acoustic cues contained in individual fragments, leading to decreased accuracy.

FIG. 4. Sentence intelligibility across initial gating rates and 50% duty cycle for three listener groups: young normal-hearing adults (YNH), older normal-hearing adults (ONH), and older hearing-impaired adults (OHI). Each panel shows performance across interruption rates with interruption type the parameter: periodic gating (PG), time-compression (PG_{TC}) and time-expansion (PG_{TE}). The error bars represent 95% confidence intervals, shown on one side of each curve for better visibility.

IV. GENERAL DISCUSSION

The findings of the present study demonstrate that the negative effects of interrupted and temporally altered speech vary across listeners of different age and hearing abilities. When periodically gated sentences were temporally compressed by concatenating adjacent speech fragments, intelligibility declined for all groups regardless of age or hearing status. When interrupted sentences were temporally expanded by doubling the silent intervals between the speech fragments, intelligibility remained at the same level as for regularly gated speech for YNH listeners but declined for ONH and OHI listeners. These group differences in performance for regularly gated and gated and temporally altered conditions may have resulted from the contributions of at least two general factors that distinguish YNH listeners: (a) more robust initial sensory representations of the speech fragments retained after gating and (b) more effective processing of sensory information due to greater cognitive abilities such as superior memory capacity, attention, or perceptual processing speed.

The involvement of working-memory constraints is supported by a large number of studies that attribute the decline in speech perception abilities in ONH and OHI adults to a decrease in working-memory function that accompanies aging, and can be exacerbated by hearing loss (Akeroyd, 2008). Reduced working-memory capacity related to aging, combined with the negative effects of hearing impairment, has been shown to exacerbate speech perception difficulties (Pichora-Fuller *et al.*, 1995; Shinn-Cunningham and Best, 2008; Schneider *et al.*, 2010; Rönnberg *et al.*, 2013). A long-standing finding in various sensory modalities is that in response to a decrease in the quality and quantity of sensory information, more processing time is needed (Fairbanks and Kodman, 1957; Huggins, 1964; Heiman *et al.*, 1986; Shinn-Cunningham and Best, 2008; Janse, 2009; Salthouse, 1996). Sensory degradation of speech input may lead to the activation of a greater number of lexical items, producing greater competition among items (Miller and Wingfield, 2010; Piquado *et al.*, 2012). Resolving increased lexical competition would in turn require a longer processing time and result in greater perceptual effort (Nahum *et al.*, 2008; Ahissar *et al.*, 2008; Rönnberg *et al.*, 2013).

In the present study, when additional time was added in the temporally expanded conditions, instead of a benefit, a decrement in performance was observed for ONH and OHI listeners with neither benefit nor decrement obtained for YNH listeners. The latter finding contradicts the initial expectations of the perceptual benefit for temporally expanded speech, derived from speculation on processing time limitations in some, but not all, past experimental studies (Schmitt and McCroskey, 1981; Schmitt, 1983; Kupryjanow and Chyzewski, 2012; Piquado *et al.*, 2012; Gygi and Shafiro, 2014). It is possible that in older and hearing-impaired listeners, the potential benefit of additional processing time in the temporally expanded conditions of the present study was attenuated by faster memory decay of the partially encoded sensory information in working memory. The greater rate of decay in older listeners, who tend to have

smaller working-memory capacity, may have been further compounded by sensory degradation due to hearing loss, resulting in the poorer performance of OHI compared to ONH listeners. The present results, however, demonstrate that the negative effect of hearing loss was only observed for overall intelligibility. Hearing loss did not seem to contribute to the decrement in performance due to temporal alteration: performance in both ONH and OHI groups in the temporally compressed and temporally expanded conditions decreased by the same amount relative to that in the regularly gated conditions.

