

Age effects on perceptual restoration of degraded interrupted sentences

Brittany N. Jaekel,^{a)} Rochelle S. Newman, and Matthew J. Goupell

Department of Hearing and Speech Sciences, University of Maryland, College Park, Maryland 20742, USA

(Received 13 June 2017; revised 18 November 2017; accepted 22 November 2017; published online 8 January 2018)

Adult cochlear-implant (CI) users show small or non-existent perceptual restoration effects when listening to interrupted speech. Perceptual restoration is believed to be a top-down mechanism that enhances speech perception in adverse listening conditions, and appears to be particularly utilized by older normal-hearing participants. Whether older normal-hearing participants can derive any restoration benefits from degraded speech (as would be presented through a CI speech processor) is the focus of this study. Two groups of normal-hearing participants (younger: age ≤ 30 yrs; older: age ≥ 60 yrs) were tested for perceptual restoration effects in the context of interrupted sentences. Speech signal degradations were controlled by manipulating parameters of a noise vocoder and were used to analyze effects of spectral resolution and noise burst spectral content on perceptual restoration. Older normal-hearing participants generally showed larger and more consistent perceptual restoration benefits for vocoded speech than did younger normal-hearing participants, even in the lowest spectral resolution conditions. Reduced restoration in CI users thus may be caused by factors like noise reduction strategies or small dynamic ranges rather than an interaction of aging effects and low spectral resolution. © 2018 Acoustical Society of America.

<https://doi.org/10.1121/1.5016968>

[DB]

Pages: 84–97

I. INTRODUCTION

Real-world listening environments may be noisy and cause important speech information to be interrupted by extraneous sounds. Perceptual restoration (PR) is one mechanism that allows listeners to increase their speech understanding in such environments (Warren, 1970; Samuel, 1981). For example, young normal-hearing (YNH) listeners typically show poor speech recognition when sentences or words are interrupted with silent gaps (SGs), but comparatively better speech recognition when SGs are “filled-in” with short noise bursts (NBs) (Warren, 1970; Samuel, 1981; Powers and Wilcox, 1977; Miller and Licklider, 1950). The addition of NBs to the SG-interrupted sentence seems to create an auditory illusion in which the speech is perceived as continuing “behind” the NBs. It is possible that this perception of an intact sentence is what leads to better speech understanding. The PR paradigm is useful for understanding to what extent listeners can repair interrupted speech and improve understanding.

Older people experience deficits in peripheral and central auditory processing as well as decreases in cognitive abilities as the result of aging (Working Group on Speech Understanding and Aging, 1988; Fitzgibbons and Gordon-Salant, 1996; Pichora-Fuller and Souza, 2003; Strouse *et al.*, 1998). Despite slower and less efficient age-related central auditory processing, older normal-hearing (ONH) listeners do appear to have access to the PR effect. For example, Saija *et al.* (2014) found that ONH listeners had a *larger* PR benefit compared to YNH listeners. Older adults may rely more

heavily on linguistic knowledge than younger adults to restore interrupted speech (Saija *et al.*, 2014), and may also be more likely to utilize sentence context in general (e.g., Pichora-Fuller *et al.*, 1995; Schneider *et al.*, 2010; Sheldon *et al.*, 2008; Pichora-Fuller, 2008). Whether older adults can still access such context benefits and utilize top-down PR mechanisms with interrupted *degraded* speech was the focus of the present study. For the purposes of this study, “signal degradation” refers to the changes in speech signals caused by passage through the speech processor of a cochlear implant (CI), aspects of which are simulated here through a process called vocoding. Furthermore, we investigated whether stronger cognitive abilities like working memory or better linguistic abilities like lexical access are linked with the ability to restore speech in such difficult, degraded listening conditions.

A. PR of degraded speech

CI users, who experience degraded speech signals through their processors, have typically shown only small or atypical PR effects (Bhargava *et al.*, 2014). It appears that CI users are less able to utilize NBs to repair an interrupted speech signal. Bhargava *et al.* (2014) concluded that CI users may need access to longer portions of the speech signal in order to utilize top-down PR mechanisms, and/or that degraded speech and interrupting noise may be difficult to distinguish from one another. In general, CI users report difficulty with speech-in-noise perception, showing sometimes drastically reduced speech understanding in noisy listening environments (Fetterman and Domico, 2002; Sladen and Zappler, 2015). Therefore, determining ways in which we

^{a)}Electronic mail: jaekel@umd.edu

can improve or provide access to speech-related PR mechanisms in this population may help ameliorate this problem. Using a vocoder, which can simulate aspects of CI processing, researchers can investigate which aspects of the signal itself may be affecting the PR mechanism. By presenting vocoded speech signals to normal-hearing listeners, one can also better control for listener factors that cannot be controlled for in CI users, like health of the auditory nerve (see Litovsky *et al.*, 2012 for a review).

An important listener factor that could affect PR ability is chronological age. In the study by Bhargava *et al.* (2014), older CI users tended to experience reduced PR, particularly for speech that was interrupted at a 50% duty cycle. In that study, seven of eight CI users over the age of 50 yrs had PR effects of approximately ≤ 0 RAUs (rationalized arcsine units; see Studebaker, 1985). While younger CI users in the study were more likely to have longer durations of CI use and greater access to residual hearing, it is possible that the interaction of degraded, interrupted speech and older age resulted in reduced PR ability. Further characterizing the ways in which the PR mechanism is available to older listeners presented degraded speech will be helpful for better understanding how some CI users might be processing the noisy, interrupted speech signals typical of real-world listening scenarios.

B. PR in ONH listeners

In contrast to the possible age-related reductions in PR observed in CI users (Bhargava *et al.*, 2014), PR appears to be enhanced in ONH listeners when speech is non-degraded (i.e., non-vocoded; Saija *et al.*, 2014). In the Saija *et al.* (2014) study, sequences of SGs or NBs interrupted speech at various rates of interruption. With SG interrupted speech, ONH listeners (average age of 66 yrs) had significantly worse speech understanding compared to YNH listeners (average age of 22 yrs) at certain interruption rates, potentially because advanced age reduced the ability to integrate speech information across gaps. In contrast, with NB-interrupted speech, performance for both listener groups improved. This indicated that both groups were able to utilize noise to improve speech perception. Specifically, ONH listeners showed significantly larger PR benefits at a 2.5-Hz interruption rate (20 to 30 percentage point improvement) compared to YNH listeners (approximately 10 percentage point improvement), which resulted in comparable absolute speech understanding scores between YNH and ONH listeners. ONH listeners therefore relied to a greater extent on the presence of NBs in order to achieve equivalent speech understanding than YNH listeners. In general, ONH listeners may rely more heavily on top-down mechanisms to achieve good speech perception. Greater reliance on sentential context cues and/or other top-down strategies in ONH listeners compared to YNH listeners has been reported for other speech tasks (Pichora-Fuller, 2008; Sommers and Danielson, 1999; Sheldon *et al.*, 2008). Pichora-Fuller *et al.* (2008) referred to this phenomenon as a “rebalance” of ONH listeners’ bottom-up versus top-down cue use: without high quality, well-encoded speech signals, ONH listeners’ use of

bottom-up auditory cues decreases while their use of top-down mechanisms (e.g., memory, vocabulary, and/or sentential context) increases.

C. Using NB interruptions to restore degraded speech

As described above, only one study thus far has analyzed PR abilities with interrupting NBs in CI users directly (Bhargava *et al.*, 2014). Other studies have examined CI users’ perception of speech interrupted with a non-target speaker, reporting very poor performance (average of 5% correct identification; Gnansia *et al.*, 2010), and with gated noise maskers, showing that CI users’ speech understanding reduces substantially compared to YNH listeners in the presence of fluctuating background noise (Nelson and Jin, 2004). While the study by Bhargava *et al.* (2014) found that, on average, CI users showed no PR at any signal-to-noise ratio (SNR) at the 50% duty cycle, individual results were highly variable, with PR effects spanning approximately -10 to $+20$ RAUs. While some of this variability may be explained by the age of the listener (seven of the eight older CI users showed negligible or negative PR effects; see Fig. 2 in Bhargava *et al.*, 2014), this variability provides a basis for further study of other individual characteristics of listeners that may be driving PR benefits in degraded listening conditions.

Adding NBs to degraded, vocoded speech produces mixed results in terms of PR benefits in YNH listeners. Başkent (2012) presented sentences interrupted with SGs or NBs at a 1.5-Hz rate to YNH listeners, and varied vocoder parameters to determine the amount of spectral resolution needed for PR to occur. Sentences were noise vocoded with 4, 8, 16, or 32 channels. Interruptions (SGs or NBs) were added to the sentences *prior* to vocoding to simulate how such interruptions would be processed by a CI. PR for vocoded speech was found only at 32 channels, which indicated a spectral degradation “limit” for restorability, after which the vocoded speech and NBs may have been too perceptually similar to be discriminated from one another. In such listening conditions, perhaps top-down mechanisms were unable to interact with speech portions of bottom-up acoustic information, and PR could not occur (Başkent, 2012).

Creating starker spectral differences between speech and interrupting noise therefore may increase PR in degraded speech conditions. Clarke *et al.* (2016) found significant PR benefits in YNH listeners with 16-channel noise-vocoded sentences interrupted with non-vocoded NBs. In that experiment, speech and noise were likely less perceptually similar because interrupting NBs were added to sentences *after* sentences had been vocoded. PR may therefore be possible at lower spectral resolutions than 32-channels if speech and noise are more perceptually distinct. However, it should be noted that the noise vocoder used in Clarke *et al.* (2016) was non-typical: it involved a software called TANDEM-STRAIGHT which, unlike more typical vocoders (e.g., Shannon *et al.*, 1995), did not simulate channel interactions and eliminated temporal fundamental frequency cues in the temporal envelope.

