Effects of age and hearing loss on the intelligibility of interrupted speech^{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*

Brian Gygi^{c)}

*National Institute for Health Research, Nottingham Hearing Biomedical Research Unit, Nottingham,
United Kingdom*

(Received 10 February 2014; revised 21 December 2014; accepted 27 December 2014)

How age and hearing loss affect the perception of interrupted speech may vary based on both the physical properties of preserved or obliterated speech fragments and individual listener characteristics. To investigate perceptual processes and interruption parameters influencing intelligibility across interruption rates, participants of different age and hearing status heard sentences interrupted by silence at either a single primary rate (0.5–8 Hz; 25%, 50%, 75% duty cycle) or at an additional concurrent secondary rate (24 Hz; 50% duty cycle). Although age and hearing loss significantly affected intelligibility, the ability to integrate sub-phonemic speech fragments produced by the fast secondary rate was similar in all listener groups. Age and hearing loss interacted with rate with smallest group differences observed at the lowest and highest interruption rates of 0.5 and 24 Hz. Furthermore, intelligibility of dual-rate gated sentences was higher than single-rate gated sentences with the same proportion of retained speech. Correlations of intelligibility of interrupted speech to pure-tone thresholds, age, or measures of working memory and auditory spectro-temporal pattern discrimination were generally low-to-moderate and mostly nonsignificant. These findings demonstrate rate-dependent effects of age and hearing loss on the perception of interrupted speech, suggesting complex interactions of perceptual processes across different time scales.

© 2015 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4906275>]

[DB]

Pages: 745–756

I. INTRODUCTION

In daily life, speech perception typically takes place in the presence of more intense extraneous sounds, which often render much of the speech signal inaudible. A remarkable and numerous replicated finding of mid twentieth-century research was that under some conditions, accurate speech perception is possible even if half of the original speech signal is periodically masked by noise or replaced by silence (Miller and Licklider, 1950; Powers and Speaks, 1973; Huggins, 1975; Powers and Wilcox, 1977; Nelson *et al.*, 2003; Nelson and Jin, 2004; Buss *et al.*, 2009; Jin and Nelson, 2010; Saija *et al.*, 2014). Despite these physical constraints, listeners are able to exploit the inherent perceptual redundancies of naturally produced speech and use the remaining low-level speech cues to obtain high-level linguistic information. To accomplish this task, listeners can rely on existing lexical, semantic, and syntactic knowledge of the language. In addition, a set of interdependent cognitive processes, including working memory, attention, and speed of processing are also known to play an important role in the process of integrating low-level speech cues with high-level

linguistic knowledge. Thus a variety of factors, including peripheral auditory status, cognitive abilities, and linguistic skills can influence the perception of interrupted speech (Benard *et al.*, 2014; Jin and Nelson, 2010).

The contributions of different factors can become confounded when speech is interrupted at a single fast rate. At slow rates of about 0.5–1 Hz with a 50% duty cycle (i.e., the proportion of speech available within each interruption cycle), the durations of the remaining and discarded speech fragments generally exceed that of whole words.¹ In this case, speech perception might be expected to rely more on centrally based contextual semantic processing to fill in the missing information with the perception of retained words requiring relatively little effort. Alternatively, at faster interruption rates (e.g., 8 Hz or above) that sample each (sub) phonemic segment within a word, speech processing may involve temporal smoothing over small gaps in the signal and rely to a greater extent on more basic auditory processing abilities. However, when sampling speech information with a single fast interruption rate, both low-level acoustic cues and higher-order contextual syntactic and semantic cues become increasingly available, obscuring the relative involvement of the associated perceptual processes.

To untangle contributions of speech cues available in the brief speech fragments produced by fast interruption rates, Shafiro *et al.* (2011) used a dual-rate gating paradigm in which speech was gated twice, first with a slow primary rate and, subsequently, with a faster secondary rate. The results

^{a)}Portions of the data were presented at Interspeech 2011 and the 165th Meeting of the Acoustical Society of America.

^{b)}Author to whom correspondence should be addressed. Electronic mail: Valeriy_Shafiro@rush.edu

^{c)}Also at: Veterans Affairs Northern California Health Care System, Martinez, CA.

revealed that the ability to perceptually integrate small speech fragments produced by the fast rates was strongly influenced by the duration of the larger speech fragments produced by the slower rates. The present study utilized the same dual-rate gating paradigm as used in [Shafiro *et al.* \(2011\)](#) to (1) investigate the effects of age and hearing loss on intelligibility of gated speech and (2) examine the relationship of basic spectro-temporal processing abilities and working memory—factors known to affect speech perception in noise—with the perception of interrupted speech.

A. Dual-rate gating approach to perception of interrupted speech

An overarching goal of most experiments with interrupted speech is to elucidate factors that may affect speech perception in noisy listening environments. Although the impetus of the early work (e.g., [Miller and Licklider, 1950](#)) was speech transmission through a bandlimited communication channel, later research explored specific stimulus parameters and listener characteristics that can affect the perception of partially audible speech. As an alternative to semi-stochastic masking properties of most real world noises, experiments with interrupted speech typically mask speech with modulated broadband noise or periodic replacement of speech intervals with silence.

Intelligibility of interrupted speech is generally thought to rely on the ability to detect and perceptually integrate multiple audible fragments of the original speech (sometimes referred to as glimpses) into existing memory templates, using some kind of an “intelligent warping” process ([Moore, 2003](#); [Cooke, 2006](#)). Because perception of speech takes place on several different time scales, the nature of the acoustic information being integrated and perceptual processes involved are likely to vary with the duration of the retained and missing speech fragments ([Giraud and Poeppel, 2012](#); [Ghitza, 2011](#); [Ghitza and Greenberg, 2009](#); [Rosen, 1992](#); [Howard-Jones and Rosen, 1993](#)). For instance, (sub) segmental acoustic cues to phonemes involve spectro-temporal variation on the order of tens of milliseconds (e.g., formant transitions, noise duration), relying on listeners’ abilities to detect fast changes in temporal fine structure (TFS), fundamental frequency, and the envelope. On the other hand, perception of intonation or post-perceptual processing of syntactic and semantic information may involve a broader set of listeners’ linguistic skills and higher-order cognitive processes such as working memory, thus requiring a longer processing time on the order of hundreds of milliseconds (for a review, see [Mattys, 1997](#)). Ultimately, successful speech perception requires integration of linguistically salient acoustic information across multiple time scales. However, in interrupted speech, rate-dependent contributions of different factors, described in the preceding text, may be compounded when only a single gating rate is used.

To separate factors involved in perceptual processing of interrupted speech at different rates, [Shafiro *et al.* \(2011\)](#) compared intelligibility of speech gated with either a single primary rate and a 50% or 25% duty cycle to that obtained

with speech gated with two rates: First, a slow primary rate with a 50% duty cycle and then also a secondary faster rate with a 50% duty cycle ([Fig. 1](#)). The results indicated that the ability to use speech information that remained after gating with a secondary rate critically depended on the slower primary interruption rate. For some primary interruption rates (i.e., 0.5 and 8 Hz), intelligibility in dual-rate conditions closely approximated that of single-rate gated speech with a 50% duty cycle. On the other hand, for middle primary rates of 2–4 Hz, which contained 25% of the original signal, intelligibility of dual-rate gated speech was considerably lower than that obtained with a single rate and a 50% duty cycle. Nevertheless, even in these conditions, intelligibility of dual-rate gated speech was significantly higher than that with a single rate and a 25% duty cycle despite the equivalent proportion of total speech per interruption cycle. This suggests that following secondary gating with a fast rate, listeners were able to utilize brief speech fragments to partially reconstruct information contained in the larger speech fragments produced by a slower primary rate.