The observed rate-dependent pattern of results further may indicate that the effect of temporal alterations on intelligibility could be mediated by the characteristics of the linguistic information contained in speech fragments after gating. For speech interrupted at a low rate of 0.5 Hz, individual fragments are likely to contain complete or nearly complete words, which could support quick and relatively effortless lexical access (Molis *et al.*, 2015). The corresponding intervals between fragments, on the other hand, may be too long to effectively benefit from further perceptual, linguistic, or cognitive processing. Consequently, performance at the 0.5 Hz interruption rate did not change significantly for any group with either type of temporal alteration compared to regularly gated conditions. The small within- and between-group differences at the rate of 0.5 Hz are especially noteworthy because the time intervals separating speech fragments at this rate were the longest, reaching 2 s in the temporally expanded condition. The minimal variation in intelligibility across interruption types or subject groups with this long interval duration could reflect the efficient coding of linguistic information contained in the individual speech fragments.

In contrast, at higher rates, starting with 2 Hz, although intervals between consecutive fragments were substantially shorter, individual speech fragments were less likely to contain enough linguistic information for accurate word identification. This could lead to greater difficulties for ONH and OHI listeners. Recent work has indicated that ONH and OHI listeners require longer word fragments for accurate word identification than YNH listeners (Molis *et al.*, 2015). With shorter fragments at 2 and 4 Hz interruption rates, large group differences between younger and older adults emerged. At these middle and higher rates, word identification would likely require integrating information across two or more sub-syllabic speech fragments (Huggins, 1975; Samuel, 1991; Shafiro *et al.*, 2015). The need to “bridge” perceptual gaps between fragments would be expected to lead to a greater involvement of working memory in order to support a protracted course of lexical access (Rönnberg *et al.*, 2013). Nevertheless, more work is needed in delineating the role of working memory in the perception of interrupted and temporally altered speech as several recent studies have not revealed clear associations between the intelligibility of interrupted speech and scores on standard working-memory tests (Shafiro *et al.*, 2015; Nagaraj and Knapp, 2015).

The nonlinearity in the effect of absolute interval duration on differences in performance between regularly gated and

temporally expanded conditions is also consistent with the view that distinguishes between perceptual processing of interrupted speech at low and high interruption rates (Huggins, 1975; Samuel, 1991; Ghitza and Greenberg, 2009; Shafiro *et al.*, 2011a; Ghitza, 2014; Shafiro *et al.*, 2015). Recent studies by Ghitza (Ghitza and Greenberg, 2009; Ghitza, 2014) have suggested a neurologically plausible cortical mechanism that may be able to account for the nonlinear effects of time alterations. From that perspective, the variable effects of interval duration in interrupted speech reflect synchronization patterns between input speech and the cortical theta rhythm, which approximately corresponds to syllable timing. Speech intelligibility improves when input speech can be synchronized with cortical oscillations, but it declines when theta oscillations cannot be synchronized with the speech rate. This synchronization process, in turn, is driven by salient acoustic landmarks of the speech signal found in the speech envelope (Doelling *et al.*, 2014). In relation to present results, the time added in the PG_{TE} conditions can be beneficial for speech processing when it changes the speech input rate to correspond to the rate of cortical oscillations but will be disruptive otherwise. This view is appealing as it provides a testable neurophysiologic basis for the nonlinear effects of temporal alterations. However, more work is needed to account for the differences in performance based on age and hearing impairment across specific rates.

The greater difficulties of older normal-hearing and hearing-impaired adults at higher interruption rates may be also related to a general decrease in auditory temporal processing abilities which accompanies aging (Füllgrabe *et al.*, 2006; Gordon-Salant *et al.*, 2007; Humes and Dubno, 2010; Sheft *et al.*, 2012). Given the importance of slow varying envelope fluctuations for the perception of sentences (Fogerty *et al.*, 2012), it is possible that changes in the sentence envelopes introduced by gating and further exacerbated by temporal alterations contributed to the decline in intelligibility in the ONH and OHI groups. It is also conceivable that lesser sensitivity to envelope changes might have obscured some of the acoustic landmarks used to synchronize speech input with cortical oscillators, thus negatively affecting the perceptual outcome (Ghitza, 2014).