No research has specifically analyzed the effects of NB-interrupted vocoded speech and the PR effect in ONH

listeners. ONH listeners generally show slower temporal processing for complex stimuli like speech (Goupell *et al.*, 2017; Fitzgibbons and Gordon-Salant, 1996) and less precise encoding of the temporal properties of speech subcortically (Anderson *et al.*, 2012). When speech is degraded by a vocoder to simulate aspects of CI processing, spectral resolution is reduced and temporal envelopes of the signal must be relied upon for accurate speech perception (Shannon *et al.*, 1995). It is unknown if ONH listeners, who are expected to have reduced temporal processing abilities, can utilize temporal envelopes to perceptually restore NB-interrupted vocoded speech. Negligible PR effects among ONH listeners presented vocoded speech could implicate aging as a factor leading to reduced PR among some CI users.

D. Individual factors driving PR

1. Linguistic skills: Verbal fluency and lexical access

Linguistic skills like vocabulary knowledge appear to be important contributors to PR (Bashford *et al.*, 1992; Benard *et al.*, 2014). Another linguistic skill that could be correlated with PR is lexical access, or one's ability to retrieve words from his or her lexicon (Levelt, 2001). Listeners with strong lexical access abilities may be better at generating potential word candidates, even with degraded interrupted speech signals. The semantic fluency test can provide information on strength of lexical access and strength of search strategies used during word retrieval (Zarino *et al.*, 2014; Harrison *et al.*, 2000). The phonological fluency subtest in the Montreal Cognitive Assessment (MoCA) may also provide information about participants' abilities to generate word candidates based on access to partial phonological information, like a single speech sound (Tombaugh *et al.*, 1999). Slower lexical access, and thus a lower semantic and/or phonological fluency test score, is predicted to be associated with worse performance on the PR task. Noise bursts, in conjunction with neighboring speech sounds and sentence context, may stimulate fewer word candidates in people with poor lexical access ability compared to people with good lexical access ability. That is, for listeners with poor lexical access ability, the addition of NBs to a SG-interrupted speech signal might not result in improved speech perception, and thus no PR effect will be observed.

2. Cognitive function: Working memory

Working memory deficits can be indicative of problems with storing and processing incoming information (Park *et al.*, 2002; Gordon-Salant and Cole, 2016). Listeners with less effective working memory skills may be less able to quickly process noisy or interrupted speech, and to hold this speech in memory long enough or accurately enough for top-down mechanisms to interact with the signal. In general, older adults are expected to show reduced working memory skills, potentially due to less efficient processing and decreased inhibition of irrelevant or distracting stimuli (Drag and Bieliauskas, 2010; Tulskey *et al.*, 2014). While working memory was not shown to mediate PR of unprocessed speech in YNH adults (Benard *et al.*, 2014), it is unknown if

working memory skills impact PR in older adults. Stronger working memory skills in older adults have previously been associated with better speech-in-noise perception and greater use of context cues (Gordon-Salant and Cole, 2016), so perhaps stronger working memory skills will be important for older adults completing the PR task. In summary, how linguistic skills like lexical access and cognitive skills like working memory affect PR of unprocessed and degraded speech will be analyzed in this study.

E. Summary and hypotheses

The addition of noise to an interrupted unprocessed sentence may promote PR by potentially improving one's ability to group speech information in the signal, allowing top-down mechanisms to interact with bottom-up acoustic information, and/or prompt lexical access by stimulating possible candidates for completing the sentence (Bhargava *et al.*, 2014; Clarke *et al.*, 2016). The present study aims to evaluate how PR works in degraded speech conditions for YNH and ONH participants, and whether PR benefits are mediated by linguistic and/or cognitive skills. Findings here will help elucidate how PR—an important and useful skill for speech understanding—may be functioning in CI users, and whether chronological age can impact one's ability to restore degraded speech. Speech was either unprocessed or degraded using noise-vocoders with 16 or 32 channels of spectral information to simulate aspects of CI processing. While CI users have been shown to effectively have access to only approximately eight channels of spectral information at any point in time (Friesen *et al.*, 2001), we chose to utilize greater numbers of spectral channels to investigate how aging affects PR at previously tested "boundaries" of vocoded speech restoration. These boundaries appear to be, specifically, at 16-channels with NBs added *after* signal vocoding (see Clarke *et al.*, 2016) and at 32-channels with NBs added *prior to* signal vocoding (see Başkent, 2012). In the present study, SGs and NBs were used to interrupt sentences, and NBs were either vocoded (in that NBs were added to the sentence *before* the sentence was vocoded) or non-vocoded (in that NBs were added to the sentence *after* the sentence was vocoded). Previous studies investigating vocoded speech and PR (e.g., Başkent, 2012; Bhargava *et al.*, 2014) typically used *vocoded* NBs, as this method approximates how NB-interrupted speech would be processed by a CI speech processor. However, due to the degraded nature of the signals, it might have been difficult for listeners to determine which parts of the signal were noise and which parts were speech—leading to reduced PR (Bhargava *et al.*, 2014). Therefore, a second NB condition was employed in the present study: adding *non-vocoded* NBs to vocoded speech was expected to create perceptual dissimilarities between speech and NB interruptions. If perceptual dissimilarity between NB interruptions and speech is necessary to induce PR, greater PR in the non-vocoded NB conditions compared to vocoded NB conditions should be observed. Although this is not a signal that CI users would normally experience, it allows us to identify the underlying causes of potential difficulties with PR in this population.

Additional hypotheses are as follows:

- ONH participants will show larger PR benefits compared to YNH participants in the unprocessed speech conditions (confirming results reported in [Saija et al., 2014](#)) and in the degraded speech conditions with non-vocoded NBs. The non-vocoded NB condition will create NBs that are spectrally dissimilar from speech segments, better allowing ONH participants to glean speech information from the sentence and allow top-down knowledge to interact with bottom-up information.
- With lower spectral resolutions (16-channels as opposed to 32-channels), PR in ONH participants—who often have poorer temporal processing and encoding skills compared to YNH participants—will decrease, as temporal cues become more important.
- Working memory and linguistic scores and their relationship to PR may help differentiate whether cognitive factors or linguistic factors are more crucial to the PR effect. Based on previous research, it is expected that working memory skills will be an important factor for ONH participants, while linguistic skills will be an important factor regardless of participant age.

II. EXPERIMENT

A. Participants

There were two participant groups, categorized by chronological age. The first group ($n=15$) was composed of YNH adults with a mean age of 23.7 yrs [standard deviation (SD)=3.8, range 20 to 30 yrs]. The second group ($n=17$) was composed of ONH adults with a mean age of 66.8 yrs (SD=4.9, range 60 to 75 yrs). Participants had hearing thresholds at or below 25 dB hearing level (HL) at octave frequencies between 250 and 4000 Hz in both ears,¹ with no threshold discrepancy between ears greater than 15 dB. The averaged

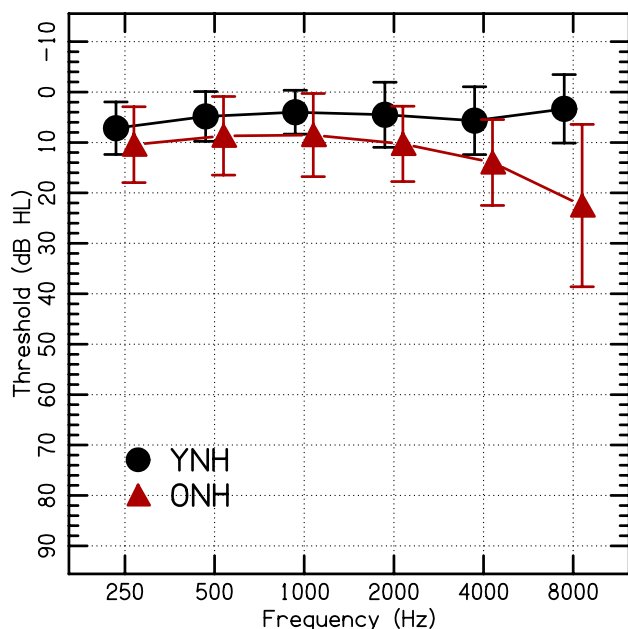


FIG. 1. (Color online) Averaged audiometric thresholds for YNH (circles) and ONH (triangles) participants.

audiometric thresholds for the two participant groups are presented in Fig. 1. Participants were also required to pass the MoCA ([Nasreddine et al., 2005](#)) with a score of at least 22 points (out of 30 possible points), which was used as a screener for the presence of mild cognitive impairment. A sub-test of phonological fluency is included in the MoCA. Finally, all participants were self-reported native speakers of American English.