In sum, when a secondary interruption rate is applied in dual-rate gating to speech already gated by the primary rate, it further obliterates some of the remaining low-level speech cues and reduces the total proportion of speech duration. On the other hand, the high-level information including semantic, syntactic, and prosodic or any other linguistic contextual information may also not exceed that which was available after initial gating with a single primary rate alone. Thus the dual-rate gating approach makes it possible to examine factors involved in the perceptual processing of speech interrupted at a fast secondary rate, while controlling access to higher-level information that may aid speech perception through systematic manipulation of the slower primary rate.

B. Age and hearing loss and the perception of interrupted speech

As they age, listeners experience increasing difficulty understanding speech in background noise. These difficulties are well documented and have been shown to be associated with a number of higher-level cognitive factors, which refer to listeners’ existing knowledge and the ability to store and manipulate information to complete perceptual tasks as well as presbycusis-related deficits in auditory acuity ([Akeroyd, 2008](#); [Schneider *et al.*, 2010](#); [Salthouse, 1996](#); [Humes and Dubno, 2010](#); [Pichora-Fuller and Souza, 2003](#); [Jin and Nelson, 2010](#); [Kidd and Humes, 2012](#)). Similar to speech in modulated background noise, perception of gated speech involves integration of temporally distributed fragments of acoustic information. It could thus be expected that a similar set of higher-level cognitive, linguistic, and lower-level basic auditory factors may play a role in the perception of both types of interrupted speech: Speech-in-noise and periodically gated speech. This expectation is supported by previous reports of correlation between intelligibility of speech interrupted by periodic gating or modulated noise ([Saija *et al.*, 2014](#); [Jin and Nelson, 2010](#)).

Chief among cognitive factors shown to influence the perception of speech under adverse listening conditions by

FIG. 1. Illustrations of dual-rate gated speech. The top panel (a) shows the continuous waveform of a sentence—"The birch canoe slid on the smooth planks"; the two middle panels below it show this sentence gated, respectively, at a single rate of 24 Hz (b) and 2 Hz (c), both with a 50% duty cycle, as was done in the S50 condition. The next panel (d) shows the sentence gated concurrently with both 2 and 24 Hz as was done in the D25 condition. When concurrent, the faster gating rate of 24 Hz affects only the speech remaining after gating at the slower primary rate of 2 Hz, leaving 25% of the original speech signal. Finally, the bottom panel (e) shows the sentence gated at a single rate and 2 Hz with a 25% duty cycle as was done in the S25 condition. In this example, the total proportion of preserved speech is equivalent between D25 and S25 condition, although, as can be seen, D25 condition produces a temporally broader sampling of the speech signal.

older adults with either normal or impaired hearing is working-memory capacity and the ability to use sentence context (Schvartz *et al.*, 2008; Akeroyd, 2008; Gordon-Salant and Fitzgibbons, 1997; Lunner, 2003; Wingfield and Tun, 2007). In a seminal study, Pichora-Fuller *et al.* (1995) examined the recognition of words that were either highly predictable from the preceding sentence context or not. The comparison of psychometric functions for high- and low-predictability words obtained across different signal-to-noise ratios (SNR) in young normal hearing (YNH), older normal hearing (ONH), and older hearing-impaired (OHI) adults demonstrated that under increasing noise masking, ONH and OHI adults tended to rely on semantic contextual cues to a greater extent than YNH adults, a finding confirmed in numerous later studies (cf. Humes *et al.*, 2007). ONH and OHI adults also appear to rely to a greater extent than YNH adults on their general knowledge of regularity in linguistic structures and semantic context to augment speech perception (Pichora-Fuller and Souza, 2003). This conclusion has been further supported by more recent work that demonstrated that perception of interrupted speech is correlated with individuals' linguistic skills such as lexical knowledge measured by the Peabody Picture Vocabulary Test (Benard *et al.*, 2014).

Recently, Kidd and Humes (2012) examined the effect of age, hearing loss, and sentence context on the intelligibility of interrupted speech. Periodically interrupted words were either presented alone or embedded in the end of sentences with high or low semantic context using similar

materials (i.e., SPIN-R) as in Pichora-Fuller *et al.* (1995) study. The stimuli were individually spectrally shaped for hearing-impaired participants to minimize the influence of basic audibility. The results revealed the effects of age (i.e., YNH listeners performed better overall than ONH or OHI listeners) and sentence context, so that listeners benefited from additional top-down information regardless of age or hearing status. However, the largest overriding factor in the intelligibility of interrupted words was the total proportion of speech duration preserved. Importantly, this stimulus-level factor also interacted with higher-order characteristics of the stimuli, such as contextual predictability of sentences and lexical difficulty of individual words, indicating that the involvement of higher-order factors may be modulated by the availability and integrity of lower-level sensory information.

In addition to stimulus parameters, the integrity of sensory information is affected by spectro-temporal processing abilities. Previous research suggests strong links between a number of basic auditory abilities in older and hearing-impaired listeners and the perception of speech interrupted by silence or modulated noise. These include forward masking, TFS processing, low-to-mid frequency audibility, and auditory filter bandwidth (Jin and Nelson, 2006, 2010; Leger *et al.*, 2012; Sheft *et al.*, 2012a). The age-related deficits in temporal processing, however, appeared most detrimental at high interruption rates, while having less of an effect at slower rates (Grose *et al.*, 2009; Füllgrabe *et al.*, 2006). Nevertheless, a consistent finding of previous research is that

unless specific steps are taken to compensate for reduced audibility, audiometric pure-tone thresholds remain the highest predictor of the intelligibility of speech in noise among older and hearing-impaired listeners, accounting for most of the variance in their performance (Jerger *et al.*, 1991; van Rooij and Plomp, 1992; Humes and Dubno, 2010).

C. Present study

The present study investigated the effects of age and hearing impairment on the perception of speech gated with either one rate or two concurrent rates. Based on previous work (Kidd and Humes, 2012; Jin and Nelson, 2006; Saija *et al.*, 2014), it was expected that compared to YNH adults tested under the same conditions by Shafiro *et al.* (2011), overall intelligibility of interrupted speech would be reduced for ONH and OHI adults for both single- and dual-rate conditions. Of particular interest, however, was the ability of ONH and OHI adults to perceive fine-grained sub-phonemic speech fragments produced by a fast gating rate, while embedded into larger size speech fragments produced by a slower concurrent rate. If perceptual integration of such small fragments is dependent on audibility, spectro-temporal processing abilities and working memory capacity, factors known to be negatively affected by age and hearing impairment, ONH and OHI adults will not be able to use information in the fast secondary rate as effectively as YNH listeners. On the other hand, if similar to YNH listeners, ONH and OHI listeners are able to integrate fine-grained sub-phonemic speech fragments produced at fast gating rates, their performance in dual-rate conditions should be superior to that in comparable single-rate conditions.