With respect to effects of sentence context, consistent with previous work (Dirks *et al.*, 1969; Shafiro *et al.*, 2011a; Kidd and Humes, 2010), interrupted and temporally altered sentences were overall more intelligible than similarly processed words. This finding confirms that listeners can successfully compensate for the sensory degradation of input signals by using preexisting knowledge of linguistic structure to maximize syntactic and semantic coherence (Miller *et al.*, 1951; Dirks *et al.*, 1969; Pichora-Fuller *et al.*, 1995; Elliot, 1995; Kidd and Humes, 2010). However, the advantage of sentences over words was present only in the regularly gated and temporally expanded conditions. In the temporally compressed conditions, when silent intervals between retained speech fragments were eliminated, sentences were no longer more intelligible than words at most rates. This may indicate that the time intervals between speech fragments in the regularly gated and temporally expanded conditions were used to

access higher-order contextual cues that were available for sentences but not for words.

The magnitude of intelligibility variation across interruption rates was also greater for sentences than for words. In PG and PG_{TE} conditions, sentence accuracy improved with rate while word accuracy was characterized by a U-shaped intelligibility function with a minimum at 1 Hz. This is in line with previous observations that the effect of speech materials differs across interruption rates (Dirks *et al.*, 1969; Shafiro *et al.*, 2011a). Overall, it appears that in addition to temporal constraints, access to sentence-level context in interrupted sentences is mediated by linguistic information contained in speech fragments of different durations.

V. SUMMARY

The present study investigated the role of sensory and temporal constraints on the perception of interrupted and temporally altered speech in adults of different age and hearing status. The intelligibility of speech gated at different rates between 0.5 and 16 Hz with a 50% duty cycle was compared to that of speech in which silent intervals at each rate were either eliminated by concatenating speech fragments or increased by doubling the silent intervals between consecutive speech fragments. An additional comparison was made in YNH listeners between two kinds of speech materials that differed in contextual information, words vs sentences. As predicted, compared to regularly gated speech, temporal alteration of gated speech by concatenation produced a decrease in performance for YNH, ONH, and OHI listeners, regardless of age and hearing status. Contrary to expectations, increasing the duration of silent intervals between speech fragments also resulted in a decrease in performance, but only for older and hearing-impaired listeners. However, the magnitude of the decrement was similar for both groups of older listeners, suggesting that the negative effect of hearing loss was limited to overall intelligibility. The changes in intelligibility across interruption rates were also rate dependent for all groups with local minima at 2 Hz. The nonlinear changes in intelligibility across interruption rates suggest an interaction between the duration of speech fragments produced by gating and the perceptual processing abilities that characterize the different listener groups.

Finally, the finding of overall greater intelligibility of sentences over words in the regularly gated and temporally expanded conditions indicates that additional coarticulatory, syntactic, and semantic cues contained in sentences make perceptual processing of interrupted speech more efficient. These higher-level contextual cues may be useful in reducing the uncertainty associated with removal of information from the speech signal during gating.

ACKNOWLEDGMENTS

Research reported in this publication was supported by NIH-NIDCD Award Nos. R03 DC008676 and R15 DC011916. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. We would like to thank

Karson Glass for his assistance with stimulus preparation and analysis.

¹Although gating with a square wave introduces a systematic distortion in the signal due to rapid signal onset, our observations and recent research (Wilson, 2014) did not indicate intelligibility differences due to variation in the rise time of the onset ramps.

²ONH and OHI listeners were also presented with PG and PG_{TC} interrupted stimuli based on 75% duty cycle. However, these additional conditions with greater proportion of preserved speech resulted in highly predictable performance patterns that mimicked those of the corresponding 50% duty cycle stimuli albeit at a higher overall level of performance. Because results from the 75% duty cycle do not provide new insights into the questions asked in this study, they will not be considered further.

Ahissar, M., Nahum, M., Nelken, I., and Hochstein, S. (2008). "Reverse hierarchies and sensory learning." *Phil. Trans. R. Soc. B*, **364**, 285–299.

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. Audiol.* **47**, S53–71.

Benichov, J., Cox, C. L., Tun, P. A., and Wingfield, A. (2012). "Word recognition within a linguistic context: Effects of age, hearing acuity, verbal ability and cognitive function." *Ear Hear.* **32**, 250–256.