B. Stimuli

1. Sentences

Two corpora of sentences were used in the present study. The corpora differ in length and complexity of sentences, and may generate different patterns of PR effects. IEEE sentences were used for training and half of the test stimuli ([Rothausser et al., 1969](#)). IEEE sentences contain 5 to 12 words, 5 of which are keywords. For each participant, 120 sentences were drawn in random order without replacement from a corpus of 720 IEEE sentences recorded by a young adult male native speaker of American English. IEEE sentence durations were on average 2.7 s (SD=0.4) and the speaker produced on average 2.9 words per second (SD=0.4). Bamford-Kowal-Bench (BKB) sentences were used for the other half of test stimuli ([Bench et al., 1979](#)). BKB sentences contain 3 to 7 words, 2 to 4 of which are keywords, and have simple syntax and vocabulary. For each participant, 120 sentences were drawn in random order without replacement from a corpus of 120 BKB sentences recorded by a young adult female native speaker of American English. BKB sentence durations were on average 1.7 s (SD=0.2) and the speaker produced on average 3.1 words per second (SD=0.6). Previous work has shown that BKB sentences have greater sentence predictability than IEEE sentences, based on an analysis of text reception thresholds in adult participants ([Schoof and Rosen, 2015](#)).

2. Vocoding

To simulate aspects of CI processing, sentences were noise vocoded. The noise vocoder created spectral degradations and removed the temporal fine structure from the speech signals ([Friesen et al., 2001](#); [Shannon et al., 1995](#)). Noise vocoding also affected the integrity of the temporal envelope by adding small random amplitude fluctuations to the envelope ([Whitmal et al., 2007](#)). Thus, the vocoder used in the present study degraded spectral properties and some temporal properties of the signal. Sentences were bandpass filtered into 16 or 32 channels, which altered the spectral resolution of the sentences. The filters were third-order Butterworth filters with forward-backward filtering applied, creating filter slopes that were -36 dB per octave. Forward-backward filtering was used to minimize distortions in the temporal envelope of the signal. The channels were contiguous and logarithmically spaced, with frequency boundaries from 200 to 4000 Hz. Temporal envelopes were extracted from these channels using the Hilbert transform, and the low pass filter envelope cutoff was 160 Hz. White noise was modulated by the extracted temporal envelopes from each

channel, and then all channels were added together and presented with the same root-mean-square energy as the original unprocessed sentence. Unprocessed sentences were the originally recorded sentences containing their original spectral and temporal properties. Both unprocessed and vocoded sentences were used in the present study.

3. Interruptions

Sentences were either interrupted at a 2.5-Hz rate or left intact (uninterrupted). The 2.5-Hz rate generated the largest age-related PR effects for VU sentences (Versfeld *et al.*, 2000) spoken at a normal rate in a previous study by Saija *et al.* (2014). Interruptions were created with a periodic nominally square wave with 1-ms raised cosine on/off ramps, which always began with a full-duration on phase. Interruptions occurred at a 50% duty cycle. Therefore, participants heard 200 ms of speech information followed by 200 ms of interruption, a cycle that repeated for the duration of the sentence. These interruptions might affect perception of various speech segments like vowels and consonant-vowel/vowel-consonant clusters depending on stress, speech rate, and location in the phrase (Crystal and House, 1990).

Two types of interruptions were applied to unprocessed speech: SGs or NBs. The NBs had the same spectrum as the long-term Fourier-transformation of the averaged spectra of the combined sentences in the respective corpus. The inverse of the periodic square wave applied to the target was applied to the noise sample, with 1-ms raised cosine on/off ramps. By adding together the interrupted speech and these interrupted noise signals, the interrupting gaps were filled with NBs.

Three types of interruptions were applied to vocoded speech: SGs, NBs added to sentences *before* vocoding was applied to the sentence (NBBV), and NBs added to sentences *after* vocoding was applied to the sentence (NBAV). All target speech was presented at 65 dB sound pressure level (SPL), and NB interruptions were presented at 70 dB SPL, resulting in a -5 -dB SNR. Presenting speech at negative rather than positive SNRs typically results in greater speech intelligibility in NB interrupted speech conditions (Powers and Wilcox, 1977). Furthermore, previous studies in PR research have utilized negative SNRs (e.g., Saija *et al.*, 2014).

C. Procedure

Measurements were conducted in the following order for all participants: semantic fluency test, working memory test, MoCA (which included the phonological fluency sub-test), vocoded speech training, and the PR task. Breaks were allowed as needed. The total experiment duration was between 1 and 2 h. Participants were seated in a sound-treated booth (Industrial Acoustics, Inc., Bronx, NY) for the duration of the experiment, and the experimenter sat in the booth beside the participant.

1. Semantic fluency test

The participant was asked to generate names of animals for 1 min, without repetition and as quickly as possible. The experimenter used a timer and tallied the number of responses from the participant. The total number of reported animals became that participant's semantic fluency score.

2. Phonological fluency test

Phonological fluency was measured by a sub-test in the MoCA. Participants were asked to generate words that began with the letter "F" for 1 min. Participants were warned that proper nouns would not be accepted, and that the same word appended with various suffixes would only be counted once. The total number of reported words became that participant's phonological fluency score.

3. List sorting working memory test

The List Sorting Working Memory Test (Age 7 + v2.1) was presented on an iPad 2 (Apple, Inc., Cupertino, CA). This test is available through the NIH Toolbox iPad Application as part of the Cognitive Battery, and is appropriate for participants aged 7 years and older (Glinberg and Associates, Inc., 2016). The test duration is approximately 7 min. The iPad screen was pointed toward the participant with full screen brightness and volume at 75%. The experimenter controlled the application and scored task performance using a wireless keyboard and answer sheet hidden from the participants' view. Participants were asked to view a sequence of pictures depicting items and, at the end of the sequence, report the names of the items in the order of smallest- to largest-sized item. Picture size correlated with the real-world size of the item. For example, the picture of the elephant (large item) filled most of the iPad screen. Each picture in the sequence was presented simultaneously with a recording of an adult American English-speaking woman pronouncing the name of the item. The second test portion added the additional rule of categorizing items based on whether they were food or animals, and reporting items smallest- to largest-sized within-category. Answers were marked correct only if all items in the sequence were reported in their correct order. A correct answer prompted the application to increase the number of items presented in the following sequence. The test portion ended either when participants answered two subsequent trials incorrectly, or when participants successfully reported all seven items for a seven-item sequence, the maximum number of items in the test. The overall "uncorrected" standard score, which is generated by the test application, was noted for each participant. This score standardizes each participant's raw score without reference to the participant's chronological age.

4. Vocoded speech training

Stimuli were played diotically through circumaural headphones (Sennheiser, HD 650; Old Lyme, CT). The experiment was controlled using MATLAB (The Mathworks, Inc., Natick, MA) and presented on a computer with a touchscreen monitor. Participants were asked to sit turned away from the monitor so that answers (presented visually on the screen to the

experimenter) could not be seen. The experimenter controlled the experiment using the touchscreen monitor.

Participants were familiarized with noise-vocoded speech through a training session prior to completing the PR task. Participants listened to 50 IEEE sentences that were noise vocoded with 8 channels. Note that BKB sentences were not used for training as there were a limited number of sentences in that corpus and thus they were retained for the testing conditions. Furthermore, an 8-channel vocoder was used during training because previous research has shown that training on difficult speech-related tasks with spectrally degraded signals generalized to performance on easier speech-related tasks (Loebach and Pisoni, 2008). Training sentences were not interrupted with SGs or NBs, and did not later re-appear during the PR task. Participants were asked to listen to the presented speech and then report what they heard to the experimenter as accurately as possible. On each trial, the experimenter pressed a “play” button, the sentence was presented to the participant, and the participant verbally repeated what they heard. The experimenter recorded correct keywords. The same sentence was then presented twice more: first as an unprocessed version of the sentence, followed by the same vocoded version of the sentence. Participants were asked to listen to these sentences and were not required to report what they heard. After these two “feedback” sentences played, the next trial could begin.

5. PR task

Before testing began, participants were informed that the test sentences might be intact or contain gaps or noise, and might be vocoded or unprocessed. Participants were also informed that no feedback would be provided, that sentences could not be repeated, and that they were encouraged to guess. On each trial, one test sentence was presented, and after presentation, participants reported verbally what they heard. The experimenter recorded the participants’ answers by selecting correct keywords. “Lax” scoring was used, in that if the reported word had an incorrect tense or suffix, this was recorded as a correct response (e.g., if the correct keyword was “helps” and the participant reported “helped,” this was marked as correct). Homophones were also recorded as correct. For each trial, participants’ scores were the number of correctly reported target words divided by the total number of target words in the sentence. To ensure accurate grading by the first experimenter, a second experimenter blind to experimental condition regraded sentences using voice recordings of the participants’ answers.

Test conditions were blocked by sentence processing type (three levels: unprocessed, 16-, and 32-channel vocoded) and sentence corpus (two levels: IEEE and BKB), resulting in six different blocks. The order of blocks was randomized for each participant. Within each block, interruption conditions were randomized. All sentences from both corpora were created in such a way that they could appear with any type of sentence processing and any interruption, and could therefore appear in any block featuring their corpus, without replacement.

The following information applies to both speech corpora used in the study. In the unprocessed speech block, there were three interruption conditions, with eight sentences in each condition, for a total of 24 sentences (per corpus). The three conditions were (1) intact speech, (2) speech interrupted with SGs, and (3) speech interrupted with NBs. In the 32-channel noise-vocoded speech block, there were four conditions, with 12 sentences in each condition, for a total of 48 sentences (per corpus). The four conditions were: (1) intact vocoded speech, (2) vocoded speech interrupted with SGs, (3) vocoded speech interrupted with NBBV, and (4) vocoded speech interrupted with NBAV. In the 16-channel noise-vocoded speech block, there were the same four conditions as in the 32-channel block, with 12 sentences in each condition, for a total of 48 sentences (per corpus). In summary, 120 BKB test sentences and 120 IEEE test sentences were presented to each participant.

