Behavioral measures of cochlear compression and temporal resolution as predictors of speech masking release in hearing-impaired listeners

Melanie J. Gregan[a] and Peggy B. Nelson
Department of Speech-Language-Hearing Science, University of Minnesota, 164 Pillsbury Drive SE, Minneapolis, Minnesota 55455

Andrew J. Oxenham
Department of Psychology, University of Minnesota, 75 East River Parkway, Minneapolis, Minnesota 55455

(Received 30 June 2012; revised 31 July 2013; accepted 5 August 2013)

Hearing-impaired (HI) listeners often show less masking release (MR) than normal-hearing listeners when temporal fluctuations are imposed on a steady-state masker, even when accounting for overall audibility differences. This difference may be related to a loss of cochlear compression in HI listeners. Behavioral estimates of compression, using temporal masking curves (TMCs), were compared with MR for band-limited (500–4000 Hz) speech and pure tones in HI listeners and age-matched, noise-masked normal-hearing (NMNH) listeners. Compression and pure-tone MR estimates were made at 500, 1500, and 4000 Hz. The amount of MR was defined as the difference in performance between steady-state and 10-Hz square-wave-gated speech-shaped noise. In addition, temporal resolution was estimated from the slope of the off-frequency TMC. No significant relationship was found between estimated cochlear compression and MR for either speech or pure tones. NMNH listeners had significantly steeper off-frequency temporal masking recovery slopes than did HI listeners, and a small but significant correlation was observed between poorer temporal resolution and reduced MR for speech. The results suggest either that the effects of hearing impairment on MR are not determined primarily by changes in peripheral compression, or that the TMC does not provide a sufficiently reliable measure of cochlear compression.
 [http://dx.doi.org/10.1121/1.4818773]

PACS number(s): 43.66.Dc, 43.71.Ky [ELP] Pages: 2895–2912

I. INTRODUCTION

Listeners with hearing impairments often have difficulty understanding speech in a noisy background. Compared with normal-hearing (NH) listeners, this difficulty can be exacerbated when the background noise contains temporal and/or spectral fluctuations. A better understanding of the causes underlying these difficulties in complex acoustic backgrounds may help in the search for better diagnostic tools and improved hearing-aid algorithms. Generally, NH listeners experience an improvement in speech understanding when the background noise is temporally modulated, compared to their performance when the noise is unmodulated. This improvement, termed "masking release" (MR), is thought to reflect the ability of listeners to take advantage of the improved signal-to-noise ratio (SNR) in the temporal dips of the fluctuating masker. Listeners with cochlear hearing loss generally show less or even no MR under similar listening conditions, leading to a greater difference between results from NH and hearing-impaired (HI) listeners in fluctuating than in steady-state maskers (Miller and Licklider, 1950; Duquesnoy, 1983; Festen and Plomp, 1990; Takahashi and Bacon, 1992; Eisenberg *et al.*, 1995; Bacon *et al.*, 1998; Peters *et al.*, 1998; Kwon and Turner, 2001; Nelson *et al.*, 2003; Nelson and Jin, 2004; Summers and Molis, 2004; George *et al.*, 2006; Jin and Nelson, 2006; Lorenzi *et al.*, 2006; Wagener *et al.*, 2006; Buss *et al.*, 2009; Desloge *et al.*, 2010). Several studies have evaluated possible explanations for this lack of MR, including audibility, temporal resolution, spectral resolution, and cochlear compression.

One potential reason for reduced MR in hearing-impaired (HI) listeners is the reduced audibility of the speech within the temporal dips in a masker as a result of the hearing loss. Several studies have attempted to control for the reduced audibility in HI listeners, either by comparing them to NH listeners with thresholds elevated by noise (e.g., Eisenberg *et al.*, 1995; Bacon *et al.*, 1998; Desloge *et al.*, 2010), or by amplifying the stimuli presented to the HI listeners according to a prescriptive formula to improve audibility (e.g., Peters *et al.*, 1998; Moore *et al.*, 1999b). It appears from these studies that for some HI listeners audibility does indeed explain the lack of MR (e.g., Bacon *et al.*, 1998; Desloge *et al.*, 2010). However, this is not the case for all HI listeners (e.g., Eisenberg *et al.*, 1995; Bacon *et al.*, 1998). Therefore, factors beyond audibility may reduce MR, at least for a subset of HI listeners.

