Spectrotemporal modulation sensitivity for hearing-impaired listeners: Dependence on carrier center frequency and the relationship to speech intelligibility

Golbarg Mehraei

Program in Speech and Hearing Bioscience and Technology, Harvard University–Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Frederick J. Gallun and Marjorie R. Leek

VA RR&D National Center for Rehabilitative Auditory Research, Portland VA Medical Center, Portland, Oregon 97239

Joshua G. W. Bernstein^{a)}

National Military Audiology and Speech Pathology Center, Walter Reed National Military Medical Center, Bethesda, Maryland 20889

(Received 13 February 2013; revised 29 April 2014; accepted 10 May 2014)

Poor speech understanding in noise by hearing-impaired (HI) listeners is only partly explained by elevated audiometric thresholds. Suprathreshold-processing impairments such as reduced temporal or spectral resolution or temporal fine-structure (TFS) processing ability might also contribute. Although speech contains dynamic combinations of temporal and spectral modulation and TFS content, these capabilities are often treated separately. Modulation-depth detection thresholds for spectrotemporal modulation (STM) applied to octave-band noise were measured for normal-hearing and HI listeners as a function of temporal modulation rate (4–32 Hz), spectral ripple density [0.5–4 cycles/octave (c/o)] and carrier center frequency (500–4000 Hz). STM sensitivity was worse than normal for HI listeners only for a low-frequency carrier (1000 Hz) at low temporal modulation rates (4–12 Hz) and a spectral ripple density of 2 c/o, and for a high-frequency carrier (4000 Hz) at a high spectral ripple density (4 c/o). STM sensitivity for the 4-Hz, 4-c/o condition for a 4000-Hz carrier and for the 4-Hz, 2-c/o condition for a 1000-Hz carrier were correlated with speech-recognition performance in noise after partialling out the audiogram-based speech-intelligibility index. Poor speech-reception and STM-detection performance for HI listeners may be related to a combination of reduced frequency selectivity and a TFS-processing deficit limiting the ability to track spectral-peak movements. [<http://dx.doi.org/10.1121/1.4881918>]

PACS number(s): 43.66.Ba, 43.71.Rt, 43.66.Fe, 43.66.Sr [TD]

Pages: 301–316

I. INTRODUCTION

One of the most common complaints reported by listeners with sensorineural hearing loss is difficulty understanding speech in noisy situations. Although amplification via a hearing aid compensates for speech perception to some extent, audibility does not account for the entire deficit in speech recognition. Specifically, individual differences in the audiogram only predict about half of the variance in speech intelligibility in noise across hearing-impaired (HI) individuals (e.g., [Smoorenburg, 1992](#)). Hearing loss is also associated with certain abnormalities in the processing of sounds presented at suprathreshold levels by the peripheral auditory system, which might contribute to an impaired ability to understand speech in noise ([Plomp, 1986](#)). The internal representation of sounds made audible via amplification might nevertheless be distorted by abnormal processing by the cochlea or auditory nerve. Investigations of the relationship between suprathreshold processing ability and speech perception by HI listeners have mainly focused on three

aspects of hearing acuity: spectral resolution, temporal resolution, and temporal fine-structure (TFS) processing. Recent results have shown that the intersubject differences in the ability to detect the presence of spectrotemporal modulation (STM) applied to a broadband noise carrier can account for a substantial proportion of the variance in speech understanding in noise ([Bernstein et al., 2013a](#)). STM signals contain elements of spectral and temporal modulation as well as TFS information. The goal of the current study was to further examine the relationship between STM sensitivity and speech-reception performance, and to elucidate the psychophysical mechanisms underlying this relationship, by examining STM detection performance for octave-band noise carriers as a function of carrier center frequency.

There is strong evidence that listeners with sensorineural hearing loss have impaired spectral resolution compared to normal-hearing (NH) listeners (e.g., [Glasberg and Moore, 1986](#)), which is thought to reflect a reduction of the active cochlear mechanism that provides frequency-specific gain at each location along the cochlear partition (e.g., [Ruggero et al., 1997](#)). A reduction in spectral resolution could smear the spectral details in the internal representation of a speech signal that relay consonant and vowel information. Reduced

^{a)}Author to whom correspondence should be addressed. Electronic mail: joshua.g.bernstein.civ@health.mil

spectral resolution could also increase the masking effect for portions of a noise signal with frequency content falling near important speech features. NH listeners presented with speech processed to smear the spectrum of the speech signal show reduced speech-reception performance (Baer and Moore, 1993), providing some evidence for a relationship between reduced spectral resolution and speech reception ability. However, the evidence for a relationship between psychophysical estimates of spectral resolution and speech reception ability in noise across HI listeners is mixed, with some studies observing a significant relationship between the two types of measure, and other studies failing to observe such a relationship (Smootenburg, 1992; Buss *et al.*, 2004; Summers *et al.*, 2013).

The evidence for reduced temporal processing ability in HI listeners is also mixed. Temporal processing ability can be measured in different ways, with orthogonal results. Studies investigating amplitude-modulation (AM) detection have generally found that performance is not significantly affected by sensorineural hearing loss (Moore *et al.*, 1992). On the other hand, certain studies have identified an increased effect of forward masking for HI listeners. This is likely related to a loss of the compressive gain of the cochlear active mechanism, whereby impaired listeners do not receive a benefit of larger gain applied to a low-level signal that immediately follows a higher-level masker (Nelson *et al.*, 2001). Temporal resolution has also been shown to be negatively impacted by age, with older listeners with near-normal audiograms nevertheless demonstrating reduced temporal modulation sensitivity (e.g., He *et al.*, 2008). Reduced temporal resolution could disrupt speech reception by smearing the temporal features present in the speech signal or by increasing the effects of forward masking. Although several studies have investigated the relationship between measures of temporal resolution and speech reception performance across HI listeners, the only evidence for such a relationship has been observed in situations involving noise maskers modulated at rates greater than 16 Hz (George *et al.*, 2006), perhaps reflecting the effects of increased forward masking. Thus, in general, reduced temporal resolution does not appear to negatively impact speech-reception ability in noise.

There is some evidence in the literature that HI listeners have a reduced ability to use TFS information, i.e., rapidly changing variation in the stimulus waveform. HI listeners are unable to understand speech that is processed to remove relatively slowly varying envelope information as much as possible, leaving only TFS information in the stimulus (Lorenzi *et al.*, 2006). HI listeners are also less affected than NH listeners by the removal of TFS information from the stimulus via vocoder processing (Hopkins and Moore, 2007), suggesting that HI listeners may not be using TFS cues. In another example, HI listeners have less-than-normal ability to detect low-rate (around 2-Hz) frequency modulation (FM) applied to a relatively low-frequency carrier tone (Moore and Skrodzka, 2002; Buss *et al.*, 2004). This task is thought to depend on the ability to use TFS to detect changes in the carrier frequency (Moore and Sek, 1996; Moore and Skrodzka, 2002). Strelcyk and Dau (2009) found that the

detection of low-rate FM for a low-frequency carrier by HI listeners was correlated with sensitivity to interaural phase, but not with frequency selectivity or audiometric thresholds, further supporting the idea that FM detection for low-frequency carriers is constrained by the ability to process TFS information. Furthermore, masked speech-reception performance is correlated with the degree of impairment in tasks thought to be dependent on TFS-processing ability (Buss *et al.*, 2004; Lorenzi *et al.*, 2006; Hopkins and Moore, 2007; Gnansia *et al.*, 2009; Strelcyk and Dau, 2009). It should be noted, however, that because the HI listeners in many of these studies were often older than the NH controls, some of the apparent effects of hearing loss on TFS processing might reflect the effects of aging, rather than hearing loss *per se*. Previous studies have shown that TFS processing is correlated with age (Grose and Mamo, 2010; Hopkins and Moore, 2011; Neher *et al.*, 2012) and suggested that age, rather than hearing loss, is responsible for deficits in TFS processing that can impair speech perception (Moore *et al.*, 2011; Sheft *et al.*, 2012).

While the studies described above have established much about the effects of hearing impairment on spectral and temporal resolution and TFS processing, treating these aspects of auditory processing separately does not directly reflect the characteristics of natural sounds that contain combined spectral and temporal modulation (STM) and variable TFS information. Significant speech features such as formant peaks, spectral edges, and amplitude modulations at onsets and offsets contribute to the energy modulations seen in speech spectrograms, both in time for any given frequency channel, and along the spectral axis at any instant. Chi *et al.* (1999) measured sensitivity to combined spectral and temporal modulations using broadband STM stimuli in NH listeners. These measurements were used as a basis for the development of a computational model that successfully predicted speech intelligibility in noise and reverberation for NH listeners based on STM strength in the auditory periphery (Elhilali *et al.*, 2003). The idea that STM sensitivity is related to speech intelligibility raises the possibility that HI listeners have difficulty in understanding speech in noise because of deficits in their ability to detect STM.

Bernstein *et al.* (2013a) investigated this hypothesis by measuring broadband STM sensitivity for HI and NH listeners and comparing these measurements to speech-intelligibility scores and other psychoacoustic measures. Their study had three major findings. First, STM sensitivity was strongly correlated with HI speech-reception performance in noise, even after partialling out audiometric differences across listeners. Second, STM sensitivity was impaired for HI listeners only for conditions that involved higher spectral ripple densities (i.e., more closely spaced spectral peaks) and low temporal modulation rates. Third, the variance in STM sensitivity observed across HI listeners was largely accounted for by psychoacoustic measures of frequency selectivity at high frequencies (4000 Hz) and estimates of TFS processing ability (FM-detection performance) at low frequencies (500 Hz). These results were interpreted to suggest that STM sensitivity, and in turn speech reception

in noise, is negatively impacted by hearing loss through a combination of reduced frequency resolution (at high frequencies) and impaired ability to use TFS information (at low frequencies). As was observed in the data, reduced frequency selectivity would be expected to affect STM sensitivity more for higher spectral ripple densities, where the close spacing of spectral peaks is more likely to interact with limited spectral resolution. Also as observed in the data, a reduced ability to use TFS information would be expected to affect STM sensitivity more for lower temporal rates and for a non-zero spectral ripple density. This is because the TFS-processing mechanism is thought to be “sluggish” even for NH listeners, and only capable of tracking the moving frequencies of spectral peaks for relatively slow modulations (Moore and Sek, 1996). Because this TFS mechanism would not be operational at higher temporal modulation rates, even for NH listeners, a reduced ability to use TFS information would not be expected to affect HI performance for such fast-moving stimuli.