D. Results

1. Subject variables

The following describes results for the subject variables of semantic fluency, phonological fluency, and working memory. Group averages were compared using two-tailed independent samples *t*-tests ($\alpha = 0.05$). For the measure of semantic fluency, YNH participants on average generated the names of 25.7 animals (SD = 6.6) and ONH participants generated the names of 21.9 animals (SD = 4.7). For the measure of phonological fluency, YNH participants on average generated 17.4 words (SD = 5.1) and ONH participants generated 15.8 words (SD = 4.6). Scores on the two linguistic measures were not significantly different between groups [semantic fluency: $t(30) = 1.92$, $p = 0.064$; phonological fluency: $t(30) = 0.95$, $p = 0.35$]. An overall linguistic skill score was calculated for each participant by averaging scores on the semantic and phonological fluency tests, and was done to reduce multicollinearity issues in the linear mixed-effect analyses below. The overall linguistic score for YNH participants was 21.6 (SD = 4.3) and for ONH participants was 18.8 (SD = 4.0), a difference which was not statistically significant [$t(30) = 1.88$, $p = 0.071$]. For the measure of working memory, YNH participants achieved on average 110.2 points (SD = 5.5) and ONH participants achieved 100.1 points (SD = 8.2), a significant difference [$t(30) = 4.04$, $p < 0.001$]. In summary, YNH participants achieved higher working memory scores compared to ONH participants, but comparable semantic and phonological fluency scores.

2. Vocoded speech training and intact speech perception

Individual participants’ speech understanding scores were transformed into RAUs, as some scores were near zero in certain experimental conditions. On average, YNH participants reported 90.4% (SD = 4.2) training words correctly (in RAUs: 94.5, SD = 7.0) and ONH participants reported 90.9% (SD = 3.6) training words correctly (in RAUs: 95.1, SD = 6.0). No significant differences in scores (in RAUs or in percent correct) between groups was observed per two-

tailed independent samples t -tests [RAUs: $t(30) = 0.29$, $p = 0.77$; Percent correct: $t(30) = 0.38$, $p = 0.71$].

To determine whether groups achieved different RAU scores on the six intact speech conditions (i.e., conditions without any interruptions), the two groups' scores were compared with multiple two-tailed independent samples t -tests ($\alpha = 0.05$). Scores on these baseline conditions were high in both groups (110.0 RAUs or higher, on average), and no differences in group performance were significant [unprocessed BKBs: $t(30) = 0.94$, $p = 0.37$; 32-channel vocoded BKBs: $t(30) = 0.16$, $p = 0.88$; 16-channel vocoded BKBs: $t(30) = 0.04$, $p = 0.97$; unprocessed IEEEs: $t(30) = 0.04$, $p = 0.97$; 32-channel vocoded IEEEs: $t(30) = 0.96$, $p = 0.34$, 16-channel vocoded IEEEs: $t(30) = 0.70$, $p = 0.49$]. Thus, age did not affect perception of intact speech, even when speech was noise vocoded with 16- or 32-channels. Individual participants' speech understanding scores on intact conditions were always greater than or equal to 93.2 RAUs.

3. PR of interrupted sentences

Inter-rater reliability for judging participant responses during the PR task was high; the second rater agreed with the first rater 98.0% of the time. Because of the substantial concordance of the two raters' scorings, the original rater's data were used for all analyses. Speech understanding scores (transformed into RAUs) for each interruption type, spectral resolution, corpus, and age group are tabulated in Table I, as are PR benefits. PR benefits were calculated by subtracting scores in the SG condition from scores in the NB, NBAV, or NBBV condition. Speech understanding scores are presented for BKB sentences in Fig. 2 and for IEEEs sentences in Fig. 3. Whether speech understanding improved significantly between SG and the various NB conditions for each participant was tested with paired samples t -

tests, corrected for multiple comparisons ($\alpha = 0.005$). YNH participants showed a significant benefit with the addition of noise in only two conditions: 32-channel vocoded BKB sentences with NBAV [$t(14) = 3.82$, $p = 0.002$, $d = 1.32$] and 32-channel vocoded IEEEs sentences with NBAV [$t(14) = 3.65$, $p = 0.003$, $d = 0.94$]. In contrast, ONH participants showed a significant benefit with the addition of noise in seven conditions: 32-channel vocoded BKB sentences with NBAV [$t(16) = 4.42$, $p < 0.001$, $d = 1.08$] and NBBV [$t(16) = 3.68$, $p = 0.002$, $d = 0.89$], and every IEEEs sentence condition [unprocessed SG vs NB: $t(16) = 4.36$, $p < 0.001$, $d = 1.06$; 32-channel vocoded SG vs NBAV: $t(16) = 7.28$, $p < 0.001$, $d = 1.79$; 32-channel vocoded SG vs NBBV: $t(16) = 5.93$, $p < 0.001$, $d = 1.50$; 16-channel vocoded SG vs NBAV: $t(16) = 6.19$, $p < 0.001$, $d = 1.51$; and 16-channel vocoded SG vs NBBV: $t(16) = 5.00$, $p < 0.001$, $d = 1.21$]. ONH participants therefore typically obtained a significant benefit from the addition of noise to interrupted sentences. Average PR benefits (as well as individual data) for each group are presented for BKB sentences (Fig. 2) and IEEEs sentences (Fig. 3), for each listening condition. While ONH participants as a group tended to show larger PR benefits than YNH participants, variability across participants was high. Large variability in PR performance has been reported previously (e.g., Bhargava *et al.*, 2014; Verschuure and Brocaar, 1983). We also note that a ceiling effect could exist for YNH participants presented unprocessed BKB sentences (Fig. 2) and a floor effect for ONH participants presented vocoded IEEEs sentences (Fig. 3). For the former case, differences in performance for SG and NB sentences were not significantly different for unprocessed BKB sentences for either group, and larger PR benefits in YNH than in ONH participants would contradict previous research in this area (e.g., see Saija *et al.*, 2014). For the latter case, floor effects in

TABLE I. Average speech understanding scores (in RAUs) in each condition and average PR benefits are presented, rounded to the nearest tenth. SDs are in parentheses. PR benefits were calculated by subtracting each participant's scores in the SG-interrupted condition from their score in the associated NB-interrupted condition; the averaged values of these PR benefits are listed. Asterisks indicate significant PR benefits per paired samples t -tests (comparing performance on SG conditions to each noise condition) by group ($\alpha = 0.005$). Blank boxes indicate untested conditions.

	Noise Vocoded					
	Unprocessed		32 channels		16 channels	
	YNH	ONH	YNH	ONH	YNH	ONH
BKBs						
SG	90.7 (13.7)	72.2 (16.9)	57.5 (11.4)	38.5 (14.7)	44.5 (14.6)	25.1 (10.1)
NBAV or NB	94.7 (14.5)	83.9 (20.5)	75.4 (11.3)	56.9 (13.3)	55.8 (12.9)	36.9 (19.1)
NBBV			60.0 (11.2)	50.6 (15.1)	48.1 (14.1)	34.0 (15.7)
PR benefit (NBAV-SG) or (NB-SG)	4.0 (18.0)	11.7 (17.8)	17.9* (18.2)	18.4* (17.2)	11.3 (18.3)	11.8 (20.7)
PR benefit (NBBV-SG)			2.6 (17.0)	12.0* (13.5)	3.7 (16.6)	8.9 (15.3)
			Noise Vocoded			
	Unprocessed		32 channels		16 channels	
IEEEs	YNH	ONH	YNH	ONH	YNH	ONH
SG	57.7 (17.6)	42.4 (16.3)	22.4 (12.9)	7.3 (9.5)	15.1 (12.5)	4.7 (10.8)
NBAV or NB	66.0 (12.9)	59.8 (14.0)	36.6 (12.3)	28.6 (12.0)	23.5 (13.4)	19.8 (12.2)
NBBV			30.8 (10.8)	23.2 (12.9)	19.2 (14.0)	17.5 (10.7)
PR benefit (NBAV-SG) or (NB-SG)	8.3 (16.7)	17.4* (16.5)	14.2* (15.0)	21.3* (12.0)	8.4 (11.8)	15.2* (10.1)
PR benefit (NBBV-SG)			8.5 (12.9)	15.8* (11.0)	4.1 (14.5)	12.9* (10.6)