Furthermore, working memory and basic auditory abilities, factors that have been shown to be involved in the perception of speech in noise, can be also expected to be involved in the perception of interrupted speech because in both cases, speech perception requires some kind of “perceptual reconstruction” based on partially audible fragments of the original signal. The importance of different perceptual processes, represented by correlation magnitudes, might vary across rates. Specifically, working memory may play a more prominent role at low gating rates, which preserve word-size speech fragments, with involvement decreasing at faster rates in which basic spectro-temporal processing abilities play a greater role. Because older listeners have been shown to effectively rely on working memory in the perception of degraded speech, their performance with gating at a single low rate may be expected to correlate with working memory scores. In dual-rate conditions, however, basic spectro-temporal processing abilities may interact with working memory, resulting in high correlations with both working memory and auditory tests.

II. METHOD

A. Subjects

All listeners were tested with the four types of materials described in the following text: Interrupted speech sentences, speech-in-noise tests, tests of basic auditory abilities, and

working-memory tests. Thirty-one participants were separated into two groups based on their hearing status: A pure-tone average (PTA) threshold in the better hearing ear based on 0.5-, 1-, 2-kHz thresholds (Fig. 2). ONH participants were 17 adults (mean age 68.2 yr; range: 61–87 yr, 12 females) with a mean better-ear PTA of 15.6 dB hearing level (HL) (range: 10–22 dB HL) and mean binaural PTA of 17.3 dB HL (range: 12–26 dB HL). Although labeled normal-hearing based on better-ear PTA, all but four participants in the ONH group exhibited at least a mild hearing loss at 4 kHz, and all but three participants at 8 kHz. OHI participants were 15 adults (mean age, 70.2 yr; range: 60–85 yr, 8 females) with a mild-to-moderate sloping hearing loss confirmed as sensorineural by bone-conduction thresholds and tympanometry. Their mean better-ear PTA was 31.1 dB HL (range: 25–41 dB HL) and mean binaural PTA was 34.5 dB HL (range: 25–57 dB HL). All study participants achieved a score of 25 or greater on the Mini Mental Status Examination (Folstein *et al.*, 1975) and spoke English as their first and primary language. In addition, data from 19 YNH participants who previously took part in a similar study by Shafiro *et al.* (2011) were included in analysis. They were normal-hearing adults (14 females) with hearing thresholds at or below 20 dB HL between 250 Hz and 8 kHz. Their mean age was 25 yr (SD = 6.7).

FIG. 2. Better-ear audiometric thresholds for older-normal hearing (ONH) participants in the top panel, and older hearing-impaired (OHI) participants in the bottom panel.

B. Stimuli, design, and procedure

Gated speech stimuli were based on HINT sentences spoken by a male talker (Nilsson *et al.*, 1994). These sentences have a simple syntactic structure and abundant semantic cues that augment their perception for ONH and OHI individuals. All sentences were gated with either one (primary) rate or two (primary and secondary) rates (see Fig. 1), following the method of Shafiro *et al.* (2011). All sentences were gated at a primary rate (0.5, 2, 4, or 8 Hz) by multiplying the original stimuli with a rectangular gating window and a random starting phase. Duty cycles of 25%, 50%, and 75%, were used for each single-rate gated sentence, with condition labels of S25, S50, S75, respectively. Single-rate gated sentences with a duty cycle of either 50% or 75% were gated at a secondary rate of 24 Hz and a 50% duty cycle to obtain dual-rate stimuli. With dual-rate gating, the total proportion of the original speech duration remaining within each interruption cycle was reduced in most dual-rate conditions to 25% or 37.5% for primary-rate duty cycles of 50% and 75%, respectively (with conditions labeled as D25 and D37).² An additional control condition with a single 24-Hz gating rate and a 50% duty cycle was used to ensure that speech interrupted at this single rate was highly intelligible for all listener groups.

The order of experimental conditions was randomized across sentence lists and subjects. 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. Tested in a double-walled sound-attenuated room, listeners heard all stimuli through Sennheiser 250 II headphones at 70 dB sound pressure level (SPL). After listening to each sentence, they were asked to repeat what they heard. Gated speech stimuli were processed and delivered using a custom interface developed in MATLAB 7.0 via an external Edirol UA-25 sound board. The intelligibility score in each condition was based on the total number of keywords repeated correctly for each list divided by the total number of keywords in the list (about 50 per list).

Speech-in-noise tests were (1) Quick-Speech-in-Noise Test or QuickSIN (Killion *et al.*, 2004) and (2) the Speech-in-Noise or SPIN-R (Elliot, 1995). Each QuickSIN list contains six sentences that are masked by four-talker babble. The SNR of the stimulus decreases in 5-dB steps from 25 to 0 dB. Four QuickSIN sentence lists were administered in an open-set format to obtain an average SNR loss, which represents the normalized SNR necessary to achieve 50% word recognition accuracy. The SPIN-R sentence test was administered in a closed-set response format with 50 response options per list. Two lists were administered to every participant. Every SPIN list included 50 sentences with the last word of each sentence being either easily predictable given prior sentence context for 25 of the sentences (high predictability) or difficult to predict for the remaining 25 sentences (low predictability). All SPIN sentences were presented in wideband stationary noise low-pass filtered at 8 kHz at an SNR of 2 dB.

Tests of basic auditory ability were assessed following the procedures of Sheft *et al.* (2012a) with two psychoacoustic tests that have been previously shown to be associated with perception of speech in noise (Sheft *et al.*, 2012b) or vocoded speech (Shafiro *et al.*, 2012). In the first test, discrimination of static spectral patterns was assessed using wideband stimuli (0.2–8.0 kHz) the amplitude spectra of which were sinusoidally rippled in terms of the logarithms of both frequency and amplitude. Ripple density was 1.5 cycles per octave with a peak-to-trough difference of 30 dB. In the cued two-interval forced-choice (2IFC) procedure, the phase of the sinusoidal spectral ripple of the standard stimulus was randomized each discrimination trial with the task to detect a change in ripple starting phase. The 500-ms rippled stimuli were shaped with a 50-ms rise/fall time, passed through a speech-shape filter emphasizing the mid frequencies, and presented to listeners at 80 dB SPL.

The second test measured the ability to discriminate 1-kHz pure tones frequency modulated by different samples of 5-Hz low-pass noise with peak frequency excursion fixed at 400 Hz. The resulting stimuli thus contained a stochastic pattern of dynamic frequency deviation within the range of 600–1400 Hz. With noise modulators drawn from a common sampling distribution, discrimination of such stimuli can rely only on the temporal pattern of frequency deviation. The 500-ms modulated stimuli were temporally centered in 1000-ms maskers with thresholds measured in terms of the SNR needed to just discriminate pattern of frequency fluctuation. Maskers were speech-shaped wideband noise that was processed to include slow random variations in local fine-structure periodicities and loudness to approximate modulation characteristics similar to speech. Both signals and maskers were shaped with a 50-ms rise/fall time. In the task, masker level was fixed at 80 dB SPL with the level of the FM tones varied to estimate the threshold SNR. The presentation levels for both psychoacoustic tasks were kept consistent with that used during earlier work in which these tests were developed and normative data obtained with a large sample ONH and OHI listeners (Sheft *et al.*, 2012a; Sheft *et al.*, 2012b).