A second reason for reduced MR in HI listeners may be abnormal temporal resolution, and the "persistence of forward masking," whereby peaks in the masker render following lower-level speech segments inaudible (Festen and Plomp, 1990). George *et al.* (2006) found that their measure of temporal resolution (a temporal window derived from counting the number of tone sweeps in an interrupted noise)

[a]Author to whom correspondence should be addressed. Electronic mail: grega005@umn.edu

was correlated with degree of MR for their HI listeners. This conclusion was not fully supported by Jin and Nelson (2006), who did not find a strong relationship between their measure of temporal resolution (recovery from forward masking) and degree of MR for sentences, although they did find a relationship between forward-masked thresholds and MR for consonant-vowel pairs. One problem with using forward masking as a measure of temporal resolution is that the decay of forward masking may be affected by changes in cochlear compression as well as any underlying changes in temporal resolution (Oxenham and Moore, 1997).

A third potential factor in MR may be age, independent of hearing loss. Although Takahashi and Bacon (1992) did not find a major effect of age on MR, results from Dubno *et al.* (2002, 2003) did suggest that MR may diminish with increasing age, even when audibility is controlled for, and George *et al.* (2006) found a significant effect of age on their measures of temporal resolution.

A fourth potential reason for reduced MR relates to the functioning of the cochlea's outer hair cells. Cochlear compression and sharp frequency tuning are both believed to be mediated by the outer hair cells in the cochlea (e.g., Ruggero and Rich, 1991; Zheng *et al.*, 2000). Compression can lead to low-level speech presented in masker dips being amplified relative to higher-level masker peaks, thereby increasing the effective long-term signal-to-noise ratio (e.g., Oxenham and Dau, 2001, 2004; Rhebergen *et al.*, 2009). Therefore, a loss of compression may impair the ability of HI listeners to hear speech in the dips of a temporally modulated masker (e.g., Peters *et al.*, 1998), just as the introduction of compression via a hearing-aid algorithm may improve performance in modulated maskers at lower SNRs (Moore *et al.*, 1999b; Rhebergen *et al.*, 2009). In addition, poorer spectral resolution due to broader auditory filter bandwidths may result in decreased ability to separate speech from noise (e.g., Peters *et al.*, 1998). Jin and Nelson (2010) found a small but significant relationship between spectral resolution and performance in gated noise. However, they also noted that performance in gated speech (in the absence of noise) was highly correlated with degree of MR found for speech in a modulated background, making it less clear that the effects underlying MR are specific to the interactions between the speech and the masker. In addition to temporal resolution, George *et al.* (2006) also evaluated measures of spectral resolution as they related to MR for speech in an amplitude-modulated noise. Beyond the effects of overall level (deteriorating spectral resolution with increasing presentation level for both NH and HI listeners), they did not find any significant correlation between spectral resolution and MR.

Finally, a recent alternative approach has suggested that MR may be predicted by the SNR measured in steady-state noise, independent of other factors (Bernstein and Grant, 2009; Bernstein and Brungart, 2011). These authors noted that greater MR is observed for both HI and NH listeners at lower (poorer) SNRs. Therefore, the fact that HI listeners are often tested at higher SNRs (where MR is less pronounced even for NH listeners) may be sufficient to account for the apparently reduced MR in many cases. However, the authors also note that even after making corrections for differences in SNR across HI and NH listeners, several studies do show some residual (1–4 dB) deficit in MR for HI listeners that may be attributable to factors other than SNR.