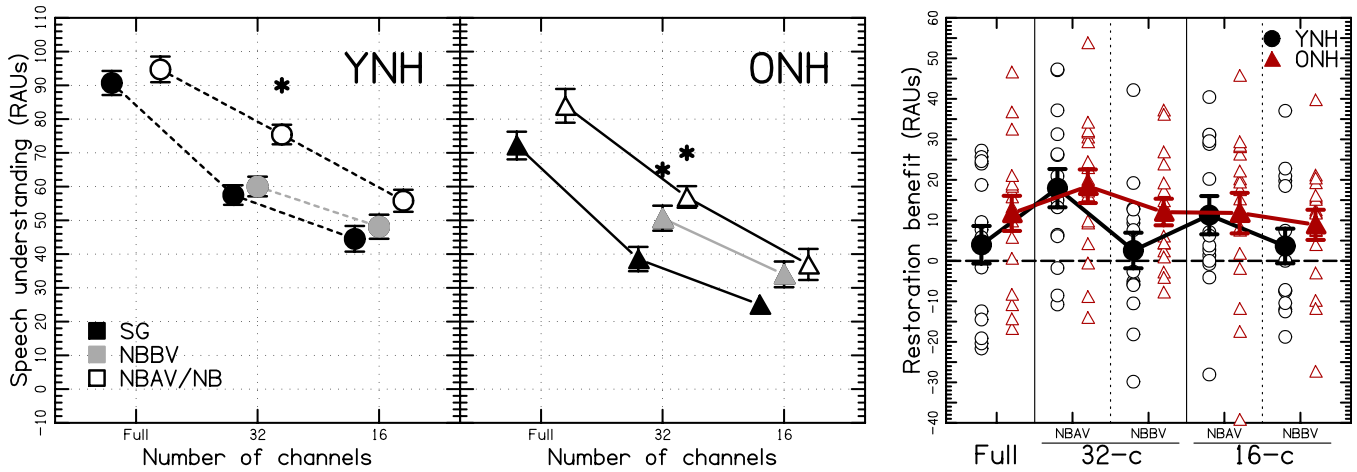


FIG. 2. (Color online) (Left): Speech understanding scores (in RAUs) for BKB sentences are presented as a function of spectral resolution. The far left panel shows results for YNH participants (circles), and the middle panel shows results for ONH participants (triangles). Black symbols represent performance in SG conditions, gray symbols represent performance in NBBV conditions, and white symbols represent performance in NBAV or NB conditions. Asterisks indicate a significant increase in performance between SG conditions and respective NB conditions, as determined by paired samples *t*-tests corrected for multiple comparisons. (Right) PR benefits (in RAUs) for BKB sentences are presented as a function of spectral resolution and NB condition, for YNH (black circles) and ONH (red triangles) participants. Solid symbols indicate group averages with standard error bars. Open symbols indicate individual data. Positive values indicate that participants obtained better speech understanding with NB-interrupted speech compared to SG-interrupted speech.

ONH participants could indicate that PR benefits are even larger than those reported here.

Linear mixed-effects (LME) models were used to understand how subject variables like overall linguistic score and working memory interacted with stimuli parameters (e.g., interruption type) for each individual participant. Four separate LME models were constructed, and each modeled individual percent-correct scores (transformed into RAUs) for interrupted sentences for each participant across specific conditions. First, from visual inspection of Figs. 2 and 3, it appeared that the factor of corpus was eliciting different patterns of responses; therefore, each corpus (BKB, IEEE) was modeled separately. Second, the unprocessed vs vocoded conditions were modeled

separately, as keeping these conditions together in a single model resulted in a rank deficiency error [vocoded conditions contained two NB types (NBAV, NBBV), while unprocessed conditions contained only one (NB)]. The four LME analyses modeled data from unprocessed BKB sentences, unprocessed IEEE sentences, vocoded BKB sentences, and vocoded IEEE sentences, respectively. The predictors in unprocessed models were age group, interruption type, linguistic score, and working memory score. The vocoded models included the additional predictor of number of channels (i.e., spectral resolution).

Discovering the following main effects and interactions through these LME analyses were of most interest: (1) a significant main effect of interruption type would indicate a PR

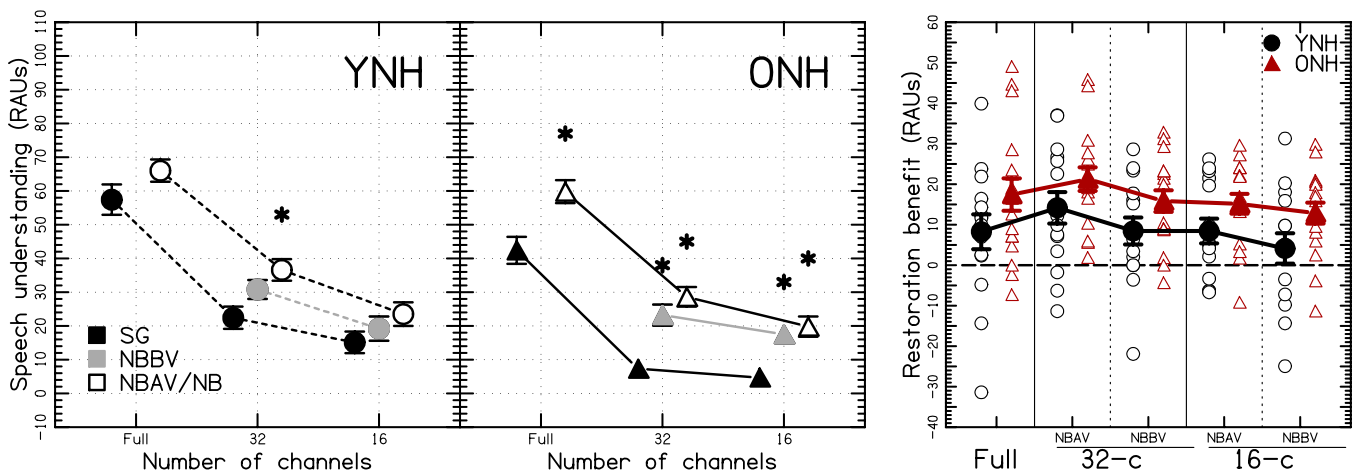


FIG. 3. (Color online) (Left): Speech understanding scores (in RAUs) for IEEE sentences are presented as a function of spectral resolution. The far left panel shows results for YNH participants (circles), and the middle panel shows results for ONH participants (triangles). Black symbols represent performance in SG conditions, gray symbols represent performance in NBBV conditions, and white symbols represent performance in NBAV or NB conditions. Asterisks indicate a significant increase in performance between SG conditions and respective NB conditions, as determined by paired samples *t*-tests corrected for multiple comparisons. (Right) PR benefits (in RAUs) for IEEE sentences are presented as a function of spectral resolution and NB condition, for YNH (black circles) and ONH (red triangles) participants. Solid symbols indicate group averages with standard error bars. Open symbols indicate individual data. Positive values indicate that participants obtained better speech understanding with NB-interrupted speech compared to SG-interrupted speech.

effect, (2) a significant interaction between age group and interruption type would indicate an “aging benefit” for PR, (3) a significant interaction between interruption type and number of channels would indicate that PR is greater at certain spectral resolutions, (4) a significant interaction(s) between linguistic score or working memory with interruption type would indicate that PR was mediated by the respective subject variable(s), and (5) a significant three-way interaction(s) among linguistic score or working memory with interruption type and age group would indicate that the aging benefit for PR was mediated by the respective subject variable(s).

a. Unprocessed BKB sentences. The first LME model analyzed participants’ RAUs for interrupted listening conditions containing unprocessed BKB sentences. We began with a maximal model structure and used a backward selection process following Barr *et al.* (2013). The full model contained fixed main effects for the subject variables of linguistic score (grand-mean centered), working memory score (grand-mean centered), and age group [two levels: 0 = YNH (reference level), 1 = ONH], and the stimulus parameter of interruption type [two levels: 0 = SG (reference level), 1 = NB]. The full model also contained the interactions of each subject variable with age group and/or interruption type, as well as the interaction of age group and interruption type. The reduced model is presented in Table II, and was obtained by repeatedly removing the highest-order fixed term that was non-significant, and re-running the model until all remaining fixed main effects and interactions were significant. Non-significant main effects and interactions were only retained in the model if they composed part of a significant interaction or higher-order interaction, an approach consistent with Hox *et al.* (2017). The random effects are also presented in Table II, and represent the maximal structure possible for model convergence. This reduced model was then compared with the full model using a χ^2 test; a non-significant χ^2 test indicated that the reduced model was sufficient for explaining the data.

Per Table II, the effect of NB interruptions was significant ($p < 0.001$), indicating a PR effect. The addition of NBs improved speech understanding scores by 12.0 RAUs, regardless of age. The effect of age was significant ($p = 0.003$), in that performance with interrupted speech in general (both SG- and NB-interrupted speech) was reduced

TABLE II. Final LME model for the unprocessed BKB sentences condition. Bolded rows indicate significant fixed terms ($\alpha = 0.05$).

Fixed effects	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	97.45	4.46	21.86	<0.001
Age (0 = YNH, 1 = ONH)	-16.03	5.07	-3.16	0.003
Interruption (0 = SG, 1 = NB)	12.03	3.30	3.65	<0.001
Random effects	Variance	SD		
Item (intercept)	390.4	19.8		
Subject (intercept)	124.6	11.2		
Residual	1163.2	34.1		

in ONH participants compared to YNH participants. The interaction of age group and NB interruptions (the aging benefit for PR) was not significant and thus removed from the final reduced model. Neither linguistic skill nor working memory significantly mediated PR for either age group.

b. Unprocessed IEEE sentences. The second LME model analyzed participants’ RAUs for interrupted listening conditions containing unprocessed IEEE sentences. The full model contained the same fixed main effects and interactions as described for the unprocessed BKB model above, and the random effects structure and final reduced model were obtained in the same manner as described above. The reduced final model for unprocessed IEEE sentences is presented in Table III.