Working-memory tests included auditory forward and backward digit span tests and a reading span test. In the digit span tests (Wechsler, 1997), the listener repeated strings of digits of increasing length in the order they were presented—forward, or in the reversed order—backward. The digits were spoken by the experimenter, who was seated directly in front of the listener at a distance of approximately 1 m at a rate of one digit per second. There were four trials for each string length. The test score for both digit span tests was the highest number of digits the participant recalled prior to missing two successive digit strings of the same length.

In the reading span test (Rönnberg, 1990) designed to tax memory storage and item processing simultaneously, listeners reported either the first or last word in a set of sentences with the number of sentences in the set increasing gradually from three to six. During the test, each sentence was displayed on a computer screen, and the subject was asked to say if the sentence made sense (i.e., was semantically coherent, such as “The girl brushed her teeth”) or not (e.g., “The train sang a song”). After a new set of sentences

FIG. 3. Mean sentence intelligibility across gating rates with gating method the parameter. Performance of each listener group is shown in a separate panel with YNH, ONH, and OHI in the left, middle, and right panels, respectively. The curves in each panel depict accuracy functions for gating with a single rate with a 50% (top) and 25% duty cycle (bottom). The middle curve indicates intelligibility of sentences gated with a dual-rate: The same primary rate and a 50% duty cycle and, concurrently gated at a 24-Hz secondary rate with a 50% duty cycle. Open circle indicate performance in the control condition with speech gated at 24 Hz and a 50% duty cycle.

was presented, the experimenter asked the subject to repeat either the first or the last word of each sentence in the batch. The order in which the subject was asked to report the first or last word was randomized across sentence sets. The final score was the percent correct of correctly recalled words (either first or last for each batch) from the total number of words presented.

III. RESULTS

Percent correct intelligibility scores, converted to rationalized arcsine-transformed units, (i.e., RAU, [Studebaker](#),

[1985](#)) are shown in Figs. 3 and 4. Along with the current data, YNH results from [Shafiro *et al.* \(2011\)](#) are shown in these figures. Separate analyses of variance (ANOVA) were performed (a) to compare the results of the present ONH and OHI groups to those obtained in the same conditions with 19 YNH adults tested by [Shafiro *et al.* \(2011\)](#) and (b) to compare performance of the ONH and OHI groups in the S75 and D37 conditions (YNH data were not collected in these conditions by [Shafiro *et al.*](#)). Finally, listener performance on the speech-in-noise tests, working-memory tests, and tests of spectro-temporal processing abilities were examined in relationship to the gated-speech conditions.

FIG. 4. Mean sentence intelligibility across gating rates with listener group the parameter. The panels show performance for interruption conditions that preserved different proportions of the original speech signal. From left to right: Single-rate gated speech with a 25% duty cycle (S25), dual-rate gated speech containing ~25% of the original signal (D25), single-rate gated speech with a 50% duty cycle (S50), dual-rate gated speech containing ~37% of the original signal (D37), and single-rate gated speech with a 75% duty cycle (S75).

Compared to the results from YNH adults, overall intelligibility of interrupted speech was reduced in ONH and, even more so, in OHI listeners. The ANOVA revealed significant main effects for all three factors: (1) Listener group, consisting of three levels: YNH, ONH, OHI, [$F(2, 45) = 39.183, p < 0.001, \eta^2_p = 0.635$], (2) primary rate, consisting of four levels: 0.5, 2, 4, 8 Hz, [$F(3, 135) = 159.36, p < 0.001, \eta^2_p = 0.780$], and (3) interruption method, consisting of three levels: S25, S50, D25 [$F(2, 90) = 402.21, p < 0.001, \eta^2_p = 0.815$]. Planned comparisons further indicated that all three groups, YNH, ONH, and OHI, and all three interruption methods were significantly different from each other (p values for all comparisons < 0.01).

In addition, there were significant interactions: A two-way interaction between primary rate and interruption method [$F(6, 270) = 23.546, p < 0.001, \eta^2_p = 0.359$], a two-way interaction between primary rate and listener group [$F(6, 135) = 10.738, p < 0.001, \eta^2_p = 0.323$], a two-way interaction between listener group and interruption method [$F(4, 90) = 4.7881, p = 0.002, \eta^2_p = 0.095$], and finally a significant three-way interaction of group, interruption rate, and interruption method [$F(12, 270) = 3.9650, p < 0.001, \eta^2_p = 0.159$]. These interactions demonstrate that changes in primary rate and interruption method differentially affected the three listener groups. They also reflect nonmonotonic changes in intelligibility across interruption rates, which, with some variation in magnitude, were observed in all three groups.

Specifically, as shown in Fig. 3, for most interruption methods, all three groups had a similar U-shaped intelligibility pattern across the primary rates with local minima around 2–4 Hz. Relative to their performance at the lowest primary rate condition of 0.5 Hz, intelligibility declined in the 2- to 4-Hz interruption region and then improved with higher primary rates. There were several exceptions to this pattern when intelligibility improved continuously with rate without a significant drop at 2 Hz: YNH listeners in S50, ONH listeners in D37, and ONH and OHI listeners in S75. Overall, the U-shaped intelligibility pattern seems more likely to arise in interruption conditions with greater reduction in preserved speech (S25, D25, D37) and for listener groups with poorer hearing sensitivity (ONH, OHI).

The magnitude of performance differences for the three groups also varied with the interruption rate (Fig. 4). As expected, at the lowest interruption rate of 0.5 Hz, intelligibility differences across the three groups, while significant, were the smallest. However, as interruption rate increased, group differences also increased up to the 8-Hz primary gating rate. This may indicate that for sentences interrupted at the rate of 0.5 Hz, ONH and OHI listeners, similar to YNH, were able to use information remaining in the relatively large (word-length and longer) fragments in essentially the same way, possibly relying on high-level contextual cues. On the other hand, the group differences decreased once again in the 24-Hz single-rate condition where all three groups achieved intelligibility over 90 RAU points: 122.3 (SD = 2.9) for YNH, 106.5 (SD = 21.1) for ONH, and 94.9 (SD = 22.1) for OHI. This indicates that with sufficient oversampling of input speech at 24 Hz, all listeners were able to utilize both low-level acoustic-phonetic and higher-level cognitive cues. Nevertheless these 24-Hz single-

rate results also demonstrate that while performance is quite high for all three groups, group differences still persist, suggesting a role of both audibility and age. It is also possible that low-level signal distortions introduced by fast gating rates prevent ONH and OHI listeners from accessing higher-level linguistic and semantic contextual cues. On the other hand, similar to YNH listeners, for both ONH and OHI listeners, the intelligibility in most dual-rate conditions was higher than that obtained with the corresponding single rate with the same proportion of speech preserved.