The broadband STM sensitivity results of Bernstein *et al.* (2013a) showed a clear relationship with speech-reception performance and provided some clues about the mechanisms underlying STM sensitivity deficits for HI listeners. Broadband stimuli were chosen in that experiment to provide a controlled psychoacoustic test with stimuli as similar to speech signals as possible. This choice could have been a factor in the strong correlations observed between speech scores and STM sensitivity because, like speech discrimination, broadband STM detection relied on a host of psychoacoustic abilities and the health of the cochlea across the cochlear partition. However, the ability to infer the underlying causes of reduced STM sensitivity from these data, and to determine which aspects of STM processing are related to speech intelligibility, is limited because of the broadband nature of the stimuli. Differences in the impact of hearing loss across frequency may differentially affect STM sensitivity.

The aim of the current study was to investigate how hearing loss affects sensitivity to STM as a function of carrier center frequency and how STM sensitivity at different carrier center frequencies relates to speech-reception performance in noise. Experiment 1 measured STM sensitivity for NH and HI listeners as a function of spectral ripple density and temporal modulation rate for octave-band carriers centered at 500, 1000, 2000, and 4000 Hz, and examined the relationship between STM sensitivity and previously published data (Summers *et al.*, 2013) measuring speech reception in noise for the same HI listeners. The goal was to elucidate the mechanisms responsible for reduced STM sensitivity and the associated speech-intelligibility deficits for HI listeners. Experiment 2 explored the possible role of a spectral-edge cue, rather than reduced TFS processing ability, in driving the pattern of conditions where STM sensitivity differed between the NH and HI listeners in experiment 1. Experiment 2 also addressed the issue of the possible role of age differences (rather than hearing loss) in the STM sensitivity differences between NH and HI listeners observed in experiment 1 by testing a subset of NH and HI listeners in a similar age range.

II. EXPERIMENT 1. STM SENSITIVITY MEASUREMENTS

A. Methods

1. Stimuli

Narrowband STM stimuli were produced similarly to the broadband (four-octave) stimuli of Bernstein *et al.* (2013a), except that the modulated noise bands were limited to one octave logarithmically centered at 500, 1000, 2000, or 4000 Hz. A schematic of the stimulus construction method is shown in Fig. 1. The noise carrier consisted of a series of 1000 equal-amplitude tones logarithmically spaced across each octave band [Fig. 1(a)]. The carrier tones were random phase, with a different starting phase randomly selected for each tone component, and a new random selection of phases generated on each stimulus trial. For clarity, only a subset of the 1000 carrier tones in each octave are depicted in Fig. 1. Sinusoidal AM was applied to each carrier tone within the octave of interest by adding sidebands above and below the carrier-tone frequency with the appropriate amplitude and for the desired modulation depth and phase. This is shown in Fig. 1(b), which is an expanded view of the region around 1000 Hz from the broadband spectrum depicted in Fig. 1(a). Spectral modulation was induced concurrently with the temporal modulation by adjusting the relative phase of the temporal modulation applied to each successive carrier tone to yield a sinusoidal envelope at each point in time along the logarithmic frequency axis. Figure 1(c) depicts this relative phase shift, with each trace representing the amplitude envelope for one carrier tone. The STM stimulus was generated by applying an inverse fast-Fourier transform to the spectrum depicted in Fig. 1. Mathematically, the STM stimulus was characterized by

FIG. 1. (a) A schematic of the STM stimulus generation method in the spectral domain, showing the carrier components and sidebands in the modulated octave band centered at 1000 Hz and the unmodulated components 15 dB lower in level in the remaining octave bands. (b) A close-up view of the portion of the spectrum near the center of the modulated octave band, showing the carrier components and modulation sidebands. (c) A schematic of the temporal envelope waveforms associated with successive carrier components illustrating how relative shifts in the phases of the spectral envelopes generate modulation in the spectral domain.

$$S(x, t) = m \sin[2\pi(\omega t + \Omega x) + \varphi], \quad (1)$$

where S represents the amplitude of each carrier tone as a function of time (t) and frequency (x), ω is the STM modulation rate (Hz), and Ω represents the spectral ripple density (cycles/octave, c/o). The starting phase (φ) of the spectral modulation was selected randomly on each stimulus presentation. The amplitude (m) of the modulation was determined by adaptively adjusting the modulation depth of all the carrier tones simultaneously (0 = no modulation and 1 = full modulation).

The experiment was initially designed so that the stimuli would be as similar to the four-octave modulated broadband stimuli of Bernstein *et al.* (2013a) by presenting the entire four-octave noise carrier to the listener, but only applying the modulation to the octave-band in question. However, STM detection proved to be too difficult for some listeners for this type of stimulus. As a compromise, the levels of the components of the unmodulated carriers in the three remaining octaves were decreased by 15 dB relative to the components of the carrier components in the modulated octave-band [Fig. 1(a)]. With stimuli presented at an overall nominal stimulus level of 86 dB sound pressure level (SPL), this resulted in the modulated octave band presented at a nominal level of 85.6 dB SPL, while each of the three non-modulated octave bands was presented at a level of 70.6 dB SPL.

An unmodulated stationary noise was also created to serve as the comparison stimulus on each trial. This stimulus consisted of only the random-phase carrier tones with no modulation sidebands added. The unmodulated stimulus had the same spectral characteristics as the modulated stimulus, with the carrier components in the relevant octave band (i.e., the band that was modulated in the STM interval of each trial) presented at a nominal level of 85.6 dB SPL, while the other three octave band components were 15 dB lower.

2. Procedure

STM detection thresholds were measured using a two-alternative forced-choice adaptive procedure. Listeners were asked to discriminate between an unmodulated stationary noise (comparison stimulus) and a STM stimulus. The STM stimulus was randomly presented in the first or second interval, each with a probability 0.5. The listener's task on each trial was to indicate, by touching one of two virtual buttons on a touch screen, which of two sequential stimulus intervals contained the modulated stimulus. The modulation depth was varied in a three-down one-up adaptive procedure tracking the 79.4% correct point (Levitt, 1971) with step sizes computed in decibels (dB, $20 \log m$, where m is the modulation depth). The starting modulation depth was 0 dB (full modulation). The modulation depth changed by 6 dB until the first reversal point, by 4 dB for the next two reversal points, and by 2 dB for the last six reversal points, for a total of nine reversals per run. The mean of the modulation depths across the last six reversal points was defined to be the threshold. The modulation depth was not allowed to exceed 0 dB (full modulation). If the tracking procedure required a modulation depth greater than this value on any given trial,

then the modulation depth was kept at 0 dB for that trial. A run was terminated early and discarded if there were five incorrect responses at full modulation depth within the run.

Because the signal interval contained additional stimulus energy in the modulated band (due to the addition of the modulation sidebands), there was a greater difference in spectrum level between the modulated and unmodulated bands for the signal interval than for the standard-noise interval of a trial. The spectrum level was always 15 dB higher within the modulated octave band than within the unmodulated bands for the standard noise, so the difference in spectrum level ranged from 15 (for zero modulation depth) to 16.76 dB (for full modulation depth) for the signal interval. Thus, it is possible that listeners could have used this spectral difference, or a comparison of the levels of the stimulus within the unmodulated bands in the two stimulus trials, as an alternative cue to identify the signal interval. To reduce the effectiveness of level cues, the actual stimulus level in each interval of the trial was independently adjusted by a value randomly selected from a uniformly distributed 5-dB range (± 2.5 dB) around this nominal level. This manipulation would have reduced the effectiveness of a strategy that compared the levels of the unmodulated carrier bands in the two intervals. However, it is still theoretically possible that listeners could have used a profile-analysis strategy by comparing the relative signal levels in the modulated and unmodulated bands, especially at greater modulation depths. A supplementary experiment (see the Appendix) investigated the potential effectiveness of this alternative cue in a group of NH listeners. This experiment found that the level difference required for the identification of the modulated interval was greater than the 1.76 dB yielded by the addition of sidebands for an STM stimulus with full modulation depth ($m = 1$). Thus, it is unlikely that listeners were able to use this alternative cue for STM detection.

Presenting the stimulus at the same relatively high level to both listener groups was intended to reduce the effects of stimulus-level differences on frequency selectivity (Glasberg and Moore, 2000) and to match the high presentation level employed by Summers *et al.* (2013) to measure speech reception in noise for the same HI listeners who participated in the current study. An uncorrelated and unmodulated four-octave Gaussian noise was presented 20 dB lower in level than the test stimulus to the non-test ear to ensure that the detection was performed using the intended test ear. This step was taken because in the absence of a contralateral masker, the open design of the Sennheiser HD580 headphones increased the likelihood that signals presented to the test ear could be detected at the non-test ear. All signals were 500 ms in duration, including 20-ms raised cosine ramps. The two intervals of the trial were separated by 200 ms.

Listeners were presented with two test runs for each combination of spectral ripple density (0.5, 1, 2, or 4 c/o), temporal modulation rate (4, 12, or 32 Hz), carrier center frequency (500, 1000, 2000, or 4000 Hz) and direction (upward or downward). A third threshold was collected for a given condition if the first two STM-detection threshold estimates differed by 3 dB or more, or if one of the two runs was

terminated early due to more than five incorrect responses at full modulation depth. A fourth threshold estimate was performed if two of the first three threshold estimates collected for a specific condition differed by more than 6 dB. A short visual feedback (correct/incorrect) was displayed after each trial.