All participants, regardless of age, showed a PR benefit, in that performance was significantly higher with NB interruptions compared to SG interruptions ($p < 0.001$). The effect of age was significant ($p = 0.036$), in that ONH participants generally achieved lower speech understanding scores compared to YNH participants. No significant aging benefit for PR was observed, as the interaction between age and interruption type was not significant and not included in the model. Neither linguistic scores nor working memory scores mediated PR benefits. Instead, better linguistic scores in both YNH and ONH participants were associated with improved interrupted speech understanding overall ($p = 0.032$).

c. Vocoded BKB sentences. The third LME model analyzed participants’ RAUs for interrupted listening conditions containing vocoded BKB sentences. The full model contained the same fixed main effects and interactions as the unprocessed sentences models above, but with the additional fixed main effect of number of channels, coded categorically [two levels: 0 = 32 channels (reference level) and 1 = 16 channels] and an interruption term containing two different NB types [three levels: 0 = SG (reference level), 1 = NBBV, and 2 = NBAV]. The full model contained these factors’ interactions with the other stimuli parameter variables and subject variables. The random effects structure and final reduced model was obtained in the same manner as described above and is presented in Table IV.

First, participants, regardless of age, were able to obtain significant PR benefits in both 32- and 16-channel vocoded listening conditions: significantly higher RAUs were

TABLE III. Final LME model for the unprocessed IEEE sentences condition. Bolded rows indicate significant fixed terms ($\alpha = 0.05$).

Fixed effects	Coefficient	SE	<i>t</i>	<i>p</i>
Intercept	55.53	4.45	12.49	<0.001
Age (0 = YNH, 1 = ONH)	-12.07	5.52	-2.19	0.036
Interruption (0 = SG, 1 = NB)	17.88	3.55	5.04	<0.001
Linguistic skill (grand-mean centered)	1.45	0.65	2.24	0.032
Random effects	Variance	SD		
Item (intercept)	838.6	29.0		
Subject (intercept)	117.3	10.8		
Residual	980.4	31.3		

TABLE IV. Final LME model for the vocoded BKB sentences condition. Bolded rows indicate significant fixed terms ($\alpha = 0.05$).

Fixed effects	Coefficient	SE	t	p
Intercept	59.13	4.79	12.34	<0.001
Age (0 = YNH, 1 = ONH)	-26.43	5.68	-4.65	<0.001
Interruption: (0 = SG, 1 = NBBV)	5.22	4.77	1.09	0.28
(0 = SG, 2 = NBAV)	20.92	5.37	3.90	<0.001
Spectral resolution (0 = 32-c, 1 = 16-c)	-18.14	3.20	-5.68	<0.001
Linguistic skill (grand-mean centered)	0.82	0.70	1.17	0.25
Working memory (grand-mean centered)	0.02	0.35	0.07	0.95
<i>Interactions</i>				
Age \times Interruption: (NBBV)	19.33	6.61	3.07	0.003
(NBAV)	7.37	7.31	1.01	0.32
Working memory \times Interruption: (NBBV)	0.91	0.38	2.39	0.019
(NBAV)	0.32	0.45	0.71	0.48
Linguistic skill \times Interruption: (NBBV)	-0.60	0.85	-0.71	0.48
(NBAV)	0.22	0.93	0.24	0.81
Spectral resolution \times Interruption: (NBBV)	-4.72	4.55	-1.04	0.30
(NBAV)	-7.15	4.52	-1.58	0.11
Spectral resolution \times Linguistic skill	-1.39	0.75	-1.85	0.06
Spectral resolution \times Linguistic skill \times Interruption: (NBBV)	1.84	1.07	1.73	0.08
(NBAV)	2.34	1.07	2.19	0.028
<i>Random effects</i>				
	Variance		SD	
Item (intercept)	731.2		27.0	
Subject (intercept)	84.0		9.2	
Subject—Interruption: (NBBV)	39.1		6.3	
(NBAV)	109.5		10.5	
Residual	1871.0		43.3	

obtained with NBAV interruptions ($p < 0.001$) than with SG interruptions and the size of this PR effect was constant across spectral resolutions. Second, an aging benefit for PR was observed for NBBV interruptions, but not for NBAV interruptions. That is, ONH participants' performance increased to a greater extent compared to YNH participants' when SGs were replaced with NBBV interruptions ($p = 0.003$), in both the 32- and 16-channel vocoded listening conditions. It should be noted that the main effect of age was significant ($p < 0.001$) in that ONH participants generally achieved lower speech understanding scores overall compared to YNH participants. Third, working memory scores significantly interacted with PR benefits obtained with NBBV interruptions ($p = 0.019$), regardless of participant age. This indicated that participants with relatively stronger, above-average working memory ability would obtain greater PR benefits obtained with NBBV interruptions compared to participants with weaker working memory. Fourth, while no subject variable significantly interacted with age and interruption type, we found that linguistic skills did interact significantly with spectral resolution and interruption type. For 32-channel vocoded interrupted speech, linguistic skills did not significantly mediate performance with SG, NBBV, or NBAV interruptions ($p > 0.05$). For 16-channel vocoded interrupted speech, in contrast, linguistic skills significantly mediated PR benefits, in that above-average linguistic skills were associated with greater PR for NBAV interruptions ($p = 0.028$).

TABLE V. Final LME model for the vocoded IEEE sentences condition. Bolded rows indicate significant fixed terms ($\alpha = 0.05$).

Fixed effects	Coefficient	SE	t	p
Intercept	12.81	3.59	3.57	<0.001
Age (0 = YNH, 1 = ONH)	-10.34	4.05	-2.56	0.015
Interruption: (0 = SG, 1 = NBBV)	13.98	2.60	5.37	<0.001
(0 = SG, 2 = NBAV)	20.23	2.65	7.64	<0.001
Spectral resolution (0 = 32-c, 1 = 16-c)	-9.38	2.92	-3.21	0.002
Linguistic skill (grand-mean centered)	0.92	0.42	2.18	0.036
Working memory (grand-mean centered)	-0.58	0.42	-1.39	0.18
<i>Interactions</i>				
Age \times Spectral resolution	7.38	3.05	2.42	0.017
Age \times Working memory	1.05	0.48	2.21	0.035
Spectral resolution \times Interruption: (NBBV)	-3.67	3.33	-1.10	0.27
(NBAV)	-6.85	3.32	-2.06	0.039
Spectral resolution \times Linguistic skill	0.90	0.36	2.50	0.014
<i>Random effects</i>				
	Variance		SD	
Item (intercept)	345.2		18.6	
Subject (intercept)	70.5		8.4	
Subject—Interruption: (NBBV)	40.0		6.3	
(NBAV)	48.3		6.9	
Subject—Spectral resolution	13.3		3.6	
Residual	878.5		29.6	

d. Vocoded IEEE sentences. The fourth LME model analyzed participants' RAUs for interrupted listening conditions containing vocoded IEEE sentences. The final reduced model was obtained in the same manner as described above, and is presented in Table V. Regardless of age, participants were able to obtain significant PR benefits. Both NB types produced a significant PR benefit for 32-channel speech ($p < 0.001$ for NBAV and NBBV). For 16-channel speech, the size of the PR benefit was significantly reduced for NBAV interruptions ($p = 0.039$) but not for NBBV interruptions ($p = 0.27$). An aging benefit for restoration was not observed, and not included in the model, though there was a significant main effect of age group ($p = 0.015$), with ONH participants performing worse overall compared to YNH participants. PR benefits were not mediated by linguistic or working memory scores. Instead, better linguistic skills were associated with improved interrupted, vocoded speech understanding in general ($p = 0.036$), and had a significantly larger positive effect on speech understanding with 16-channel vocoded speech in particular ($p = 0.014$).

III. GENERAL DISCUSSION

The present study investigated PR benefits in degraded listening conditions simulating aspects of CI processing in both YNH and ONH participants. PR benefits in degraded listening conditions were observed in both age groups but were more consistently observed among ONH participants (Figs. 2 and 3). On an individual level, PR benefits were highly variable (Figs. 2 and 3), which potentially indicated the presence of subject-specific variables mediating the effect. Four main hypotheses were tested, and are each discussed in turn below.

A. Characteristics of NB interruptions

The first hypothesis focused on the characteristics of the NB interruptions in vocoded sentences and their effect on PR. CI speech processors provide low spectral resolution and omit temporal fine structure; such degradation may reduce the perceptual dissimilarity of the speech and noise. Therefore, the PR effect may itself be reduced in such listening conditions, as top-down mechanisms would be less able to interact with bottom-up acoustic speech information. We hypothesized that creating greater perceptual dissimilarities between speech and NB interruptions would lead to stronger PR effects in vocoded conditions. In the present study, two interrupting NB conditions were used to test this hypothesis: first, NBBV, or *vocoded* NBs, represented a “weaker” perceptual dissimilarity condition, and second, NBAV, or *non-vocoded* NBs, represented a “stronger” perceptual dissimilarity condition.