A separate ANOVA on the intelligibility of two additional interruption methods, single-rate, S75, and the corresponding dual-rate, D37, also revealed main effects of interruption method [$F(1, 29) = 78.034, p < 0.001, \eta^2_p = 0.729$], primary gating rate [$F(3, 87) = 66.458, p < 0.001, \eta^2_p = 0.696$], and group [$F(1, 29) = 10.492, p = 0.003, \eta^2_p = 0.626$]. Overall intelligibility of OHI listeners was again significantly lower than that of ONH listeners. Within each group, intelligibility with S75 and D37 interruption methods once again indicated a similar rate-dependent intelligibility pattern across primary rates. Intelligibility differences between a single-rate condition with a 75% duty cycle vs the corresponding dual-rate condition, while significant for all rates except 8 Hz, were again the highest around 2–4 Hz and smallest at the lowest and highest primary rates. Remarkably, for both groups, the overall intelligibility of dual-rate gated speech that contained about 37% of the original signal was higher than that for a single rate and a 50% duty cycle even though the latter contained a greater amount of the original speech (Fig. 4). Specifically, for ONH listeners, the average intelligibility across rates in the D37 condition was 86.9 (SD = 20.1) vs 62.1 (SD = 21.1) RAU points in the S50 condition, while for OHI listeners, the average intelligibility across rates in the D37 condition was 66.2 (SD = 30.2) vs 44.5 (SD = 29.1) in the S50 condition. Independent sample t -tests revealed that these differences were significant for each group at $p < 0.05$ level. These results indicate that listeners in all three groups, regardless of age and hearing status, were able to utilize sub-phonemic speech cues retained after additional gating with a fast secondary rate.

Correlational analyses were conducted to determine the strength of associations between the intelligibility of 21 interrupted speech conditions and listeners' other auditory, speech, and working memory abilities. Separate analyses were conducted for ONH and OHI listeners. Contrary to expectations, generally weak correlations were observed between interrupted speech and tests of spectro-temporal pattern discrimination. Similarly weak and nonsignificant correlations were observed between interrupted speech and tests of working memory, age, or hearing loss. Although correlation magnitudes were in the moderate-high range for several interrupted-speech conditions, none were significant when the significance level was adjusted for the number of comparisons with a Bonferroni correction. Some exceptions to this were observed for one nonspeech measure (i.e., FM pattern discrimination at varying SNR levels) and one speech-in-noise test (QuickSIN). As can be seen in Table I, a greater number of correlations that could have been observed by chance were evident for these two conditions. However, no pattern of correlations between any of the listener

TABLE I. Pearson correlations in older normal hearing (ONH) and older hearing impaired (OHI) listeners between 21 gating conditions with two performance metrics, which showed the greatest number of significant correlations: Modulated frequency patterns (FM), and a speech-in-noise tests—QuickSIN. For the QuickSIN test, the values on the right of the slash represent correlation magnitudes obtained after partialling out contributions of better ear PTA. Correlations significant at $p < 0.05$ are marked in bold.

Interrupted speech condition	FM (ONH)	FM (OHI)	QuickSIN (ONH)	QuickSIN (OHI)
0.5 Hz/S25	-0.13	0.35	-0.10/-0.09	-0.56 /-0.44
0.5 Hz/S50	-0.10	0.00	-0.02/0.00	-0.51 /-0.45
0.5 Hz/S75	-0.16	0.03	-0.12/-0.05	-0.48 /-0.29
0.5 Hz/D25	-0.36	0.00	-0.30/-0.29	-0.17/-0.12
0.5 Hz/D37	-0.50	0.23	-0.63 /-0.61	-0.30/-0.30
2 Hz/S25	-0.19	-0.07	0.03/0.07	-0.26/-0.11
2 Hz/S50	-0.38	-0.10	-0.18/-0.14	-0.38/-0.29
2 Hz/S75	-0.53	0.12	-0.64 /-0.63	-0.31/-0.38
2 Hz/D25	-0.03	0.32	-0.09/-0.07	-0.46/-0.36
2 Hz/D37	-0.39	0.12	-0.39/-0.38	-0.68 /-0.61
4 Hz/S25	0.09	-0.10	-0.09/-0.06	-0.19/-0.02
4 Hz/S50	-0.56	-0.03	-0.53/-0.50	-0.63 /-0.48
4 Hz/S75	-0.47	0.02	-0.61 /-0.59	-0.49/-0.49
4 Hz/D25	-0.37	0.13	-0.48/-0.47	-0.22/-0.24
4 Hz/D37	-0.56	0.23	-0.42/-0.40	-0.57 /-0.52
8 Hz/S25	-0.74	-0.16	-0.44/-0.43	-0.48/-0.56
8 Hz/S50	-0.27	0.04	-0.46/-0.44	-0.64 /-0.62
8 Hz/S75	-0.18	0.10	-0.43/-0.38	-0.28/-0.31
8 Hz/D25	-0.60	0.19	-0.64 /-0.62	-0.62 /-0.54
8 Hz/D37	-0.72	-0.12	-0.38/-0.41	-0.28/-0.26
24 Hz/S50	-0.03	0.36	-0.35/-0.29	-0.63 /-0.58

characteristics and specific interruption conditions could be discerned for any specific measure.

The results from the ONH and OHI groups on the two speech-in-noise tests revealed only small group differences in the means. The SNR-Loss scores (i.e., the normalized SNR ratio required to achieve 50% accuracy) of ONH listeners was 2.71 dB, while it was 3.9 dB for OHI listeners. This difference, however, was not statistically significant in a two-tailed t -test ($p = 0.18$). Similarly, the results of the SPIN test, averaged across two lists of 50 words each, indicated that group differences in mean scores were present only in the low-predictability sentences: 85% vs 77% correct for ONH and OHI, respectively. In high-predictability sentences, both groups scored equally high: 96% correct. However, even in the low-predictability sentences, the difference was not significant in a Student's t -test ($p = 0.11$), although it was in the direction consistent with what was previously reported for this population (Pichora-Fuller *et al.*, 1995).

Several significant correlations were observed between the intelligibility of speech interrupted by gating and that masked by speech-babble noise in the QuickSIN procedure (Table I). These correlations were mostly in the low-to-moderate range with correlation magnitudes generally in a slightly higher range for the OHI than ONH listeners. Partialling out the variance due to hearing thresholds (i.e., PTA) led to small decreases in correlation magnitudes for both groups. These correlations, although small in magnitudes, suggest shared aspects of perceptual processing for speech systematically interrupted by silence and by multi-

talker babble. However, the shared processes are likely to be procedurally constrained because correlations with SPIN scores were weak and not significant, potentially owing to the differences in the types of maskers (i.e., broadband vs speech babble) and masking procedures (i.e., variable vs fixed SNR and use of closed-set format).

IV. DISCUSSION

The present findings demonstrate that both age and hearing loss affect the perception of interrupted speech, a finding consistent with previous work (Kidd and Humes, 2012; Jin and Nelson, 2006; Saija *et al.*, 2014; Baskent *et al.*, 2010). Even with generally mild losses, OHI listeners performed significantly poorer than ONH listeners, who, in turn, performed poorer than YNH adults (Fig. 4). Nevertheless, in most conditions, dual-rate performance of each listener group exceeded single-rate performance obtained with a 25% duty cycle and in some cases approximated their single-rate performance obtained with a 50% duty cycle. Thus despite differences in overall performance, in dual-rate conditions, listeners in all three groups were able to effectively utilize sub-phonemic speech cues retained after 24-Hz secondary gating, while these cues were obliterated for single primary rates with equivalent total proportions of retained speech.