Digital stimuli were generated using MATLAB at a sampling frequency of 24.414 kHz. On each interval of the stimulus trial, a two-channel signal was sent to a Tucker-Davis Technologies System III Enhanced Real-Time Processor (TDT RP 2.1), where it was stored in a buffer. Digital-to-analog conversion was then initiated by a command from the PC. The audio signal was passed through a pair of programmable attenuators (TDT PA5) that were used to implement the ± 2.5 dB level roving and the 20-dB attenuation of the signal in the non-test ear, then through a headphone buffer (TDT HB7) before being presented to the listener through Sennheiser HD580 open circumaural headphones. The system was calibrated by placing each earpiece on a flat-plate Artificial Ear Coupler (Larson-Davis AEC101) and measuring the acoustic level (dB SPL, unweighted) of an unmodulated broadband (four octaves, 353–5656 Hz) reference noise stimulus applied to the earpiece at a root-mean-squared level of 4 V.

3. Subjects

Eight NH listeners (four female, mean age: 44.5, age range: 24–60) and 12 HI listeners (one female, mean age: 75.7, age range: 70–87) took part in this study. These were the same listeners that participated in the broadband STM study of Bernstein *et al.* (2013a). Of the 20 listeners, 15 (six NH, nine HI) were tested at Walter Reed Army Medical Center, Washington, DC, and five (two NH, three HI) were tested at the National Center for Rehabilitative Auditory Research, Portland, OR. The mean audiogram (± 1 standard error) of the tested ear for each listener group is shown in Fig. 2. NH listeners had pure-tone thresholds better than or equal to 20 dB hearing level (HL) at octave frequencies between 250–8000 Hz +3000 and 6000 Hz. The HI listeners all had sensorineural hearing loss, as evidenced by the absence of an

FIG. 2. Mean audiograms ($\pm$ one standard error) for the 12 HI and 8 NH listeners that participated in experiment 1.

air-bone gap greater than 10 dB. On average, the HI listeners had high-frequency hearing loss and thresholds of 30 dB HL or better at 1000 Hz and below. The better ear, defined as the ear that had thresholds equal to or better than the other ear at each of these audiometric frequencies, was tested for each HI listener. For cases where different ears had better thresholds at different frequencies, the ear that produced a better (lower) threshold for a 1000-Hz tone was tested. NH listeners were tested in the ear of their choice.

Each subject completed a minimum of one hour of training before data collection began. Training was done on a pseudorandom sampling of the STM conditions presented in the experiment, with a greater number of training runs presented for stimuli with higher spectral ripple densities and lower carrier center frequencies where listeners experienced the most difficulty.

B. Results

Mean STM detection thresholds across eight NH (open symbols) and 12 HI (filled symbols) listeners are shown in Fig. 3 as a function of temporal modulation rate (ω , horizontal axis). Each column in Fig. 3 represents a given carrier center frequency and each row represents a given spectral ripple density. More negative (lower) dB values in Fig. 3 indicate better performance, with STM detectable at smaller modulation depths. Some data points are shifted horizontally for clarity. The main goal of this experiment was to identify the combinations of carrier center frequency, rate, and density for which STM sensitivity was negatively impacted by hearing loss. Three specific hypotheses regarding the influence of spectral resolution, temporal resolution and TFS processing on STM sensitivity for HI listeners were tested by examining the combination of stimulus parameters for which performance was significantly poorer for the HI than for the NH listeners:

- (1) An influence of *impaired temporal resolution* on STM sensitivity would be suggested by a greater impact of hearing loss mainly at higher temporal modulation rates. This is because the longer modulation period associated with low temporal modulation rates would be relatively immune to increased forward masking or an increased temporal integration window.
- (2) An influence of *impaired spectral resolution* on STM sensitivity would be suggested by a greater impact of hearing loss mainly at higher spectral ripple densities, especially at higher carrier center frequencies where Summers *et al.* (2013) showed elevated frequency selectivity for the same HI subjects who participated in the current study.
- (3) An influence of *impaired TFS processing ability* on STM sensitivity would be suggested by an impact of hearing loss mainly at lower temporal modulation rates, where TFS information is most likely to be used by NH listeners to detect changes in the frequencies of spectral peaks (Moore and Sek, 1996). The effect would be greatest for lower carrier center frequencies where phase-locking information is best represented in auditory-nerve firing patterns (Johnson, 1980).

FIG. 3. Group-mean STM-detection thresholds averaged across upward- and downward-moving conditions. Columns represent individual carrier center frequencies (500, 1000, 2000, and 4000 Hz). Rows represent individual spectral ripple densities (0.5, 1, 2, and 4 c/o). The horizontal axis in each panel represents temporal modulation rate. The HI data are shifted horizontally for clarity. Error bars indicate $\pm$ one standard error across the listeners in each group. Asterisks indicate conditions where NH and HI STM-detection thresholds were significantly different.

The data were analyzed using a repeated-measures analysis of variance (ANOVA) conducted with four within-subjects factors (carrier center frequency, temporal modulation rate, spectral ripple density, and direction of movement) and one between-subjects factor (hearing status; i.e., NH or HI). Floor effects were observed in some conditions involving low carrier center frequencies, and higher rates and densities, whereby some listeners were not able to achieve a threshold level of performance at full modulation depth. For analysis purposes, detection thresholds for individual listeners in these cases were set to 0 dB (full modulation depth). To address the main experimental question regarding the impact of hearing loss on STM sensitivity performance, we were particularly interested in interactions between hearing status and each of the within-subjects factors. For example, an interaction between hearing status and carrier center frequency would suggest that hearing loss had a larger impact

in some frequency regions than in others. The STM thresholds presented in Fig. 3 were averaged across upward- and downward-moving conditions because there were no statistically significant interactions between hearing status and direction ($p > 0.05$).

Although there was not a significant main effect of hearing status ($p = 0.33$), the ANOVA identified significant interactions between hearing status and spectral ripple density [$F(2.35, 42.3) = 14.1$, $p < 0.005$] between hearing status, temporal modulation rate, and spectral ripple density [$F(6, 108) = 3.65$, $p < 0.01$], and between hearing status, spectral ripple density and carrier center frequency [$F(8.83, 158) = 5.18$, $p < 0.005$]. None of the other interactions involving hearing status were significant ($p > 0.05$). *Post hoc* tests were conducted to determine which conditions were responsible for the significant interactions between hearing loss and the experimental variables (temporal

modulation rate, spectral ripple density and carrier center frequency) observed in the ANOVA. One-tailed *t*-tests comparing NH and HI performance were computed on the STM data collapsed across direction. Although there were a total of 48 conditions for which performance for the NH and the HI group was compared (three temporal modulation rates $\times$ four spectral ripple densities $\times$ four carrier center frequencies), *p*-values were Bonferroni corrected for only 12 multiple comparisons (three temporal modulation rates $\times$ four spectral ripple densities). The previous results of Bernstein *et al.* (2013a) identified a particular pattern of the effect of hearing loss on STM sensitivity as a function of spectral ripple density and temporal modulation rate for a broadband (four-octave) carrier. Performance differences between the two listener groups were mainly observed at low temporal modulation rates (suggesting an influence of reduced TFS processing ability) and at high spectral ripple densities (suggesting an influence of reduced frequency selectivity on STM sensitivity for the HI listeners). The question in this study was whether this pattern of results would be reflected in each frequency region. Thus, the comparisons at each carrier center frequency were planned, but the *post hoc* analyses included corrections for the tests involving the 12 combinations of spectral ripple density and temporal modulation rate within each carrier-frequency condition (Keppel and Wickens, 2004).

Significant sensitivity differences between NH and HI listener groups ($p < 0.05$) were observed for three combinations of spectral ripple density, temporal modulation rate, and frequency. The conditions for which HI demonstrated poorer-than-normal performance, indicated by asterisks in Fig. 3, mainly involved lower temporal modulation rates (4 and 12 Hz) and higher spectral ripple densities (2 and 4 c/o) for carrier center frequencies of 1000 and 4000 Hz. For a 1000-Hz carrier center frequency, performance was significantly poorer for the HI listeners for a spectral ripple density of 2 c/o and temporal modulation rates of 4 and 12 Hz. For a 4000-Hz carrier center frequency, performance was significantly poorer for the HI listeners for a spectral ripple density of 4 c/o and a temporal modulation rate of 4 Hz. Performance differences between NH and HI listeners were non-significant ($p > 0.05$) for all other conditions. It should be noted that the ability to observe effects of hearing loss on STM sensitivity at 4 c/o for the lower carrier center frequencies (500 and 1000 Hz) were likely limited by floor effects, as STM detection thresholds were measured near 0 dB or could not be measured at all for some listeners in both groups for these conditions.

C. The relationship between STM sensitivity and speech intelligibility

The second goal of the study was to investigate whether individual differences in STM sensitivity are related to inter-subject variability in speech-reception scores. To determine the relationship between STM sensitivity and speech intelligibility, speech-intelligibility data for the HI subjects were taken from the study of Summers *et al.* (2013). Summers *et al.* measured speech intelligibility for IEEE (1969)

sentences spoken by a female talker for NH and HI listeners, including the 12 HI listeners who participated in the current study. Summers *et al.* presented the target speech at a level of 92 dB SPL in an attempt to overcome audibility limitations. They measured performance for speech presented in stationary and modulated speech-shaped noise across a range of SNRs (−6, −3, 0, and +3 dB).