LME analyses revealed that for ONH participants, *both* NB types resulted in significant PR effects in vocoded conditions (Tables IV and V; Figs. 2 and 3). For YNH participants, NBAV consistently elicited PR benefits, while NBBV only elicited PR benefits with IEEE sentences (Tables IV and V). For both groups, PR benefits with NBAV appeared to be substantially larger than those obtained with NBBV. These findings extend our understanding of the “spectral limit” for restorability, as well as whether perceptual dissimilarities between speech and noise interruptions are necessary for PR to work in a degraded sentence context. The spectral limit for restorability with NBBV in YNH participants was previously reported to be 32 noise-vocoded channels (Başkent, 2012), and for NBAV in YNH participants was reported to be 16 channels with a specialized noise-excited vocoder (see Clarke *et al.*, 2016). In the present study, the spectral limit (among tested spectral resolutions) for both NB types and both age groups was 16 noise-vocoded channels, at least with IEEE sentences (Table V). Therefore, PR is possible for both YNH and ONH participants presented degraded sentences, and while perceptual dissimilarities between speech information and noise interruptions appear to further enhance PR, they are not a prerequisite for the effect to occur. In the present study, ONH participants (mean age = 66.8 yrs) obtained average PR benefits ranging from 8.9 to 15.8 RAUs with NBBV-interrupted vocoded speech and 11.8 to 21.3 RAUs with NBAV-interrupted vocoded speech (Table I). In comparison, previous research showed that older CI users (aged 52 to 65 yrs) presented NB-interrupted speech at a 50% duty cycle typically showed only negligible or negative PR effects (see Fig. 2 of Bhargava *et al.*, 2014). Whether perceptual dissimilarities are prerequisites for PR and/or aging benefits to occur at relatively lower spectral resolutions (e.g., 8 channels), which would provide participants a similar spectral resolution to that generally available to CI users (Friesen *et al.*, 2001), is unclear. One study (Clarke *et al.*, 2016) found no significant PR with NBAV in YNH participants presented 8-channel speech processed with a specialized noise-excited vocoder.

B. Aging benefits for PR of degraded speech

The second hypothesis focused on the aging benefit associated with PR. First, it was expected that an aging benefit in unprocessed speech conditions would occur, as reported in Saija *et al.* (2014). Surprisingly, the LME model analyses of the unprocessed speech conditions did not show significant interactions of NB interruptions and age group (Tables II and III). Analyses of average group performance in the SG and NB unprocessed conditions, however, did reveal a significant aging benefit for PR for IEEE sentences (Fig. 3). Variability in performance with BKB sentences may have impacted our ability to find a significant aging benefit for that corpus in the unprocessed condition. Second, it was expected that ONH participants would show an aging benefit for PR in degraded, vocoded conditions containing NBAV interruptions. NBAV interruptions would introduce starker perceptual dissimilarities between speech and NB interruptions, which in turn would provide more clearly delineated speech information to top-down PR mechanisms. Group data presented in Figs. 2 and 3 show that ONH participants consistently showed significant increases in performance with the addition of noise—regardless of NB type—across vocoded speech conditions, and particularly for IEEE sentences.

YNH participants, in contrast, were less likely to show significant improvements in vocoded conditions. Per the LME model analyses, the interaction of age group and interruption type was significant for vocoded BKB sentences (Table IV); specifically, an aging benefit for PR was observed with NBBV interruptions. No aging benefit for PR was observed with vocoded IEEE sentences. With that corpus, both YNH and ONH participants obtained significant PR benefits with both noise types.

Slower temporal processing and less precise temporal encoding associated with aging may have influenced how ONH participants perceived and utilized NB interruptions in the present study. Older participants have less ability to perceive and encode sudden temporal changes in the speech signal (Anderson *et al.*, 2012; Fitzgibbons and Gordon-Salant, 1996; Goupell *et al.*, 2017), so the perception of the rapid changes between speech and NB interruptions may have resulted in a stronger, more fused percept of the two stimuli. A more fused percept would result in a stronger auditory illusion of the speech continuing through the noise, and therefore a stronger PR effect. Less ability to encode fast temporal changes in the signal could explain why creating starker differences between vocoded speech and noise interruptions had no effect on ONH participants, as PR was high regardless of NB type. PR of degraded speech in older participants therefore may be less dependent on the ability to distinguish speech from noise, and instead dependent on utilizing whatever context can be gleaned from the noisy speech percept (Sheldon *et al.*, 2008; Pichora-Fuller, 2008; Pichora-Fuller *et al.*, 1995).

YNH participants typically failed to utilize NBBV interruptions to restore degraded speech, but did utilize NBAV interruptions (Figs. 2 and 3). Compared to ONH participants, in everyday listening conditions, YNH participants likely do

not have to rely as heavily on PR mechanisms to perceive speech (Saija *et al.*, 2014). Therefore, YNH participants may have less experience utilizing PR and context to repair a speech signal, especially with degraded, vocoded sentences (Sheldon *et al.*, 2008; Pichora-Fuller, 2008). Overall, average speech understanding in YNH participants was always higher than that of ONH participants, in every condition (Figs. 2 and 3), especially in conditions with SG interruptions. This could indicate that for YNH participants, the utilization of NBs through PR mechanisms was less necessary for accurate speech perception. NBAV interruptions may have been particularly useful to YNH participants for prompting PR, as the more spectrally complex noise may have helped stimulate more potential lexical candidates and led to an increased chance of perceiving the correct word (Bhargava *et al.*, 2014).

C. Spectral resolution effects on PR

The third hypothesis focused on whether lower spectral resolutions would significantly reduce PR benefits in ONH participants. Vocoding reduces the amount of spectral information in the speech signal, while mostly preserving temporal envelope information (Shannon *et al.*, 1995). PR with vocoded speech might therefore depend on one's ability to restore temporal envelopes, the processing of which is less efficient in ONH participants (Goupell *et al.*, 2017). LME model analyses revealed that PR benefits in ONH participants were not affected by spectral resolution, as the interaction of age group, number of channels, and interruption type was not significant and was not included in the models (Tables IV and V). It is unclear whether an effect of spectral resolution on PR in older participants would emerge with a vocoder containing fewer channels; at 16 noise-vocoded channels, however, processing of the temporal envelope and/or the PR of temporal envelopes is still possible for ONH participants.

D. Effects of linguistic and cognitive skills on PR

The fourth hypothesis focused on how linguistic and cognitive skills might mediate PR effects. In the present study, it was posited that in degraded/vocoded listening conditions, linguistic factors and the cognitive factor of working memory would mediate PR, and would do so differently for the two age groups. Linguistic knowledge and vocabulary are "crystallized" forms of knowledge that are generally not impacted negatively by aging; in fact, vocabulary size tends to grow across the lifetime (Park *et al.*, 2002; Drag and Bieliauskas, 2010). Therefore, it was expected that higher linguistic scores would be associated with higher PR in both age groups. Despite the intuitive relationship between being able to generate words based on a single phoneme and being able to identify words based on partial information, the LME model analyses revealed that linguistic skills significantly mediated PR benefits with vocoded BKB sentences containing NBAV interruptions only, and only in the low spectral resolution condition (Table IV). Failure to observe broader effects of linguistic skills on PR, even in unprocessed speech conditions, may indicate that the linguistic measures used in

the present study were inadequate for measuring participants' language skills.

While working memory has not previously been shown to mediate PR in young participants (Benard *et al.*, 2014), we posited that the skill may have a mediating effect in older participants. Working memory ability decreases with age (Park *et al.*, 2002; Tulsy *et al.*, 2014), and better working memory skills have been associated with better speech-in-noise processing in older participants and better ability to utilize context cues to perceive speech (Gordon-Salant and Cole, 2016). Since top-down repair mechanisms like PR are believed to use context to restore speech, the ability to store and apply context information to incoming interrupted speech in working memory may be crucial, especially in older participants. However, in the present study, none of the four LME model analyses showed an effect of working memory specifically on PR, for either age group (Tables II, III, IV, and V). When considering interrupted speech understanding in general, across interruption types, higher working memory scores were associated with significantly better performance in ONH participants compared to YNH participants in vocoded IEEE sentence contexts (Table V). In summary, neither linguistic nor cognitive skills were strongly implicated as mediating PR in YNH and ONH participants, in either unprocessed or vocoded listening conditions.

E. Conclusion

In conclusion, ONH participants can perceptually restore NB-interrupted sentences, even when speech signals are degraded to simulate aspects of CI processing. That is, the interaction of aging and signal degradation does not erase older participants' ability to access and utilize context and other top-down PR mechanisms. Spectral resolution appears to have little effect on the magnitude of the PR effect for the resolutions that were tested; ONH participants obtained similar PR benefits in both 16- and 32-channel noise-vocoded speech. The quality of the speech signal—the bottom-up acoustic information—did not significantly impair interaction with top-down repair mechanisms, at least in ONH participants. Furthermore, ONH participants used both vocoded and non-vocoded NB interruptions to restore degraded, vocoded speech. The presence of *any* noise was useful to ONH participants, and could be captured into the target speech stream to create a continuous auditory illusion percept. These findings help inform future directions for PR research in CI users. Factors beyond spectral resolution and spectral differences between speech and interrupting noise should be investigated next. For example, noise reduction strategies and small dynamic ranges in CIs may be impacting the perception of the relationships between speech and noise (Mauger *et al.*, 2012), affecting the normal course of repair mechanisms like PR. Ensuring that CI users have access to the beneficial aspects of noise may be one way to improve speech perception in real-life listening conditions.

ACKNOWLEDGMENTS

Thank you to Kelly Miller, Stefanie Kuchinsky, Hannah Cohen, Stephen Fong, Hannah Johnson, Emily Waddington,

Tracy Wilkinson, and Lauren Wilson for their assistance with data collection and analysis. Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award No. R01AG051603. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work was also supported by training Grant No. T32DC00046 from the National Institute of Deafness and Communicative Disorders of the National Institutes of Health. Portions of this work were presented at the 40th MidWinter Meeting of the Association for Research in Otolaryngology and the 2017 Aging and Speech Communication Conference.