The goal of the present study was to assess the role of peripheral compression in MR. There is some limited support in the literature for a potential relationship between compression and MR. Moore *et al.* (1999b) evaluated the potential benefit of fast-acting multi-channel compression for speech understanding in modulated noise and found a small but significant increase in performance with compression. Less is known about whether differences in peripheral compression between individual HI listeners, or the differences between HI and NH listeners, can account for observed differences in MR. Behavioral estimates of cochlear compression can vary among HI listeners, especially those with mild-to-moderate losses in ways that are not predicted by the audiogram (e.g., Plack *et al.*, 2004; Lopez-Poveda *et al.*, 2005; Lopez-Poveda and Johannesen, 2012). Therefore, residual cochlear compression may explain the differences in MR seen among HI listeners with similar audiograms.

In the current study, MR for speech and pure tones was measured in HI listeners and was compared with that in a group of age-matched NH listeners, who served as controls. The measures of MR were then compared with measures of cochlear compression derived from the temporal masking curve (TMC) method (Nelson *et al.*, 2001), and with measures of temporal resolution derived from the recovery of forward masking. Audibility was matched across the two listener groups by presenting the stimuli in a background of noise.

II. EXPERIMENT 1: MASKING RELEASE WITH SPEECH

A. Listeners

Twenty-four listeners participated in the study. Twelve had varying degrees of sensorineural hearing loss (as confirmed by air and bone conduction audiometry, as well as tympanometry), and their ages ranged from 21 to 69 yr (mean age of 50.2 yr). The remaining 12 listeners had audiometrically normal hearing for at least one ear (which was used for testing), defined as thresholds of no more than 20 dB hearing level (HL) at octave frequencies between 250 and 4000 Hz. Their ages ranged from 23-69 yr (mean age of 50.1 yr). The NH listeners were each selected so that their age generally matched that of one of the HI listeners, for whom they would serve as a control. The age of each listener is listed in Table I. Two hours of training was provided for each listener prior to data collection. All listeners were paid an hourly rate for their participation. There were two HI listeners who had asymmetrical sensorineural hearing losses. These listeners had one poorer ear due to a sudden hearing loss, and a better ear that had hearing loss consistent with outer hair cell damage (i.e., due to presbycusis and/or noise exposure). The ears with hearing loss that was reported as sudden in onset were not used as test ears.

B. Threshold elevation

To control for differences in audibility between the NH and HI listener groups, all listeners had their thresholds

TABLE I. Demographic and threshold information for each HI/NMHI listening pair. See text for details.