A. Effects of age and hearing loss across interruption rates

The effects of age and hearing loss were rate-dependent, increasing in magnitude with primary interruption rate up to 8 Hz before decreasing again at 24 Hz. The smallest group differences among YNH, ONH, OHI listeners were observed in the single-rate condition with the lowest interruption rate of 0.5 Hz. Similarly, small group differences were also observed in the control condition with single-rate 24-Hz gating in which high intelligibility was obtained from all three groups. In contrast, in the 2- to 8-Hz region, notwithstanding floor effects, the differences among all three groups were quite large for both single- and dual-rate conditions, reaching up to and above 50 RAU points in the 2-Hz/S50, 8-Hz/D25, and 8-Hz/S25 conditions. The better performance of YNH than ONH listeners indicates that perception of interrupted speech is (a) strongly affected by listener age even when the effects of audibility are minimized and (b) depends on specific interruption parameters that control the amount of remaining and deleted speech. In turn, better performance of ONH than OHI listeners, despite relatively mild losses of the latter, indicates that hearing impairment was also a significant contributor to the perception of gated speech.

The general pattern of group differences in the present study resembles that found in earlier research with interrupted speech in normal-hearing and hearing-impaired listeners (Miller and Licklider, 1950; Nelson and Jin, 2004; Jin and Nelson, 2010; Shafiro *et al.*, 2011). It was most recently reported by Saija *et al.* (2014) for sentences interrupted by either gating with silence or modulated noise. In that study, the differences in accuracy between YNH and ONH adults were also highest with interruption rates of 1.25–5 Hz. Such a rate-dependent nature of the effects of age and hearing loss may

result from variation in the perceptual processes that play a primary role in the processing of interrupted speech at specific rates. Linguistic and cognitive skills and word knowledge may, for instance, play a greater role at the slowest interruption rates that often retain complete uninterrupted words, while redundant speech cues may play a greater role at the highest interruption rates, which sample each phoneme in a given word.

On the other hand, the greater influence of age and hearing loss in the 2- to 4-Hz interruption region, where group differences were the highest, may indicate a general vulnerability of perceptual processing of speech fragments at these rates. Speech fragments produced at these rates are generally shorter than full words but do not sample each word more than once (cf. Miller and Licklider, 1950; Shafiro *et al.*, 2011). The absence of complete words in a sentence may result in a greater lexical uncertainty and make it more difficult to benefit from contextual cues. This, in turn, may lead to an information processing bottle-neck during the integration of low-level peripheral (or bottom-up) and higher-level central (or top-down) cues during lexical search. In the absence of strong bottom-up cues, more exhaustive lexical searches, augmented by any acoustic-phonetic, semantic and syntactic context cues, maybe needed to determine correct words (Gygi and Shafiro, 2014). The group differences observed among YNH, ONH, and OHI subjects are consistent with this interpretation. They suggest that differences in both cognitive processes associated with aging and more peripheral hearing abilities associated with hearing loss affect perception of interrupted speech.

This interpretation is also consistent with the Reverse Hierarchy Theory (RHT) account of speech perception in which perceptual accuracy deteriorates under conditions that require simultaneous access to fine-grained low-level stimulus information and high-level response categories (Nahum *et al.*, 2008; Ahissar *et al.*, 2008). Perceptual difficulty with syllabic interruption rates may also result from structural neurophysiological limitations in a hierarchical array of differentially timed neural oscillators, which concurrently process acoustic elements of speech on different time scales (Ghitza and Greenberg, 2009; Ghitza, 2011; Giraud and Poeppel, 2012). Remarkably, however, YNH listeners, unlike either ONH or OHI listeners, still maintained a gradually rising intelligibility from slow to fast rates or produced a mid-rate dip of a smaller magnitude around 2-Hz region (Fig. 4). This cross-rate pattern of performance indicates that YNH listeners were able to compensate for incomplete acoustic information for individual words to a much greater extent than ONH and, especially, OHI listeners. It might have been the result of greater access to low-level speech cues due to higher audibility and superior spectro-temporal processing abilities in YNH listeners or greater working memory capacity and processing speed that allowed YNH listeners to conduct a greater number of word searches before partial speech information decayed with some combination of these bottom-up and top-down factors also being likely. A better understanding of the factors underlying the considerable variation of speech intelligibility in the 2–4 Hz interruption rate may thus provide a greater insight into fundamental processes of speech perception.

B. Interruption parameters and speech intelligibility

The present findings have further implications with respect to the ongoing discussion of the role of stimulus parameters that affect the intelligibility of interrupted speech (cf. Kidd and Humes, 2012; Fogerty *et al.*, 2012; Stilp, 2014; Stilp *et al.*, 2013). Specifically, better performance in dual-rate conditions than those with the same or larger total proportion of speech per interruption cycle also suggests that the frequency of speech sampling and the duration of the speech interval being sampled had a greater effect on intelligibility than the total proportion of speech preserved. The observed differences between single- and dual-rate rate conditions with equivalent proportions of speech may thus reflect the way speech information is distributed in each case. This is in line with previous findings of Shafiro *et al.* (2011), who used the same parameters for dual-rate gating. However, at face value, it is inconsistent with the findings by Kidd and Humes (2012), who reported that the total proportion of the speech duration preserved was the dominant factor in accounting for intelligibility of interrupted sentences.

One reason for this contradiction may be the differences in interruption rates used in the two sets of studies. Relatively slow rates of 0.5–8 Hz were used by Shafiro *et al.* (2011) and in the present work, while faster interruption rates of approximately 10–30 Hz were used by Kidd and Humes (2012).³ While these faster rates would effectively sample the majority of phonemes within each word, the slower rates used by Shafiro *et al.* (2011) omitted many phoneme, syllable, and even word-size speech fragments. Therefore the nature of the low-level acoustic-phonetic cues and more complex linguistic information contained in the speech fragments retained across different rates and associated perceptual processing would be quite different. Thus as also suggested by Kidd and Humes (2012), the effect of the total proportion of speech preserved may be less important at slower interruption rates. Nevertheless the present results can also be accommodated for by a more relaxed version of Kidd and Humes (2012) proposal in which the proportion of total duration of a speech interval being sampled rather than the total speech duration retained after interruption is the main parameter that controls intelligibility. Furthermore, as shown by Stilp *et al.* (2013), the total information-bearing acoustic change retained in the speech signal may provide a more sensitive measure than the total proportion of speech duration retained after interruption. The present result appear to be consistent with this view because the total proportion of information-bearing acoustic changes retained in the dual-rate conditions, which sample a larger total duration of speech, would be greater than that in corresponding single-rate conditions that retain the equivalent duration of the original speech.

C. Peripheral and cognitive factors

Contrary to expectations, the findings from correlational analysis for the older normal-hearing and hearing-impaired groups in the present work do not reveal clear associations of intelligibility of interrupted speech with either age or hearing loss. Although a relationship between perception of

interrupted speech and hearing acuity in low- and mid-frequencies has been reported by Jin and Nelson (2006, 2010), in the present analysis, these correlations were in the low range and not significant when assessed separately for each group. On the other hand, the present result is consistent with the findings of Kidd and Humes (2012), who did not find significant correlations between intelligibility of interrupted speech and hearing acuity. While speech stimuli in Kidd and Humes were individually spectrally shaped for each listener to compensate for reduced audibility, the same presentation level was used for all listeners in the present study. Presumably the relatively mild losses of the present subjects did not have a strong influence on individual performance with interrupted speech even though as a group OHI performed poorer than either ONH or YNH listeners. Although age was also a significant factor in the group data, with intelligibility of ONH listeners lower than that of YNH, the lack of significant correlations with interrupted speech conditions, when only ONH and OHI listeners were included, indicates that it too may not be a strong predictor of individual performance with interrupted speech.