The stationary-noise data were considered first. To extract a single performance metric from the data collected at four SNRs, the speech-reception threshold at a performance level of 50% keywords correct (SRT₅₀) was calculated from the speech-reception scores in stationary noise for each individual subject by fitting a sigmoidal function to the percentage-correct data and estimating the SNR required for 50% correct performance. The key question was whether the STM metric provided predictive power for speech-reception performance in noise beyond that provided by the audiogram. To ask this question, it was first necessary to determine the extent to which the audiogram could predict the speech scores. This was done by calculating an SII value for speech presented in noise at each of the four SNRs tested by Summers *et al.* (2013).

To make a prediction of the speech-reception performance based on the SII, it was necessary to transform the SII values to percentage-correct scores. Rather than use the ANSI-standard transformation, we instead plotted speech-reception performance for all of the HI listeners and SNRs against the calculated values, and fit a sigmoidal curve to the data. The fit suggested that, on average, speech performance of 50% keywords correct required an SII of 0.3649. The SII-predicted SRT₅₀ was therefore defined as the SNR that yielded an SII of 0.3649. This value was extracted for each HI listener by fitting a sigmoidal curve to the SII values plotted as a function of SNR for that individual.

Pairwise correlations were computed for the HI listeners between the SRT₅₀ metric extracted from the stationary-noise speech-reception data of Summers *et al.* (2013) (their Table 1) and the octave-band STM sensitivity scores measured in the current study. Only the three conditions for which HI listeners showed significantly poorer STM sensitivity performance than the NH listeners (asterisks in Fig. 3) were tested. One-tailed *p*-values were Bonferroni corrected for (three) multiple comparisons. STM detection thresholds were found to be significantly correlated with speech scores for the 4 c/o, 4 Hz condition for a 4000 Hz carrier ($R = 0.66$, $p < 0.05$). The correlations were not found to be significant for the two other STM conditions for which the HI group showed poorer STM sensitivity than the NH group (1000 Hz, 4 Hz, 2 c/o: $p = 0.08$; 1000 Hz, 12 Hz, 2 c/o: $p = 1$).

Correlations between STM sensitivity and speech-reception performance were re-computed after partialling out the SII-based SRT₅₀ prediction to determine whether the octave-band STM measure provided predictive power for speech-reception performance beyond that provided by the audiogram. After partialling out the contribution of the SII prediction, performance for both the 2 c/o, 4 Hz condition for a 1000-Hz carrier ($R = 0.74$, $p < 0.05$) and the 4 c/o, 4 Hz condition for a 4000-Hz carrier ($R = 0.70$, $p < 0.05$) were found to be significantly correlated to speech-reception

FIG. 4. The measured SRT_{50} is plotted as a function of the SII-based predictions of the SRT_{50} for individual HI listeners.

performance, while the third STM condition examined (1000 Hz, 12 Hz, 2 c/o) was not ($p = 0.21$).

A stepwise regression analysis was then conducted to determine the combined predictive power of the STM-sensitivity estimates for these particular conditions and the SII. As shown in Fig. 4, SII-based SRT_{50} predictions fell in a narrow range of SNRs between -0.2 and $+0.9$ dB, reflecting the fact that the SII values in noise are dominated by the statistics of the noise (the same for all subjects) rather than the differences in audiograms. Thus, audibility cannot account for the wide variation in measured SRT_{50} (-4 to $+4$ dB). The SII-based SRT_{50} predictions were nevertheless highly correlated with the measured SRT_{50} values, accounting for 59.4% of the variance in speech intelligibility ($R = 0.77$, $p < 0.005$). The addition of STM sensitivity for the low-frequency carrier (2 c/o, 4 Hz, 1000 Hz) as a second predictor variable significantly increased ($p < 0.05$) the overall proportion of the variance in speech-reception performance accounted for to 81.7% (not shown). The addition of STM sensitivity for the high-frequency carrier (4 c/o, 4 Hz, 4000 Hz) into the analysis as a third predictor variable significantly increased ($p < 0.05$) the overall variance accounted for to 89.9% (Fig. 5). Thus, performance for these two STM conditions accounted for an additional 30% of the variance in speech-reception performance beyond that accounted for by the audiogram-based SII.

Previous results have suggested a larger impact of hearing loss and suprathreshold auditory processing deficits on speech perception in modulated noise (e.g., Strelcyk and Dau, 2009). An additional analysis was carried out to determine the relationship between the SII, STM sensitivity, and the SRT_{50} obtained from the speech scores in speech-modulated noise reported by Summers *et al.* (2013). Pairwise correlations between this SRT_{50} metric and the octave-band STM sensitivity scores revealed no significant correlations for any of the three STM conditions for which the HI listeners performed significantly worse than the NH listeners. However, after partialling out the SII-based SRT_{50}

FIG. 5. The SRT_{50} measured for individual HI subjects is plotted as a function of the SRT_{50} predicted by a linear regression model with three inputs: the SII and STM sensitivity for the [4 Hz, 4 c/o, 4000 Hz] and [4 Hz, 2 c/o, 1000 Hz] conditions.

prediction, the STM sensitivity for the 2 c/o, 4 Hz condition for a 1000-Hz carrier ($R = 0.74$, $p < 0.05$) were significantly correlated with the SRT_{50} in speech-modulated noise.

A stepwise linear regression analysis showed that the SII-based SRT_{50} predictions were highly correlated with the measured SRT_{50} values, accounting for 72.1% of the variance in speech intelligibility in modulated noise ($R = 0.84$, $p < 0.005$), a somewhat larger proportion of the variance than in the stationary-noise case. This is consistent with the idea that audibility plays a larger role in limiting speech-reception performance in modulated noise, because absolute thresholds would tend to play a more significant role during silent gaps in the masker when little or no noise is present. The addition of STM sensitivity for the low-frequency carrier (2 c/o, 4 Hz, 1000 Hz) to the linear regression model increased ($p < 0.05$) the overall proportion of the variance in speech-reception performance accounted for to 87.5% (not shown). Thus, performance for this STM condition accounted for an additional 16% of the variance in speech-reception performance beyond that accounted for by the audiogram-based SII. Thus, the results for the speech-modulated noise data were similar to the results for the stationary-noise speech-reception data, except that STM sensitivity for the high-frequency carrier (4 c/o, 4 Hz, 4000 Hz) STM condition did not account for a significant proportion of the variance in speech scores for the modulated noise case.

III. EXPERIMENT 2. STM SENSITIVITY WITH SPECTRAL-EDGE CUES REMOVED

A. Rationale

Certain aspects of the data from experiment 1 and the broadband STM detection experiment of Bernstein *et al.*

(2013a)—i.e., reduced HI performance for STM stimuli containing low frequencies and low (e.g., 4 Hz) but not high (32 Hz) temporal modulation rates—were consistent with the idea that STM sensitivity is impaired by hearing loss due to a reduction in the ability to use TFS information. However, there is an alternative possible explanation for the pattern of results observed in these experiments. In both studies, the STM stimulus was generated by adding spectral sidebands with variable phase to each frequency component of the noise carrier. This stimulus generation process resulted in the introduction of a spectral-edge cue that listeners could have used to detect the presence of the modulation. Because the modulation was induced by adding sidebands to each carrier component, the lower and upper edges of the modulated band extended to slightly higher and lower frequencies for the modulated signal than for the unmodulated reference stimulus. This spectral-edge cue would have been most salient for higher temporal modulation rates (e.g., 32 Hz), where the sidebands would have extended to frequencies farther from the spectral edge of the noise carrier than for lower modulation rates. This cue would have also been more salient for stimuli with a low-frequency spectral edge (i.e., the broadband carrier and lower frequency octave-band carriers), because the change in frequency of the spectral edge induced by the added modulation would constitute a larger proportion of the edge frequency. It is possible that HI listeners might have had a deficit that affected STM detection equally across modulation rate, but that they were able to use the alternative spectral-edge cue to perform the task at higher temporal modulation rates just as well as NH listeners, especially at low frequencies where frequency selectivity was likely to be normal. If this were the case, then the HI listeners would show a more substantial deficit at lower temporal modulation rates (where the spectral-edge cue would be less salient) as was observed in experiment 1 and the study of [Bernstein et al. \(2013a\)](#)

Experiment 2 aimed to determine whether the pattern of STM-sensitivity deficits for the HI listeners observed in experiment 1 and in the broadband-carrier conditions of [Bernstein et al. \(2013a\)](#) would also be observed with this spectral-edge cue removed. If the trends observed previously (poorer performance for HI listeners at low temporal modulation rates and stimuli containing low carrier center frequencies) persisted, this would further support the idea that a reduced ability to use TFS information contributed to poorer STM sensitivity for low-frequency carriers for HI listeners. If not, this would instead suggest that the interaction between hearing loss and temporal modulation rate reflected the use of a spectral-edge cue by both listener groups for higher temporal modulation rates.

A second issue addressed by experiment 2 was that of age effects in STM detection. In experiment 1, all of the HI listeners were older than all of the NH listeners, such that the pattern of degradation in STM sensitivity for the HI group might have been attributable to their hearing loss or their age. As discussed in the introduction, previous studies have suggested that aging effects, rather than audiometric hearing loss *per se*, may be responsible for reduced TFS processing deficits (e.g., [Moore et al., 2011](#); [Sheft et al., 2012](#)) such as

those evidenced by the results of experiment 1. A subset of the NH listeners that were recruited to participate in experiment 2 had ages similar to those of the HI participants, allowing an examination of the effects of hearing loss on STM sensitivity, independent of age.

B. Methods

STM stimuli were generated as in experiment 1, with several changes implemented to reduce the availability of spectral-edge cues. The first change was that a given modulation sideband was added to the carrier only if its frequency fell between low- and high-frequency edges of the modulated band. The second change was that additional half-octave unmodulated noise bands were added below the 500-Hz octave band and above the 4000-Hz octave band. These noise bands had a nominal level of 70 dB/octave (15 dB/octave lower than the modulated carrier noise band), which was the same spectrum level as the noise presented in three octaves of the four-octave stimulus bandwidth that were left unmodulated in experiment 1. The purpose of these noise bands was to mask any spectral components that extended above or below the cutoff frequencies of the modulated band. Note that in experiment 1, unmodulated noise was present above and below the carrier band in the 1000-Hz and 2000-Hz octave-band conditions, but was missing above the 4000-Hz and below the 500-Hz octave-band carriers. In the broadband (four-octave) condition of [Bernstein et al. \(2013a\)](#) there was no unmodulated noise above or below the four-octave carrier bandwidth.