Bernstein, J. G. W., Summers, V., Iyer, N., and Brungart, D. S. (2012). "Set-size procedures for controlling variations in speech-recognition performance with a fluctuating masker." *J. Acoust. Soc. Am.* **132**, 2676–2689.

Boothroyd, A. (2010). "Adapting to changed hearing: The potential role of formal training." *J. Am. Acad. Audiol.* **21**, 601–611.

Bronkhorst, A. W., Bosman, A. J., and Smoorenburg, G. F. (1993). "A model for context effects in speech perception." *J. Acoust. Soc. Am.* **93**, 499–509.

Buss, E., Whittle, L. N., Grose, J. H., and Hall, J. W. (2009). "Masking release for words in amplitude-modulated noise as a function of modulation rate and task." *J. Acoust. Soc. Am.* **126**, 269–280.

Cooke, M. (2006). "A glimpsing model of speech perception in noise." *J. Acoust. Soc. Am.* **119**, 1562–1573.

Dirks, D. D., Wilson, R. H., and Bower, D. R. (1969). "Effect of pulsed masking on selected speech materials." *J. Acoust. Soc. Am.* **46**, 898–906.

Doelling, K. B., Arnal, L. H., Ghitza, O., and Poeppel, D. (2014). "Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing." *NeuroImage* **85**, 761–768.

Elliott, L. L. (1995). "Verbal auditory closure and the Speech Perception in Noise (SPIN) test." *J. Speech Hear. Res.* **38**, 1363–1376.

Fairbanks, G., and Kodman, F., Jr. (1957). "Word Intelligibility as a function of time compression." *J. Acoust. Soc. Am.* **29**, 836–841.

Fogerty, D., Ahlstrom, J. B., Bologna, W. J., and Dubno, J. R. (2015). "Sentence intelligibility during segmental interruption and masking by speech-modulated noise: Effects of age and hearing loss." *J. Acoust. Soc. Am.* **137**, 3487–3501.

Fogerty, D., Kewley-Port, D., and Humes, L. E. (2012). "The relative importance of consonant and vowel segments to the recognition of words and sentences: Effects of age and hearing loss." *J. Acoust. Soc. Am.* **132**, 1667–1678.

Füllgrabe, C., Berthommier, F., and Lorenzi, C. (2006). "Masking release for consonant features in temporally fluctuating background noise." *Hear. Res.* **211**, 74–84.

Garvey, W. D. (1953). "The intelligibility of speeded speech." *J. Exp. Psych.* **45**, 102–108.

Ghitza, O. (2011). "Linking speech perception and neurophysiology: Speech decoding guided by cascaded oscillators locked to the input rhythm." *Front. Psychol.* **2**, 130.

Ghitza, O. (2014). "Behavioral evidence for the role of cortical theta oscillations in determining auditory channel capacity for speech." *Front. Psychol.* **5**, 652.

Ghitza, O., and Greenberg, S. (2009). "On the possible role of brain rhythms in speech perception: Intelligibility of time-compressed speech with periodic and aperiodic insertions of silence." *Phonetica* **66**, 113–126.

Gordon-Salant, S., and Fitzgibbons, P. J. (1997). "Selected cognitive factors and speech recognition performance among young and elderly listeners." *J. Speech Hear. Res.* **40**, 423–431.

Gordon-Salant, S., Fitzgibbons, P. J., and Friedman, S. A. (2007). "Recognition of time-compressed and natural speech with selective temporal enhancements by young and elderly listeners." *J. Speech Hear. Res.* **50**, 1181–1193.

Gustafsson, H. A., and Arlinger, S. D. (1994). "Masking of speech by amplitude-modulated noise." *J. Acoust. Soc. Am.* **95**, 518–529.

Gygi, B., and Shafiro, V. (2012). "Spatial and temporal factors in a multi-talker dual listening task." *Acta Acust. Acust.* **98**, 142–157.

Gygi, B., and Shafiro, V. (2014). "Spatial and temporal modifications of multitalker speech can improve speech perception in older adults." *Hear. Res.* **310**, 76–86.