¹Two of the 17 ONH participants had a threshold of 30 dB HL at one or more frequencies at or below 4000 Hz. *Post hoc* testing of the relationships between audiometric thresholds and PR task performance showed no significant correlations. Furthermore, each of these two participants' average performances with interrupted speech was within ± 1 SD of the average performance of the rest of the participants, respectively. Thus, these two participants' data were retained, and included in the present study's results and analyses.

- Anderson, S., Parbery-Clark, A., White-Schwoch, T., and Kraus, N. (2012). "Aging affects neural precision of speech encoding," *J. Neurosci.* **32**, 14156–14164.
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). "Random effects structure for confirmatory hypothesis testing: Keep it maximal," *J. Mem. Lang.* **68**, 255–278.
- Bashford, J. A., Riener, K. R., and Warren, R. M. (1992). "Increasing the intelligibility of speech through multiple phonemic restorations," *Percept. Psychophys.* **51**, 211–217.
- Başkent, D. (2012). "Effect of speech degradation on top-down repair: Phonemic restoration with simulations of cochlear implants and combined electric-acoustic stimulation," *J. Assoc. Res. Otol.* **13**, 683–692.
- 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.
- Bench, J., Kowal, A., and Bamford, J. (1979). "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children," *Brit. J. Audiol.* **13**, 108–112.
- Bhargava, P., Gaudrain, E., and Başkent, D. (2014). "Top-down restoration of speech in cochlear-implant users," *Hear. Res.* **309**, 113–123.
- Clarke, J., Başkent, D., and Gaudrain, E. (2016). "Pitch and spectral resolution: A systematic comparison of bottom-up cues for top-down repair of degraded speech," *J. Acoust. Soc. Am.* **139**, 395–405.
- Crystal, T. H., and House, A. S. (1990). "Articulation rate and the duration of syllables and stress groups in connected speech," *J. Acoust. Soc. Am.* **88**, 101–112.
- Drag, L. L., and Bieliauskas, L. A. (2010). "Contemporary review 2009: Cognitive aging," *J. Geriatr. Psych. Neur.* **23**, 75–93.
- Fetterman, B. L., and Domico, E. H. (2002). "Speech recognition in background noise of cochlear implant patients," *Otolaryng. Head Neck* **126**, 257–263.
- Fitzgibbons, P. J., and Gordon-Salant, S. (1996). "Auditory temporal processing in elderly listeners," *J. Am. Acad. Aud.* **7**, 183–189.
- Friesen, L. M., Shannon, R. V., Başkent, D., and Wang, X. (2001). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," *J. Acoust. Soc. Am.* **110**, 1150–1163.
- Glinberg and Associates, Inc. (2016). NIH Toolbox (Version 1.7) [Mobile application software]. Retrieved from <http://itunes.apple.com> (Last viewed June 13, 2017).
- Gnansia, D., Pressnitzer, D., Pean, V., Meyer, B., and Lorenzi, C. (2010). "Intelligibility of interrupted and interleaved speech for normal-hearing listeners and cochlear implantees," *Hear. Res.* **265**, 46–53.
- Gordon-Salant, S., and Cole, S. S. (2016). "Effects of age and working memory capacity on speech recognition performance in noise among listeners with normal hearing," *Ear Hear.* **37**, 593–602.
- Goupell, M. J., Gaskins, C. R., Shader, M. J., Walter, E. P., Anderson, S., and Gordon-Salant, S. (2017). "Age-related differences in the processing of temporal envelope and spectral cues in a speech segment," *Ear Hear.* **38**(6), e335–e342.
- Harrison, J. E., Buxton, P., Husain, M., and Wise, R. (2000). "Short test of semantic and phonological fluency: Normal performance, validity, and test-retest reliability," *Brit. J. Clin. Psychol.* **39**, 181–191.
- Hox, J. J., Moerbeek, M., and van de Schoot, R. (2017). *Multilevel Analysis: Techniques and Applications*, 3rd ed. (Routledge, New York).
- Levelt, W. J. M. (2001). "Spoken word production: A theory of lexical access," *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13464–13471.
- Litovsky, R. Y., Goupell, M. J., Godar, S., Grieco-Calub, T., Jones, G. L., Garadat, S. N., Agrawal, S., Kan, A., Todd, A., Hess, C., and Misurelli, S. (2012). "Studies on bilateral cochlear implants at the University of Wisconsin's Binaural Hearing and Speech Laboratory," *J. Am. Acad. Audiol.* **23**, 476–494.
- Loebach, J. L., and Pisoni, D. B. (2008). "Perceptual learning of spectrally degraded speech and environmental sounds," *J. Acoust. Soc. Am.* **123**, 1126–1139.
- Mauger, S. J., Dawson, P. W., and Hersbach, A. A. (2012). "Perceptually optimized gain function for cochlear implant signal-to-noise ratio based noise reduction," *J. Acoust. Soc. Am.* **131**, 327–336.
- Miller, G. A., and Licklider, J. C. R. (1950). "The intelligibility of interrupted speech," *J. Acoust. Soc. Am.* **22**, 167–173.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., and Chertkow, H. (2005). "The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment," *J. Am. Geriatrics Soc.* **53**, 695–699.
- 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.
- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., and Smith, P. K. (2002). "Models of visuospatial and verbal memory across the adult life span," *Psychol. Aging.* **17**, 299–320.
- Pichora-Fuller, M. K. (2008). "Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing," *Int. J. Audiol.* **47**, S72–S82.
- Pichora-Fuller, M. K., Schneider, B. A., and Daneman, M. (1995). "How young and old adults listen to and remember speech in noise," *J. Acoust. Soc. Am.* **97**, 593–608.
- Pichora-Fuller, M. K., and Souza, P. E. (2003). "Effects of aging on auditory processing of speech," *Int. J. Audiol.* **42**, S11–S16.
- 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.
- Rothauer, E., Chapman, W., Guttman, N., Nordby, K., Silbiger, H., Urbanek, G., and Weinstock, M. (1969). "IEEE recommended practice for speech quality measurements," *IEEE Trans. Audio Electroacoust.* **17**, 225–246.
- Saija, J. D., Akyurek, E. G., Andringa, T. C., and Başkent, D. (2014). "Perceptual restoration of degraded speech is preserved with advancing age," *J. Assoc. Res. Otol.* **15**, 139–148.
- Samuel, A. G. (1981). "Phonemic restoration: Insights from a new methodology," *J. Exp. Psychol. Gen.* **110**, 474–494.
- Schneider, B. A., Pichora-Fuller, K., and Daneman, M. (2010). "Effects of senescent changes in audition and cognition on spoken language comprehension," in *The Aging Auditory System*, edited by S. Gordon-Salant, R. D. Frishta, A. N. Popper, and R. R. Fay (Springer, New York), pp. 167–210.
- Schoof, T., and Rosen, S. (2015). "High sentence predictability increases the fluctuating masker benefit," *J. Acoust. Soc. Am.* **138**, EL181–EL186.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Sheldon, S., Pichora-Fuller, M. K., and Schneider, B. A. (2008). "Priming and sentence context support listening to noise-vocoded speech by younger and older adults," *J. Acoust. Soc. Am.* **123**, 489–499.
- Sladen, D. P., and Zappler, A. (2015). "Older and younger adult cochlear implant users: Speech recognition in quiet and noise, quality of life, and music perception," *Am. J. Audiol.* **24**, 31–39.
- Sommers, M. S., and Danielson, S. M. (1999). "Inhibitory processes and spoken word recognition in young and older adults: The interaction of lexical competition and semantic context," *Psychol. Aging* **14**, 458–472.
- Strouse, A., Ashmead, D. H., Ohde, R. N., and Grantham, D. W. (1998). "Temporal processing in the aging auditory system," *J. Acoust. Soc. Am.* **104**, 2385–2399.

- Studebaker, G. A. (1985). "A 'rationalized' arcsine transform," *J. Speech Hear. Res.* **28**, 455–462.
- Tombaugh, T. N., Kozak, J., and Rees, L. (1999). "Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming," *Arch. Clin. Neuropsychol.* **14**, 167–177.
- Tulsky, D. S., Carlozzi, N., Chiaravalloti, N. D., Beaumont, J. L., Kisala, P. A., Mungas, D., Conway, K., and Gershon, R. (2014). "NIH Toolbox Cognition Battery (NIHTB-CB): List Sorting Test to measure working memory," *J. Int. Neuropsych. Soc.* **20**, 599–610.
- Verschuure, J., and Brocaar, M. P. (1983). "Intelligibility of interrupted meaningful and nonsense speech with and without intervening noise," *Percept. Psychophys.* **33**, 232–240.
- Versfeld, N. J., Daalder, L., Festen, J. M., and Houtgast, T. (2000). "Method for the selection of sentence materials for efficient measurement of the speech reception threshold," *J. Acoust. Soc. Am.* **107**, 1671–1684.
- Warren, R. M. (1970). "Perceptual restoration of missing speech sounds," *Science* **167**, 392–393.
- Whitmal, N. A., Poissant, S. F., Freyman, R. L., and Helfer, K. S. (2007). "Speech intelligibility in cochlear implant simulations: Effects of carrier type, interfering noise, and subject experience," *J. Acoust. Soc. Am.* **122**, 2376–2388.
- Working Group on Speech Understanding and Aging (1988). "Speech understanding and aging," *J. Acoust. Soc. Am.* **83**, 859–895.
- Zarino, B., Crespi, M., Launi, M., and Casarotti, A. (2014). "A new standardization of semantic verbal fluency test," *Neurol. Sci.* **35**, 1405–1411.