The removal of some of the spectral sidebands reduced the overall saliency of the STM, resulting in performance near floor levels (i.e., threshold modulation depth near 0 dB) for some conditions for some listeners. Two changes were made to increase performance levels, thereby reducing floor effects. First, a two-down, one-up adaptive algorithm was employed, tracking the 70.7%-correct performance level, instead of the 79.4% level tracked by the three-down, one-up algorithm in experiment 1 ([Levitt, 1971](#)). Because this resulted in fewer trials per reversal in the adaptive track, two additional reversals were added to each run, for a total of 11 reversals, with the threshold modulation depth calculated to be the mean of the last eight reversal points. Second, an initial reference interval containing unmodulated noise was added. This resulted in a three-interval, two-alternative, forced-choice task, with the target STM signal always presented in the second or third interval.

The question addressed in this experiment was whether the interaction between hearing loss and carrier center frequency observed in experiment 1 and the interaction between temporal modulation rate and hearing loss observed in experiment 1 and in the study of [Bernstein et al. \(2013a\)](#) would persist in the absence of a spectral edge cue. To examine this question, a subset of the conditions from experiment 1 and [Bernstein et al. \(2013a\)](#) were tested. The broadband condition from [Bernstein et al. \(2013a\)](#) was examined along with two octave-band carrier conditions: one low-frequency (1000 Hz) and one high-frequency (4000 Hz) condition. Two temporal modulation rates were tested (4 and 32 Hz),

representing the highest and lowest temporal modulation rates from the previous experiment. In the broadband and 4000-Hz octave-band conditions, only the spectral ripple density of 2 c/o was examined. In the 1000-Hz octave-band conditions, the slightly lower spectral ripple density of 1.5 c/o was tested in addition to 2 c/o because performance at 2 c/o approached floor levels for several listeners.

Twelve HI (three female) and 14 NH listeners (ten female) participated in this experiment. Twenty-three of these listeners were tested at Walter Reed National Military Medical Center, Bethesda, MD and three listeners were tested at the National Center for Rehabilitative Auditory Research, Portland, OR. One NH and one HI listener had also participated in experiment 1. The HI listeners had mean age of 64.3 years (range 53–71). The NH listeners had a mean age of 54.2 years (range 33–68). The data were also analyzed for a subset of eleven NH listeners above 50 years (mean age 60.5 years, range 53–68), selected to more closely match the ages of the HI listeners.

Mean audiograms for the ear tested ($\pm$ one standard error) are shown in Fig. 6. As in experiment 1, the HI listeners had, on average, high-frequency hearing loss, and mild hearing loss at 1000 Hz or below. The goal of testing NH listeners in the range of 50–70 years old made it difficult to recruit subjects with strictly normal hearing (audiometric thresholds better than or equal to 20 dB HL, bilaterally). All NH subjects had thresholds in the tested ear better than or equal to 20 dB HL at 250, 500, 1000, and 2000 Hz; 25 dB HL at 3000, 4000, and 6000 Hz; and 30 dB HL at 8000 Hz. Listeners were tested in their better ear.

C. Results

Group-mean STM detection thresholds are plotted as a function of temporal modulation rate in Fig. 7. Results for each carrier center frequency condition (and the two spectral ripple densities tested for the 1000-Hz carrier) are shown in separate panels. Data for the full listener groups are plotted as open (NH) and filled (HI) squares; data for the subset of 11 age-matched NH listeners are represented by open

FIG. 6. Mean audiograms ($\pm$ one standard error) for the 12 HI and 14 NH listeners that participated in experiment 2.

triangles. In comparing NH and HI performance for each condition, no Bonferroni correction was applied, because the comparisons between NH and HI listeners were planned (Keppel and Wilkins, 2004) to test the hypothesis that the same pattern of results would be obtained even in the absence of a spectral-edge cue. Asterisks (*) indicate conditions where performance was significantly different ($p < 0.05$) between the HI listeners (filled squares) and the full group of NH listeners (open squares) listeners. Carets (^) indicate conditions where performance was significantly different ($p < 0.05$) between the HI listeners (filled squares) and the age-matched subset of NH listeners (open triangles). The HI listeners showed poorer performance than the NH listeners, but only for the subset of conditions that involved a low temporal modulation rate (4 Hz) and carriers that contained low frequencies (1000 Hz and broadband). For the comparison between the HI listeners and the full group of NH listeners, one-tailed t-tests showed significantly poorer performance ($p < 0.05$) for the HI listeners for the following 4-Hz temporal modulation rate conditions: [1000 Hz, 2 c/o], [1000 Hz, 1.5 c/o] and [broadband, 2 c/o]. None of the 32-Hz temporal modulation rate or 4000-Hz carrier conditions showed significantly poorer performance for the HI listeners. The comparison between the HI listeners and the age-matched subgroup of NH listeners showed the same pattern of group differences, except that the difference for the 2-c/o, 4-Hz STM condition for a 1000-Hz carrier was not found to be significant in the age-matched analysis.

The large range of ages in the NH listener group allowed for an additional test of the relationship between age and STM sensitivity. Correlations were examined between age and STM detection thresholds for the NH listeners for the eight STM conditions tested in experiment 2. One-tailed tests showed significant correlations between age and STM sensitivity for the 1000-Hz carrier, 32-Hz temporal modulation-rate conditions for spectral ripple densities of 1.5 c/o ($R = 0.62$, $p < 0.01$) and 2 c/o ($R = 0.50$, $p < 0.05$). Age was not significantly correlated with performance in any other STM condition.

IV. DISCUSSION

The goals of this study were twofold: (1) to identify the combinations of carrier center frequency, spectral ripple density and temporal modulation rate where HI listeners have reduced STM sensitivity relative to NH listeners, and (2) to determine for which, if any, of these conditions estimates of STM sensitivity predict speech-reception performance for listeners with sensorineural hearing loss.

A. The effect of hearing loss on STM sensitivity

Reduced STM detection performance for the HI listeners was only observed for certain combinations of temporal modulation rate, spectral ripple density and carrier center frequency suggesting that reduced “processing efficiency”—i.e., cognitive factors such as memory and attention—are not the cause of these deficits (Patterson *et al.*, 1982; Hall and Grose, 1994). The patterns of reduced STM sensitivity for carrier center frequencies of 1000 and 4000 Hz are

FIG. 7. The results of experiment 2 showing STM sensitivity with spectral-edge cues removed. Group-mean STM detection thresholds are plotted as function of temporal modulation rate. Asterisks indicate conditions where performance was significantly different between HI and full NH listener groups. Carets (^) indicate conditions where performance was significantly different between the HI listeners and the age-matched subgroup of NH listeners.

discussed separately below. Specifically, we argue that the results suggest an influence of reduced TFS processing ability for the lower (1000-Hz) carrier center frequency, and an influence of reduced frequency selectivity for the higher (4000-Hz) carrier center frequency.

1. Impaired STM sensitivity at 1000 Hz: Reduced TFS processing ability

For the 1000-Hz carrier center frequency, reduced STM sensitivity was observed for HI listeners for a high spectral ripple density (2 c/o) and a low temporal modulation rate (4 or 12 Hz). This pattern of results is inconsistent with an explanation based on reduced temporal resolution, where a roll-off in performance at high modulation rates would be expected. The results of experiment 2 argue against the possibility that the observed pattern of results is attributable to the use of a spectral-edge to detect the STM at higher temporal modulation rates: the same pattern of reduced STM sensitivity for a 1000-Hz carrier center frequency persisted even with the spectral-edge cue removed. The fact that poorer STM sensitivity for a 1000-Hz carrier was observed for higher (2 c/o) but not for lower spectral ripple densities (0.5 and 1 c/o) does suggest a possible influence of reduced frequency selectivity. However, [Summers et al. \(2013\)](#) measured frequency selectivity for the same group of HI listeners from the current study using the notched-noise method and found their average auditory-filter bandwidths at 1000 Hz to be comparable to those for a group of NH listeners.

Instead, the observed pattern of reduced STM sensitivity for a 1000-Hz carrier center frequency—with performance negatively impacted by hearing loss for low but not for high temporal modulation rates—appears to be most consistent with an explanation based on the inability to use TFS information to track dynamic spectral information. [Moore and Sek \(1996\)](#) proposed that the detection of frequency modulation (FM) can be accomplished either by detecting AM cues, or by using phase-locking information that encodes changes in the instantaneous frequency of the carrier. They showed that for low carrier frequencies, when AM was added to both intervals of a FM-detection trial to disrupt induced-AM cues, performance for NH listeners worsened for high (but not for low) carrier frequencies, and for high (but not for

low) FM rates. This suggested that listeners were using TFS information to detect FM only for low carrier frequencies and low temporal modulation rates. Their interpretation was that at higher carrier frequencies, TFS information was not available due to roll-off in auditory-nerve phase locking to the cycle-by-cycle variation in the carrier frequency ([Johnson, 1980](#)). At higher temporal modulation rates, TFS information was not available due to the sluggish nature of the TFS encoding mechanism. [Moore and Skrodzka \(2002\)](#) showed similar results while investigating the effects of hearing loss on FM detection performance. For low carrier frequencies, hearing loss impacted FM detection performance more for low than high temporal modulation rates, in contrast to high carrier frequencies where the effect of hearing loss on FM detection was constant across temporal modulation rate. These results were therefore consistent with the idea that for low carrier frequencies, NH listeners were able to make use of TFS information to detect FM, but that this process was impaired for the HI listeners.