Heiman, G. W., Leo, R. J., Leighbody, G., and Bowler, K. (1986). "Word intelligibility decrements and the comprehension of time-compressed speech." *Percept. Psychophys.* **40**(6), 407–411.

Huggins, A. W. (1964). "Distortion of the temporal pattern of speech: Interruption and alternation." *J. Acoust. Soc. Am.* **36**, 1055–1064.

Huggins, A. W. (1975). "Temporally segmented speech." *Percept. Psychophys.* **18**, 149–157.

Humes, L. E., and Dubno, J. R. (2010). "Factors affecting speech understanding in older adults," in *The Aging Auditory System: Perceptual Characterization and Neural Bases of Presbycusis*, Springer Handbook of Auditory Research, edited by S. Gordon-Salant, R. D. Frisina, A. Popper, and D. Fay (Springer, Berlin), Chap. 8, pp. 211–258.

Janse, E. (2009). "Processing of fast speech by elderly listeners." *J. Acoust. Soc. Am.* **125**(4), 2361–2373.

Janse, E., and Jesse, A. (2014). "Working memory affects older adults' use of context in spoken word recognition." *Q. J. Exp. Psych.* **67**, 1842–1862.

Jerger, J., Jerger, S., and Pirozzolo, F. (1991). "Correlational analysis of speech audiometric scores, hearing loss, age, and cognitive abilities in the elderly." *Ear Hear.* **12**, 103–109.

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. R., and Humes L. E. (2010). "Effects of age and hearing loss on the recognition of interrupted words in isolation and in sentences." *J. Acoust. Soc. Am.* **131**, 1434–1448.

Kupryjanow, A., and Chyżewski, A. (2012). "Methods of improving speech intelligibility for listeners with hearing resolution deficit." *Diagnos. Pathol.* **7**, 129.

Lee, F. F. (1971). "Time compression and expansion of speech by the sampling method." *J. Audio Eng. Soc.* **20**, 738–742.

Mattys, S. L. (1997). "The use of time during lexical processing and segmentation: A review." *Psychonomic. Bull. Rev.* **4**, 310–329.

Miller, G. A., Heise, A., and Lichten, W. (1951). "The intelligibility of speech as a function of the context of the test materials." *J. Exp. Psych.* **41**, 329–335.

Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech." *J. Acoust. Soc. Am.* **22**, 167–173.

Miller, P., and Wingfield, A. (2010). "Distinct effects of perceptual quality on auditory word recognition, memory formation and recall in a neural model of sequential memory." *Front. Syst. Neurosci.* **4**, 14 (2010).

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.* **58**, 481–496.

Moore, B. (2003). "Temporal integration and context effects in hearing." *J. Phonetics* **31**, 563–574.

Moulines, E., and Charpentier, F. (1990). "Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones." *Speech Comm.* **9**, 453–467.

Nagaraj, N. K., and Knapp, A. N. (2015). "No evidence of relation between working memory and perception of interrupted speech in young adults." *J. Acoust. Soc. Am.* **138**, EL145–EL150.

Nahum, M., Nelken, I., and Ahissar, M. (2008). "Low-level information and high-level perception: The case of speech in noise." *PLoS Biol.* **6**, e126.

Nejime, Y., and Moore, B. C. J. (1998). "Evaluation of the effect of speech-rate slowing on speech intelligibility in noise using a simulation of cochlear hearing loss." *J. Acoust. Soc. Am.* **103**, 572–576.

Nelson, P. B., and Jin, S. (2004). "Factors affecting speech understanding in gated interference: Cochlear implant users and normal-hearing listeners." *J. Acoust. Soc. Am.* **115**, 2286–2294.

Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise." *J. Acoust. Soc. Am.* **95**, 1085–1099.