Past research also suggests that effective use of working-memory capacity may have a compensatory effect on speech perception in hearing-impaired listeners (Schneider *et al.*, 2010; George *et al.*, 2006; George *et al.*, 2007; Wingfield and Tun, 2007). This was not seen in the current correlational analysis using digit and reading span tests (Table I). Although several correlations were in the moderate range, no clear cross-rate pattern could be discerned. This could possibly result from the relatively mild hearing losses of the OHI listeners that did not require greater involvement of cognitive resources or from the lack of listener-specific amplification of the stimuli (Humes *et al.*, 2007). Recent research provides further support for the relevance of high-level linguistic skills such as semantic word knowledge for the perception of interrupted speech. Benard *et al.* (2014) reported moderate-to-strong associations between perception of interrupted speech and linguistic skills (measured by Peabody Picture Vocabulary Test) but found no associations with listeners' intelligence (measured by Wechsler Adult Intelligence Scale). Thus it is possible that linguistic skills, which were not measured in the present study, of the participants might have better accounted for their speech performance.

The unexpected lack of correlations between tests of spectro-temporal processing abilities with the perception of interrupted speech could, in turn, indicate a relatively minor role of these factors in the perception of interrupted speech by healthy older adults. It is also possible that the basic auditory tests in the present study, although previously shown to be associated with speech perception in noise, were not sufficiently sensitive to the perceptual processes involved in the processing of gated speech. Alternatively, they may evaluate aspects of spectro-temporal processing that are not strongly relevant to the perception of gated speech. Previous reports of associations between basic auditory abilities and speech systematically interrupted by gating with silence or modulated noise used different measures. For instance, Jin and Nelson (2010) examined equivalent rectangular bandwidths

(ERB) and auditory filter slopes at 2 and 4 kHz using a notched-noise method. These measures of spectral resolution are likely to involve some perceptual processes that are different from the current tests of frequency-pattern and spectral-ripple discrimination. In addition, in Jin and Nelson (2010), only ERB scores at 2 kHz had consistent moderate correlations with speech interrupted at 8 and 16 Hz, while correlations with other measures were weaker and more variable. Together these results suggest a relatively minor role of these aspects of auditory processing in the perception of gated speech.

Correlational analysis also indicated a possible relationship between speech interrupted by gating and masked by modulated noise. Although only some of the correlations were significant, the number of significant correlations was greater for OHI than ONH listeners. This suggests that aspects of perception of gated speech and speech interrupted by multi-talker speech babble involved overlapping perceptual processes. Previous research (Kidd and Humes, 2012; Miller and Licklider, 1950; Jin and Nelson, 2010; Saija *et al.*, 2014) demonstrated similar performance with speech periodically interrupted by either silent gaps or modulated noise. At the same time, the number of nonsignificant correlations in the present findings suggests that this relationship points to differences between speech interrupted by noise vs that interrupted by silence. Moreover, in previous work, Başkent (2010) found that the accuracy differences between speech interrupted by noise vs silence strongly correlated with speech perception for normal hearing listeners, while only weak and non significant correlations were found for listeners with mild or moderate hearing losses. This suggests that the ability to perceptually restore speech information missing from interrupted signals may require a relatively high level of audibility (i.e., mild losses), beyond which it is difficult to effectively restore missing information.

On the other hand, the complete lack of significant correlations with SPIN scores may indicate some conditions that limit these associations between interrupted speech and speech masked by noise. This might reflect the differences in the composition and scoring of the two tests and the differences in the masker properties because unlike the four-talker babble used in QuickSIN, SPIN masker was unmodulated broadband noise. As an adaptive test, QuickSIN also provides a more sensitive measure of speech-in-noise perception that likely explains its correlation with interrupted speech intelligibility. For QuickSIN, the difficulty of the test is progressively increasing with the lowering of the SNR for each sentence, while for the SPIN test, in the present experiment, the SNR stayed the same for all sentences. The graded SNR composition of QuickSIN provides greater sensitivity to listener speech-in-noise performance compared to the fixed-level SPIN task.

D. Summary

The present study investigated the influence of age and hearing loss on the intelligibility of interrupted speech. Sentences were gated with silence using either a single

primary rate (0.5–8 Hz, 25%–75% duty cycle) or at the primary rate with subsequent gating at 24 Hz and a 50% duty cycle. The findings demonstrate highly rate-dependent effects of age and hearing loss on the perception of interrupted speech. Intelligibility differences across groups were small at the slowest gating rate of 0.5 Hz but increased substantially with rate. Compared to YNH, performance of ONH and OHI adults was also more negatively affected by reduction of duty cycle with either single- or dual-rate gating. However, dual-gating with a high secondary rate of 24 Hz preserved some sub-phonemic cues for speech intelligibility that were deleted for single primary rates at equivalent total proportions of speech preserved. Importantly, listeners in all three groups, regardless of age and hearing status, were able to effectively integrate small temporally segregated sub-phonemic speech fragments in interrupted speech retained after gating with the secondary rate. On the other hand, individual performance with interrupted speech was not significantly correlated with any of the factors that might be expected to play a role: Age, hearing sensitivity, working memory (as measured by digit and reading spans), or basic spectro-temporal processing abilities. Potentially other listener characteristics' such as linguistic skills may have a stronger relationship with individual performance with interrupted speech or other tests of cognitive processing may reveal stronger associations. Overall, the rate-dependent variation in the effects of age and hearing loss on the perception of interrupted speech indicates aspects of perceptual integration of speech fragments that appear relatively resilient to aging and hearing impairment.

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.

¹Read English-language sentences are reported to have a rate of three to four syllables per second (Jacewicz *et al.*, 2009) or about three words per second. These rate values are generally in agreement, if slightly lower, than those for the present speech materials described in Sec. II. Measurements across 60 sentences from six randomly selected sentence lists used in the present study revealed a syllable rate of 4.3 syllables per second and 3.3 words per second.

²Across primary rates, there was some variation in the proportion of the original speech preserved in dual-rate gated sentences. Due to phase relationships between the primary and secondary rates, in several conditions, the total proportion of speech remaining after dual gating with primary and secondary rates was greater than would be obtained by simply dividing by half the total proportion of speech remaining after primary gating alone. Specifically, for dual-rate gated sentences with a primary rate of 8 Hz and a 50% duty cycle, 33.36% of the original signal remained after the application of the 24-Hz secondary rate with a 50% duty cycle. For dual-rate gated sentences with a primary rate of 4 and 8 Hz and a 75% duty cycle, 41.67% of the original signal remained after the application of the 24 Hz secondary rate with a 50% duty cycle. In all other dual-rate conditions, the original speech remaining after the application of the secondary rate was one half of the amount left after the primary rate (i.e., 25% or 37.5%).