The STM stimuli employed in the current study were similar to the FM stimuli utilized by [Moore and Sek \(1996\)](#) and [Moore and Skrodzka \(2002\)](#) in that they involve spectral peaks with frequencies that change over time. If HI listeners had a reduced ability to use TFS information to track the changing frequencies of the spectral peaks in the STM stimuli, STM detection should also be disrupted at low carrier center frequencies and low temporal modulation rates. There was some evidence for this pattern of results in the current study, with a significant performance difference observed between the NH and HI listeners for a 2-c/o, 4-Hz STM stimulus at a carrier center frequency of 1000 Hz. The fact that such an effect was not also observed for a 500 Hz carrier might reflect the influence of floor effects on the data. Nevertheless, the same trend was apparent for the 2-c/o spectral ripple density at 500 Hz, although the difference in performance between the two groups did not reach significance.

There is some evidence for TFS-processing deficits for HI listeners in previous psychophysical ([Moore and Skrodzka, 2002](#); [Buss et al., 2004](#); [Lacher-Fougère and Demany, 2005](#)) and speech-perception studies ([Lorenzi et al., 2006](#); [Hopkins and Moore, 2007](#)). There is little physiological evidence for reduced fidelity of phase-locking information in individual auditory-nerve fibers (ANFs) (e.g., [Kale](#)

and Heinz, 2010) in response to tonal stimuli presented in quiet. Recent findings have suggested that the strength of phase locking could be reduced by hearing loss for stimuli presented in noise (Henry and Heinz, 2012). However, this effect is related to an increase in the influence of the noise on the ANF response due to a broadening of the ANF tuning curve. Alternatively, the ability to use TFS information might require redundant coding available across many ANFs. Recent results have suggested that noise exposure can destroy the function of up to 50% of ANFs without substantially impacting audiometric thresholds (Kujawa and Liberman, 2009), which could greatly reduce the redundancy of TFS coding in the peripheral auditory system.

Although the pattern of impaired STM sensitivity for HI listeners across temporal modulation rates was consistent with an explanation based on reduced TFS processing ability, it is not clear why a TFS-processing deficit would affect STM detection performance for a higher spectral ripple density of 2 c/o but not lower spectral ripple densities of 0.5 or 1 c/o, as was also observed here for the 1000-Hz carrier center frequency. One possibility is that for lower spectral ripple densities, the fact that the spectral peaks in the STM stimulus were relatively broad might have reduced the usefulness of TFS information in the identification of, and detection of changes in, the instantaneous frequencies of the spectral peaks. If neither the NH nor the HI listeners were able to use TFS information for STM detection in these conditions, then performance for the HI listeners would not be impacted by poor TFS processing ability.

2. Impaired STM sensitivity at 4000 Hz: Reduced frequency selectivity

HI listeners performed significantly more poorly than the NH listeners for a high carrier center frequency of 4000 Hz, a high spectral ripple density of 4 c/o, and a low temporal modulation rate of 4 Hz. This general pattern of reduced STM sensitivity for high carrier center frequencies and high spectral ripple densities is consistent with the idea that STM detection performance was adversely affected by reduced frequency selectivity in this region.

The well-documented loss of frequency selectivity as a result of hearing loss (e.g., Glasberg and Moore, 1986; Leek and Summers, 1993) is believed to reflect the broadening of the narrowly tuned cochlear filters. HI listeners mainly show reduced frequency selectivity in frequency regions where their audiometric thresholds are elevated relative to NH listeners (Moore, 2007), which in this study and many others involves higher frequency regions. Furthermore, reduced frequency selectivity is most likely to affect the highest spectral ripple densities where the close peak spacing begins to interact with the bandwidth of the auditory filter. At 4000 Hz, the equivalent rectangular bandwidth (ERB) of the NH auditory filter is 456 Hz, or about 1/6 of an octave (Glasberg and Moore, 1990), at least for low-level sounds. There are considerable discrepancies in the literature regarding the degree to which frequency selectivity is negatively impacted by hearing loss. Earlier studies suggested that it can be reduced by a factor of 2–4 for mild-to-moderate impairment or a

factor of 5 for moderate-to-severe impairment (e.g., Glasberg and Moore, 1986). Others have argued that a substantial portion of the auditory-filter broadening relative to NH listeners is attributable to the higher sound levels at which HI listeners are tested to overcome audibility limitations (e.g., Houtgast and Festen, 2008). A conservative estimate of this effect—a factor of 2 increase in the bandwidth of the auditory filter accompanying hearing loss (Summers *et al.*, 2013)—would yield a $\frac{1}{3}$ octave ERB. In this case, reduced frequency selectivity would not affect STM sensitivity performance until the spectral ripple density exceeds roughly 3 c/o. The fact that reduced STM sensitivity for HI listeners for a 4000-Hz carrier center frequency was observed for a high spectral ripple density of 4 c/o but not for lower spectral ripple densities is consistent with the hypothesis that STM detection performance is negatively impacted by reduced frequency selectivity.

One caveat to this interpretation is that the impact of hearing loss at 4 c/o was observed more for lower (4 Hz) than for higher (12–32 Hz) temporal modulation rates. This is inconsistent with an explanation based on frequency selectivity, in that we would expect to see an impact at all temporal modulation rates. While the observed pattern would tend to favor an explanation based on TFS processing ability, as was argued for a 1000-Hz carrier center frequency, such an explanation seems unlikely for the 4000-Hz carrier center frequency given the roll-off in phase locking at higher absolute frequencies (Johnson, 1980).

B. The relationship between STM sensitivity and speech-reception performance for HI listeners

STM detection performance in two of the conditions where the HI listeners showed poorer STM sensitivity than the NH listeners were found to be significantly correlated with the SRT_{50} after partialling out the audiogram-based SII: the 2-c/o, 4-Hz STM condition for the 1000-Hz carrier center frequency and the 4-c/o, 4-Hz STM condition for the 4000-Hz carrier center frequency. The SII-based SRT_{50} predictions captured about 60% of the variance in speech-reception performance in stationary noise. When the SII was combined with sensitivity measures for these two STM conditions, approximately 90% of the variance in SRT_{50} was accounted for. To the extent that poorer STM sensitivity for HI listeners reflects a reduced ability to use TFS information (for the low-frequency 1000-Hz carrier) and reduced frequency selectivity (for the higher-frequency 4000 Hz carrier), these results suggest that deficits in TFS processing and frequency selectivity also contribute to the poorer speech-reception performance in noise exhibited by HI subjects.

The predictive power of STM-detection performance for a low carrier center frequency is consistent with previous results that have suggested that an inability to use TFS information contributes to the poor speech-reception performance in noise exhibited by HI listeners (Buss *et al.*, 2004; Lorenzi *et al.*, 2006; Hopkins and Moore, 2007; Strelcyk and Dau, 2009). Furthermore, this result suggests that a suprathreshold psychoacoustic measure in a less-impaired frequency region

can predict speech-reception deficits in HI listeners, consistent with previous studies that have shown impaired speech perception in HI listeners in frequency regions where these listeners have normal or near-normal audiometric thresholds (e.g., [Lorenzi et al., 2009](#); [Léger et al., 2012](#)).

The predictive power of STM-sensitivity performance for a high-frequency carrier at a high spectral ripple density, suggesting an influence of frequency selectivity on speech-reception performance, was somewhat surprising. [Summers et al. \(2013\)](#) examined the same group of HI listeners, and found a significant correlation between speech and FM sensitivity but not between speech and frequency selectivity measured using the notched-noise method. This result corroborates with other studies that have identified a relationship between TFS processing ability and speech intelligibility but not between frequency selectivity and speech intelligibility (e.g., [Buss et al., 2004](#)). It might be that the role of frequency selectivity suggested by the current study result from the nature of the test stimulus employed. STM stimuli were designed to represent aspects of the speech signal, whereas the notched-noise stimuli employed often employed in measurements of frequency selectivity were designed to more carefully measure the physical characteristics of cochlear processing. The dynamic nature of the STM stimuli employed here may also more closely mimic the dynamic speech signal, resulting in a stronger relationship to speech-reception performance.

C. Age effects

As with many studies of hearing loss, the makeup of the NH and HI listener groups in experiment 1 made it difficult to disassociate the effects of hearing loss and age on psychophysical performance. An analysis of the data from experiment 2 conducted for a subset of the NH listeners that were age-matched to the HI listeners showed the same pattern of results that were observed in experiment 1 (for a 1000-Hz carrier) and in [Bernstein et al. \(2013a\)](#) (for a broadband carrier). For a spectral ripple density of 2 c/o, performance differences between the NH and HI listener groups were observed for a low temporal modulation rate but not for a high temporal modulation rate. This suggests that this deficit, argued above to reflect a deficit in the ability to use TFS information, reflects hearing loss and not age. This is consistent with previous findings that have suggested that hearing impairment adversely affects the ability to use TFS cues regardless of age ([Lorenzi et al., 2006, 2009](#)). Because the 4 c/o spectral ripple density for a 4000-Hz carrier was not examined for the age-matched listeners groups in experiment 2, we cannot say whether the difference observed between NH and HI listeners in experiment 1—which we argue above may be related to reduced frequency selectivity—is attributable to age or hearing loss. However, previous studies have shown little evidence that frequency selectivity is negatively impacted by age independently of hearing loss (e.g., [Sommers and Humes, 1993](#); [Hopkins and Moore, 2011](#)).

There was one piece of evidence from experiment 2 suggesting a relationship between age and STM sensitivity: significant correlations were observed between age and STM

sensitivity in the NH listener group for a 1000-Hz carrier, 32-Hz temporal modulation rate, and 1.5- or 2-c/o spectral ripple densities. The fact that these correlations were observed for the higher tested temporal modulation rate of 32 Hz suggests an influence of age on temporal modulation processing, consistent with previous findings (e.g., [He et al., 2008](#)). Significant correlations between age and STM sensitivity were not observed for any conditions where the HI listeners showed poorer STM sensitivity than the NH listeners, nor for any low temporal modulation rate conditions for which TFS cues might play a role. This provides further evidence that the observed pattern of STM sensitivity differences between the NH and HI listeners groups were attributable to hearing loss rather than age.