- Peterson, G., and Lehiste, I. (1962). "Revised CNC lists for auditory tests," *J. Speech Hear. Dis.* **27**, 62–70.
- 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**, 593–608.
- Pichora-Fuller, M. K., and Souza, P. E. (2003). "Effects of aging on auditory processing of speech," *Int. J. Audiol.* **42**, S11–16.
- Piquado, T., Benichov, J., Brownell, H., and Wingfield, A. (2012). "The hidden effect of hearing acuity on speech recall, and compensatory effects of self-paced listening," *Int. J. Audiol.* **51**, 576–583.
- Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, O., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., and Rudner, M. (2013). "The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances," *Front. Syst. Neurosci.* **7**, 31.
- Saija, J. D., Akyürek, 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. (1996). "The processing-speed theory of adult age differences in cognition," *Psychol. Rev.* **103**, 403–428.
- Samuel, A. G. (1991). "Perceptual degradation due to signal alternation: Implications for auditory pattern processing," *J. Exp. Psychol. Hum. Percept. Perform.* **17**, 392–403.
- Schneider, B. A., Daneman, M., and Murphy, D. R. (2005). "Speech comprehension difficulties in older adults: Cognitive slowing or age-related changes in hearing?," *Psychol. Aging* **20**, 261–271.
- Schneider, B. A., Pichora-Fuller, M. K., and Daneman, M. (2010). "The effects of senescent changes in audition and cognition on spoken language comprehension," in *The Aging Auditory System: Perceptual Characterization and Neural Bases of Presbycusis*, Springer Handbook of Auditory Research, edited by S. Gordon-Salant, R. D. Frisina, A. Popper, and D. Fay (Springer, Berlin), Chap. 7, pp. 167–210.
- Schmitt, J. F. (1983). "The effects of time compression and time expansion on passage comprehension by elderly listeners," *J. Speech Hear. Res.* **26**, 373–377.
- Schmitt, J. F., and McCroskey, R. L. (1981). "Sentence comprehension in elderly listeners: The factor of rate," *J. Gerontol.* **36**, 441–445.
- Shafiro, V., Sheft, S., and Risley, R. (2011a). "Perception of interrupted speech: Effects of dual-rate gating on the intelligibility of words and sentences," *J. Acoust. Soc. Am.* **130**, 2076–2087.
- Shafiro, V., Sheft, S., and Risley, R. (2011b). "Perception of interrupted speech: Cross-rate variation in the intelligibility of gated and concatenated sentences," *J. Acoust. Soc. Am.* **130**, EL108–EL114.
- 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.
- Sheft, S., Shafiro, V., Lorenzi, C., McMullen, R., and Farrell, C. (2012). "Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech perception," *Ear. Hear.* **33**, 709–720.
- Shinn-Cunningham, B. G., and Best, V. (2008). "Selective attention in normal and impaired hearing," *Trends Amplif.* **12**, 283–299.
- Studebaker, G. A. (1985). "A 'rationalized' arcsine transform," *J. Speech. Hear. Res.* **28**, 455–462.
- Tulving, E., and Gow, C. (1963). "Stimulus information and contextual information as determinants of tachistoscopic recognition of words," *J. Exp. Psychol.* **66**, 319–327.
- Vaughan, N. E., Furukawa, I., Balasingam, N., Mortz, M., and Fausti, S. A. (2002). "Time-expanded speech and speech recognition in older adults," *J. Rehabil. Res. Dev.* **39**, 559–566.
- Verschuure, J., and Brocaar, M. P. (1983). "Intelligibility of interrupted meaningful and nonsense speech with and without intervening noise," *Percept. Psychophys.* **33**, 232–240.
- Wang, X., and Humes, L. E. (2010). "Factors influencing recognition of interrupted speech," *J. Acoust. Soc. Am.* **85**, 2100–2111.
- Wilson, R. H. (2014). "Variables that influence the recognition performance of interrupted words: Rise-fall shape and temporal location of the interruptions," *J. Am. Acad. Audiol.* **25**, 688–696.
- Wilson, R. H., McArdle, R., Betancourt, M. B., Herring, K., Lipton, T., and Chisolm, T. H. (2010). "Word-recognition performance in interrupted noise by young listeners with normal hearing and older listeners with hearing loss," *J. Am. Acad. Audiol.* **21**, 90–109.
- Wingfield, A., and Tun, P. A. (2007). "Cognitive supports and cognitive constraints on comprehension of spoken language," *J. Am. Acad. Audiol.* **18**, 548–558.