³In Kidd and Humes (2012), all interruption rates were reported per word. The interruption rates per second were estimated based on mean word durations and number of periodic interruptions per word as reported by the researchers.

- Ahissar, M., Nahum, M., Nelken, I., and Hochstein, S. (2008). "Reverse hierarchies and sensory learning," *Philos. Trans. R. Soc. London, Ser. 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.
- Başkent, D. (2010). "Phonemic restoration in sensorineural hearing loss does not depend on baseline speech perception scores," *J. Acoust. Soc. Am.* **128**, EL169–EL174.
- 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.
- 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.
- Elliott, L. L. (1995). "Verbal auditory closure and the speech perception in noise (SPIN) test," *J. Speech Lang. Hear. Res.* **38**, 1363–1376.
- 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.
- 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., Berthommier, F., and Lorenzi, C. (2006). "Masking release for consonant features in temporally fluctuating background noise," *Hear. Res.* **211**, 74–84.
- George, E. L. J., Festen, J. M., and Houtgast, T. (2006). "Factors affecting masking release for speech in modulated noise for normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **120**, 2295–2311.
- George, E. L., 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.
- Ghitza, O. (2011). "Linking speech perception and neurophysiology: Speech decoding guided by cascaded oscillators locked to the input rhythm," *Front. Psychol.* **2**, 1–13.
- 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.
- Giraud, A. L., and Poeppel, D. (2012). "Cortical oscillations and speech processing: Emerging computational principles and operations," *Nat. Neurosci.* **15**, 511–517.
- 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.
- Grose, J. H., Mamo, S. K., and Hall, J. W. (2009). "Age effects in temporal envelope processing: Speech unmasking and auditory steady state responses," *Ear Hear.* **30**, 568–575.
- 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.
- Howard-Jones, P. A., and Rosen, S. (1993). "The perception of speech in fluctuating noise," *Acustica* **78**, 258–272.
- Huggins, A. W. (1975). "Temporally segmented speech," *Percept. Psychophys.* **18**, 149–157.
- Humes, L. E., Burk, M. H., Coughlin, M. P., Busey, T. A., and Strauser, L. E. (2007). "Auditory speech recognition and visual test recognition in younger and older adults: Similarities and differences between modalities and the effects of presentation rate," *J. Speech Hear. Res.* **50**, 283–303.
- 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*, edited by S. Gordon-Salant, R. D. Frisina, A. Popper, and D. Fay (Springer Handbook of Auditory Research, Springer, Berlin), Chap. 8, pp. 211–258.

- Jacewicz, E., Fox, R. A., O'Neil, C., and Salmons, J. (2009). "Articulation rate across dialect, age, and gender," *Lang. Var. Change* **21**, 233–256.
- 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. (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. 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.
- Killion, M. C., Niquette, P. A., Gudmundsen, G. I., Revit, L. J., and Banerjee, S. (2004). "Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **116**, 2395–2405.
- Leger, A., Moore, B. C. J., Gnansia, D., and Lorenzi, C. (2012). "Effects of spectral smearing on temporal and spectral masking release in low- and mid-frequency region," *J. Acoust. Soc. Am.* **131**, 4114–4123.
- Lunner T. (2003). "Cognitive function in relation to hearing aid use," *Int. J. Audiol.* **42**, S49–S58.
- Mattys, S. L. (1997). "The use of time during lexical processing and segmentation: A review," *Psychonomic Bull. Rev.* **4**, 310–329.
- Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech," *J. Acoust. Soc. Am.* **22**, 167–173.
- Moore, B. (2003). "Temporal integration and context effects in hearing," *J. Phon.* **31**, 563–574.
- Nahum, M., Nelken, I., and Ahissar, M. (2008). "Low-level information and high-level perception: The case of speech in noise," *PLoS Biol.* **6**, 978–991.
- 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.
- Nelson, P. B., Jin, S.-H., Carney, A. E., and Nelson, D. A. (2003). "Understanding speech in modulated interference: Cochlear implant users and normal hearing listeners," *J. Acoust. Soc. Am.* **113**, 961–968.
- 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.
- 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.
- Powers, G. L., and Speaks, C. (1973). "Intelligibility of temporally interrupted speech," *J. Acoust. Soc. Am.* **54**, 661–667.
- 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.
- Rönnberg, J. (1990). "Cognitive and communicative function: The effects of chronological age and 'handicap age,'" *Eur. J. Cog. Psychol.* **2**, 253–273.
- Rosen, S. (1992). "Temporal information in speech: Acoustic, auditory and linguistic aspects," *Philos. Trans. R. Soc. London, Ser. B* **336**, 367–373.
- 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.
- 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*, edited by S. Gordon-Salant, R. D., Frisina, A., Popper, and D. Fay (Springer Handbook of Auditory Research, Springer, Berlin), Chap. 7, pp. 167–210.
- Schwartz, K., Chatterjee, M., and Gordon-Salant, S. (2008). "Recognition of spectrally degraded phonemes by younger, middle-aged, and older normal-hearing listeners," *J. Acoust. Soc. Am.* **124**, 3972–3988.
- Shafiro, V., Sheft, S., Gygi, B., and Ho, K. T. N. (2012). "The influence of environmental sound training on the perception of spectrally degraded speech and environmental sounds," *Trends Ampl.* **16**, 83–101.
- Shafiro, V., Sheft, S., and Risley, R. (2011). "Perception of interrupted speech: Effects of dual-rate gating on the intelligibility of words and sentences," *J. Acoust. Soc. Am.* **130**, 2076–2087.
- Sheft, S., Risley, R., and Shafiro, V. (2012a). "Clinical measures of static and dynamic spectral- pattern discrimination in relationship to speech perception," in *Speech Perception and Auditory Disorders*, edited by T. Dau, M. L. Jepsen, T. Poulsen, and J. C. Dalsgaard (Danavox Jubille Foundation, Ballerup, Denmark), pp. 481–488.
- Sheft, S., Shafiro, V., Lorenzi, C., McMullen, R., and Farrell, C. (2012b). "Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech perception," *Ear. Hear.* **33**, 709–720.
- Stilp, C. E. (2014). "Information-bearing acoustic change outperforms duration in predicting intelligibility of full-spectrum and noise-vocoded sentences," *J. Acoust. Soc. Am.* **135**, 1518–1529.
- Stilp, C. E., Goupell, M. J., and Kluender, K. R. (2013). "Speech perception in simulated electric hearing exploits information-bearing acoustic change," *J. Acoust. Soc. Am.* **133**, EL136–EL141.
- Studebaker, G. A. (1985). "A 'rationalized' arcsine transform," *J. Speech. Hear. Res.* **28**, 455–462.
- van Rooij, J. C. G. M., and Plomp, R. (1992). "Auditive and cognitive factors in speech perception by elderly listeners. III. Additional data and final discussion," *J. Acoust. Soc. Am.* **91**, 1028–1033.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale (WAIS-3®)*, 3rd ed. (Harcourt Assessment, San Antonio, TX).
- Wingfield, A., and Tun, P. A. (2007). "Cognitive supports and cognitive constraints on comprehension of spoken language," *J. Am. Acad. Audiol.* **18**, 548–558.