D. Implications for models of speech intelligibility

There is a rich history of attempts to model speech intelligibility for HI listeners. The general approach has been to adjust the features of a model of speech intelligibility for NH listeners to simulate perceptual degradation due to hearing loss. The most common approach is to base such degradations on the audiogram for a particular HI listener. For example, the Speech Intelligibility Index ([ANSI, 1997](#)) includes a set of parameters to account for the reduced audibility for HI listeners due to elevated audiometric thresholds. Recognizing that reduced audibility only partially accounts for the poorer speech-reception performance associated with hearing-loss, other models have incorporated generic suprathreshold distortions in their predictions of speech intelligibility for individual HI listeners (e.g., [Plomp, 1986](#); [Ching et al., 2001](#)). Although such attempts represent an improvement over earlier models that only simulate reduced audibility, they are nevertheless limited in their ability to successfully predict speech intelligibility for individual HI listeners because the degree of suprathreshold distortion is tied to audiometric measures.

The results presented here, as well as previous results (e.g., [Buss et al., 2004](#); [Bernstein et al., 2013a](#); [Summers et al., 2013](#)), suggest that to successfully account for individual differences among HI listeners, computational models of speech intelligibility must incorporate suprathreshold deficits that are not directly related to the audiogram (e.g., [Jepsen and Dau, 2011](#)). The data have mapped out the particular set of STM stimuli for which HI listeners demonstrate reduced detection ability, and outlined the subset of these conditions for which STM sensitivity is predictive of speech intelligibility. These results therefore lend themselves to the STM-based speech-intelligibility modeling approach of [Elhilali et al. \(2003\)](#) and [Chi et al. \(1999\)](#). Speech predictions generated by this model are based directly on the internal representations of STM at the output of the auditory periphery. The current results suggest that to accurately account for individual differences in speech intelligibility, an STM-based model should model deficits in both TFS processing at low frequencies and frequency selectivity at high frequencies.

[Bernstein et al. \(2013b\)](#) and [Grant et al. \(2013\)](#) recently showed that incorporating peripheral processing deficits to

model individual differences in STM sensitivity within the speech-intelligibility modeling framework of Elhilali *et al.* (2003) can predict speech intelligibility for individual HI listeners more accurately than a method that only incorporates differences in the audiogram. Bernstein *et al.* (2013b) modeled reduced STM sensitivity by adjusting the strength of a lateral inhibition network (LIN) posited at the output of the auditory periphery. Although this manipulation increased the model's ability to account for variance in speech intelligibility across individual HI listeners, they noted that adjustments to the LIN are unlikely to account for the nuances of the pattern of reduced STM sensitivity for HI listeners observed here and in the study of Bernstein *et al.* (2013a), whereby performance was impacted by hearing loss mainly for lower temporal modulation rates. To address this shortcoming, Grant *et al.* (2013) proposed a TFS-based autocorrelation mechanism to extract spectral information from the signal, consistent with the TFS-based explanation suggested by the pattern of results for HI listeners in the current study for the 1000-Hz carrier center frequency. By incorporating a temporal-integration window for the TFS-based extraction of spectral information, this approach was able to capture the temporal modulation-rate dependence of the impact of hearing loss on STM sensitivity, while also improving the model's ability to account for individual variability in speech-reception performance in noise.

The result of the current study suggests that the model's ability to account for individual speech-reception scores might be further improved by incorporating individual differences in frequency selectivity in the 4000-Hz range in addition to modeling TFS deficits in the 1000-Hz range. Broadening the filter bandwidths in the model would tend to produce a poorer representation of STM at higher spectral ripple densities, as was observed in the 4000-Hz data in the current study. However, Bernstein *et al.* (2013b) found that incorporating into the speech-intelligibility model the individualized auditory filter bandwidths as estimated using the notched-noise method did not improve the model's predictions, most likely because of the lack of a correlation between these estimates of frequency selectivity and speech intelligibility.

V. CONCLUSIONS

HI listeners showed impaired sensitivity to STM for both a low (1000 Hz) and a high carrier center frequency (4000 Hz), but only for a few combinations of spectral ripple density and temporal modulation rate. The combinations of temporal modulation rate and spectral ripple density for which these deficits occurred in these two frequency regions suggest that STM sensitivity is negatively impacted by hearing loss through a combination of reduced frequency selectivity (at 4000 Hz) and an impaired ability to use TFS information (at 1000 Hz). STM sensitivity to 4-Hz, 2-c/o modulation applied to a 1000-Hz carrier center frequency and to 4-Hz, 4-c/o modulation applied to a 4000-Hz carrier center frequency accounted for significant portions of the variance in speech-reception performance that were not accounted for by the audiogram-based SII. Together, performance in these two conditions accounted for about 30%

of the variance in speech intelligibility in stationary noise for HI listeners beyond the 60% accounted by the SII-based SRT₅₀ predictions (for a total of 90%). The results are consistent with the idea that impairment in aspects of STM detection based on TFS processing (for low carrier center frequencies) and frequency selectivity (for high carrier center frequencies) are detrimental to speech perception in noise for HI listeners.

ACKNOWLEDGMENTS

This work was supported by a grant from the Oticon Foundation, Smørum, Denmark (J.G.W.B. and M.R.L.). Additional support was provided by the VA Rehabilitation Research & Development Service [Career Development Grant (F.J.G.), Senior Research Career Scientist (M.R.L.)] and a grant from the National Institute of Health [NIH NIDCD R01 DC011828 (F.J.G.)]. Some resources and facilities were provided by the Portland VA Medical Center and the Department of Research Programs (DRP) and the Audiology and Speech Center (ASC), Walter Reed National Military Medical Center. The authors thank Ken Grant, Matthew Makashay, Sandeep Phatak, Shihab Shamma, Chris Stecker, Van Summers, Sarah Theodoroff, and Nirmal Srinivasan for their contributions to this project. Experiment 1 formed part of the thesis work of G.M. (Department of Electrical and Computer Engineering, University of Maryland-College Park) conducted at Walter Reed Army Medical Center. The views expressed in this article are those of the authors and do not reflect the official policy of the Department of the Army, the Department of the Navy, the Department of the Air Force, the Department of Defense, the Department of Veterans Affairs, or the U.S. Government.

APPENDIX: THE RELIABILITY OF AN ALTERNATIVE SPECTRAL-PROFILE CUE

1. Rationale

The construction of the STM stimuli in experiments 1 and 2 was carried out by adding spectral sideband to each tone component in the octave-band noise carrier. As a result, the difference between the spectrum level of the stimulus within the modulated octave band and that of unmodulated noise bands outside was greater for the modulated target interval than for the unmodulated reference interval. To determine whether this alternative cue could have influenced the results of experiments 1 and 2, a supplementary experiment measured the threshold spectral-level difference listeners needed to perform the task in the absence of any modulation applied to the signals.

2. Methods

The stimulus construction and threshold measurement methods were similar to experiment 2, except that no modulation sidebands were added to the target signal. Instead, the levels of the carrier components within the carrier octave band in question were raised to produce a spectrum-level

difference, relative to the other unmodulated carrier bands, that was equivalent to that produced by the addition of modulation sidebands in experiment 2 for a particular modulation depth. Specifically, the amount of gain (G , in dB) added to the carrier band for a given equivalent modulation depth was

$$G = 10 \log_{10}[1 + 2(m/2)^2], \quad (\text{A1})$$

where m is the modulation depth. For example, the amount of gain applied to produce a spectral difference equivalent to that of a fully modulated signal ($m = 1$) was 1.76 dB. After the gain was applied, the level of the stimulus was adjusted to yield a nominal overall stimulus level of 80 dB SPL before the application of the ± 2.5 dB level rove. As in experiment 2, each trial consisted of three intervals. Two of the intervals contained the reference signal while the other interval contained the modulation-equivalent gain in the carrier band in question. The target was always presented in the second or third interval of the trial. The listener was directed to identify the interval that had a different quality than the others. Feedback was given following each response by briefly presenting an asterisk above the button that was associated with the correct answer.

As in experiment 2, a two-down, one-up adaptive tracking procedure was used to determine the threshold equivalent modulation depth. The only significant departure from the threshold-measurement procedure employed in experiment 2 was that there was no mathematical limit to the amount of gain that could be applied to the “modulated” band. Thus, in this follow up experiment the spectral difference between the modulated band and the “unmodulated” sidebands was allowed to exceed 1.76 dB—that is, the amount of spectral difference in experiment 1 that was associated with full-depth modulation. As a result the “equivalent modulation depth” was allowed to exceed a value of 0 dB (i.e., $m = 1$). The idea was that if the threshold equivalent modulation depth was found to be greater than 0 dB, then it would not have been possible for listeners to detect the presence of the modulation based on the spectral-profile cue alone. If the threshold modulation depth was found to be less than 0 dB, this would indicate that for some stimulus with a physically plausible modulation depth, listeners could have completed the modulation-detection task in experiments 1 and 2 based on the spectral-profile cue rather than by actually detecting the presence of modulation.

In each adaptive run, the initial equivalent modulation depth ($20 \log_{10} m$) was set to 12 dB (i.e., $m = 4$), equivalent to the spectrum-level cue that would be generated if each modulation sideband had twice the amplitude of each carrier component. The equivalent modulation depth then changed following each sequential pair of correct responses or single incorrect response by 6 dB for the first three reversals, by 4 dB for the next two reversals, and by 2 dB for the last eight reversals. The detection threshold was calculated to be the mean of the last eight reversal points.

The threshold equivalent modulation depth was estimated for octave-band carriers centered at 1000 and 4000 Hz and for the four-octave carrier. Five runs were completed by each subject for each carrier center frequency condition.

Conditions were presented in pseudo-random order, with one run presented for each condition before a second run was presented for any given condition. The first two runs for each condition were considered to be training runs and discarded, and the threshold for each subject was calculated as the mean of the last three measurements for each condition. Eight NH listeners participated (three female, mean age 39 years, range 26–55 years). Two of these listeners had also participated in experiment 2.

3. Results

The mean equivalent modulation-depth detection thresholds for each carrier condition ($\pm$ one standard error) were 1.5 ± 1.1 dB for the 1000-Hz octave-band carrier, 5.7 ± 2.3 dB for the 4000-Hz octave-band carrier, and 16.2 ± 2.1 dB for the four-octave carrier.

4. Discussion

The maximum modulation depth allowed in experiments 1 and 2 was 0 dB (full modulation). This supplementary experiment showed that for all of the carrier center frequency conditions tested in experiment 2, the mean equivalent modulation depth required to identify the target interval based on spectrum-level information exceeded this limit. This result suggests that the group-mean thresholds measured in experiment 2 did not reflect the use of this alternative cue. Furthermore, the threshold estimation methodology employed in experiment 1 would have made it even more difficult for listeners to make use of this alternative spectrum-level cue in that experiment. First, the use of a two-interval task as in experiment 1, rather than a three-interval “odd-one-out” paradigm employed here, would have increased the task difficulty. Second, the use of an adaptive algorithm that tracked the 79.4% correct performance level, as in experiment 1, rather than the 70.7% correct performance level used here would have resulted in higher threshold. Thus, it is even more unlikely than for experiment 2 that the presence of an alternative spectrum-level cue influenced the results of experiment 1.

- ANSI (1997). S3.5, *Methods for Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).
- Baer, T., and Moore, B. C. J. (1993). “Effects of spectral smearing on the intelligibility of sentences in the presence of noise,” *J. Acoust. Soc. Am.* **94**, 1229–1241.
- Bernstein, J. G. W., Mehraei, G., Shamma, S., Gallun, F., Theodoroff, S. M., and Leek, M. R. (2013a). “Spectrotemporal modulation sensitivity as a predictor of speech intelligibility for hearing-impaired listeners,” *J. Am. Acad. Audiol.* **24**, 293–306.
- Bernstein, J. G. W., Summers, V., Grassi, E., and Grant, K. W. (2013b). “Auditory models of suprathreshold distortion and speech intelligibility in persons with impaired hearing,” *J. Am. Acad. Audiol.* **24**, 307–328.
- Buss, E., Hall, J. W., and Grose, J. H. (2004). “Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss,” *Ear Hear.* **25**, 242–250.
- Chi, T., Gao, Y., Guyton, M. C., Ru, P., and Shamma, S. (1999). “Spectrotemporal modulation transfer functions and speech intelligibility,” *J. Acoust. Soc. Am.* **106**, 2719–2732.
- Ching, T. Y. C., Dillon, H., Katsch, R., and Byrne, D. (2001). “Maximizing effective audibility in hearing aid fitting,” *Ear Hear.* **22**, 212–224.

- Elhilali, M., Taishih, C., and Shamma, S. A. (2003). "A spectro-temporal modulation index (STMI) for assessment of speech intelligibility," *Speech Commun.* **41**, 331–348.
- 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.
- Glasberg, B. R., and Moore, B. C. J. (1986). "Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments," *J. Acoust. Soc. Am.* **79**, 1020–1033.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**, 103–138.
- Glasberg, B. R., and Moore, B. C. J. (2000). "Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise," *J. Acoust. Soc. Am.* **108**, 2318–2328.
- Gnansia, D., Pean, V., Meyer, B., and Lorenzi, C. (2009). "Effects of spectral smearing and temporal fine structure degradation on speech masking release," *J. Acoust. Soc. Am.* **125**, 4023–4033.
- Grant, K. W., Bernstein, J. G. W., and Summers, V. (2013). "Predicting speech intelligibility by individual hearing-impaired listeners: The path forward," *J. Am. Acad. Audiol.* **24**, 329–336.
- Grose, J. H., and Mamo, S. K. (2010). "Processing of temporal fine structure as a function of age," *Ear Hear.* **31**, 755–760.
- Hall, J. W., and Grose, J. H. (1994). "Development of temporal resolution in children as measured by the temporal modulation transfer function," *J. Acoust. Soc. Am.* **96**, 150–154.
- He, N., Mills, J. H., Ahlstrom, J. B., and Dubno, J. R. (2008). "Age-related differences in the temporal modulation transfer function with pure-tone carriers," *J. Acoust. Soc. Am.* **124**, 3841–3849.
- Henry, K. S., and Heinz, M. G. (2012). "Diminished temporal coding with sensorineural hearing loss emerges in background noise," *Nat. Neurosci.* **15**, 1362–1364.
- Hopkins, K., and Moore, B. C. J. (2007). "Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information," *J. Acoust. Soc. Am.* **122**, 1055–1068.
- Hopkins, K., and Moore, B. C. J. (2011). "The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise," *J. Acoust. Soc. Am.* **130**, 334–349.
- Houtgast, T., and Festen, J. M. (2008). "On the auditory and cognitive functions that may explain an individual's elevation of the speech reception threshold in noise," *Int. J. Audiol.* **47**, 287–295.
- IEEE (1969). *IEEE Recommended Practice for Speech Quality Measures* (Institute of Electrical and Electronic Engineers, New York).
- Jepsen, M. L., and Dau, T. (2011). "Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss," *J. Acoust. Soc. Am.* **129**, 262–281.
- Johnson, D. H. (1980). "The relationship between spike rate and synchrony in the responses of auditory-nerve fibers to single tones," *J. Acoust. Soc. Am.* **68**, 1115–1122.
- Kale, S., and Heinz, M. G. (2010). "Envelope coding in auditory nerve fibers following noise-induced hearing loss," *J. Assoc. Res. Otolaryngol.* **11**, 657–673.
- Keppel, G., and Wickens, T. D. (2004). *Design and Analysis: A Researcher's Handbook*, 4th ed. (Prentice Hall, Englewood Cliffs, NJ), pp. 111–130.
- Kujawa, S. G., and Liberman, M. C. (2009). "Adding insult to injury: Cochlear nerve degeneration after 'temporary' noise-induced hearing loss," *J. Neurosci.* **29**, 14077–14085.
- Lacher-Fougère, S., and Demany, L. (2005). "Consequences of cochlear damage for the detection of interaural phase differences," *J. Acoust. Soc. Am.* **118**, 2519–2526.
- Leek, M. R., and Summers, V. (1993). "The effect of temporal waveform shape on spectral discrimination by normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **94**, 2074–2082.
- Léger, A. C., Moore, B. C. J., and Lorenzi, C. (2012). "Temporal and spectral masking release in low- and mid-frequency regions for normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **131**, 1502–1514.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Lorenzi, C., Debrulle, L., Garnier, S., Fleuriet, P., and Moore, B. C. J. (2009). "Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal," *J. Acoust. Soc. Am.* **125**, 27–30.
- Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (2006). "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," *Proc. Natl. Acad. Sci.* **103**, 18866–18869.
- Moore, B. C. J. (2007). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues* (Wiley, Chichester, England), pp. 82–83.
- Moore, B. C. J., Glasberg, B. R., Stoev, M., Füllgrabe, C., and Hopkins, K. (2011). "The influence of age and high-frequency hearing loss on sensitivity to temporal fine structure at low frequencies," *J. Acoust. Soc. Am.* **131**, 1003–1006.
- Moore, B. C. J., and Sek, A. (1996). "Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking," *J. Acoust. Soc. Am.* **100**, 2320–2331.
- Moore, B. C. J., Shailer, M. J., and Schooneveldt, G. P. (1992). "Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss," *Br. J. Audiol.* **26**, 229–237.
- Moore, B. C. J., and Skrodzka, E. (2002). "Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation," *J. Acoust. Soc. Am.* **111**, 327–335.
- Neher, T., Lunner, T., Hopkins, K., and Moore, B. C. J. (2012). "Binaural temporal fine structure sensitivity, cognitive function, and spatial speech recognition of hearing-impaired listeners," *J. Acoust. Soc. Am.* **131**, 2561–2564.
- Nelson, D. A., Schroder, A. C., and Wojtczak, M. (2001). "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners," *J. Acoust. Soc. Am.* **110**, 2045–2064.
- Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982). "The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold," *J. Acoust. Soc. Am.* **72**, 1788–1803.
- Plomp, R. (1986). "A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired," *J. Speech Hear. Res.* **29**, 146–154.
- Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (1997). "Basilar-membrane responses to tones at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.* **101**, 2151–2163.
- Sheft, S., Shafiro, V., Lorenzi, C., McMullen, R., and Farrell, C. (2012). "Effects of age and hearing loss on the relationship between discrimination of stochastic frequency modulation and speech perception," *Ear Hear.* **33**, 709–720.
- Smoorenburg, G. F. (1992). "Speech reception in quiet and in noisy conditions by individuals with noise-induced hearing loss in relation to their tone audiogram," *J. Acoust. Soc. Am.* **91**, 421–437.
- Sommers, M. S., and Humes, L. E. (1993). "Auditory filter shapes in normal-hearing, noise-masked normal, and elderly listeners," *J. Acoust. Soc. Am.* **93**, 2903–2914.
- Strelcyk, O., and Dau, T. (2009). "Relations between frequency selectivity, temporal fine-structure processing and speech reception in impaired hearing," *J. Acoust. Soc. Am.* **125**, 3328–3345.
- Summers, V., Makashay, M. J., Theodoroff, S., and Leek, M. R. (2013). "Suprathreshold auditory processing and speech perception in noise: Hearing-impaired and normal-hearing listeners," *J. Am. Acad. Audiol.* **24**, 274–292.