			Audiometric thresholds (dB HL)				Thresholds in quiet (dB SPL)				Thresholds in TEN (dB SPL)			
	Age years	TEN level (dB SPL/ERB)	500 Hz	1000 Hz	2000 Hz	4000 Hz	500 Hz	1500 Hz	4000 Hz	AVG	500 Hz	1500 Hz	4000 Hz	AVG
HI 1	50	60	40	55	60	65	56.4	65.4	64.4	62.1	74	77.3	73.7	75.0
NMNH1	52	55	5	0	5	0	33.4	31.8	25	30.1	77.2	77.7	72.9	75.9
HI 2	45	45	15	20	25	40	34.4	52.9	49.8	45.7	63.7	65.3	58.8	62.6
NMNH2	50	45	5	15	20	15	40.1	40.9	33.8	38.3	60.1	62.2	60.8	61.0
HI 3	53	50	10	5	30	50	23.2	28	56.7	36.0	75.1	71.2	67.8	71.4
NMNH3	56	50	15	10	0	20	33.1	26.7	27.9	29.2	66.4	77.9	66	70.1
HI 4	59	55	20	30	40	50	39.4	44.3	53	45.6	73.3	72.1	70	71.8
NMNH4	51	50	10	10	0	0	25.9	29.2	24.1	26.4	70.1	70.9	73.1	71.4
HI 5	52	45	15	25	35	35	34.2	40.8	45.8	40.3	60.6	63.4	59.4	61.1
NMNH5	61	45	10	15	10	20	24.3	25.6	27.3	25.7	63.3	65.2	61.7	63.4
HI 6	21	60	55	45	65	60	69.3	62.9	72.9	68.4	82.9	80.1	77	80.0
NMNH6	25	60	10	10	15	15	31.8	24.9	28.3	28.3	79.3	84	76.3	79.9
HI 7	23	60	55	60	50	10	78	72	26.8	58.9	78	78.7	76.3	77.7
NMNH7	23	55	−5	−5	−5	−10	23.1	28.9	23.3	25.1	74.8	81.7	74.6	77.0
HI 8	57	60	25	40	40	45	37	54.4	65	52.1	77.1	76.7	72.2	75.3
NMNH8	56	55	10	5	0	10	35.8	28.2	30.1	31.4	76.1	83.9	75.2	78.4
HI 9	54	45	35	30	25	15	39.1	40.2	30.8	36.7	61.1	65.1	63.5	63.2
NMNH9	59	45	5	5	10	5	23.1	26	23.4	24.2	65.1	64.7	62.4	64.1
HI 10	64	55	45	60	50	55	56	67.1	59.4	60.8	69.9	73.8	68.8	70.8
NMNH10	61	50	10	5	10	15	26.2	39.3	26.3	30.6	77.8	76.7	69.2	74.6
HI 11	66	40	0	10	15	45	25.3	30.6	46.3	34.1	60.2	62.1	54.8	59.0
NMNH11	69	40	10	5	10	15	30.7	31.9	31.9	31.5	59.2	59.3	55.2	57.9
HI 12	58	65	45	45	55	70	50.6	58.3	75.9	61.6	80.2	82.6	85.2	82.7
NMNH12	62	65	10	0	0	15	30.6	27.7	28.6	29.0	86.9	83	84.6	84.8

elevated using threshold equalizing noise (TEN, Moore *et al.*, 2000), which is broadband noise designed to produce equal pure-tone masked thresholds in dB sound pressure level (SPL) at all frequencies. Both groups listened in noise to remove any possible confounds that may have been caused by having one group listen in background noise and the other in quiet. Based on the findings by Gregan *et al.* (2010), it was expected that NMNH listeners should have normally compressive peripheral responses to sounds above masked threshold in the noise, and that the background noise should only serve to decrease their audibility. This approach has been used in previous studies to better equate thresholds across groups of listeners (e.g., Eisenberg *et al.*, 1995; Dubno *et al.*, 2002, 2003). Pure-tone masked thresholds were used to determine the appropriate TEN level for each subject. The tones were short-duration sinusoids at 500, 1500, and 4000 Hz. The total probe duration was 4 ms (including 2-ms raised-cosine onset and offset ramps; no steady-state) for 1500 and 4000 Hz and 10 ms (5-ms ramps) for 500 Hz. A two-down one-up adaptive procedure using a three-interval three-alternative forced-choice task was used to estimate threshold. The TEN was on continuously throughout all three intervals and the signal was presented at random in one of the three intervals. The listener's task was to select the interval containing the signal. The initial step size was 8 dB for the first two reversals, then changed to 4 dB for the next two reversals and then remained at 2 dB for the final six reversals. The threshold was taken as the average of the signal level at the last six reversal points. Three such averages were used as the final threshold for a given TEN level. The TEN levels were selected for each HI listener individually. The lowest level of TEN that raised their pure-tone thresholds in quiet at 500, 1500, and 4000 Hz by at least 5 dB was used. An exception was made for HI7 who had a reverse slope audiogram (i.e., higher thresholds in dB HL at low than at high frequencies); to avoid using uncomfortably loud noise levels, this listener was tested with a TEN level that shifted her 4000 Hz threshold to match her 500 and 1500 Hz thresholds. Each NMNH subject listened in a level of TEN that yielded similar pure-tone masked thresholds (across all frequencies) to those found for their HI counterpart. (Note that this meant some NH listeners listened in a different level of TEN than did their HI counterpart, as the goal was to equate masked thresholds and not TEN levels *per se*.) As expected, the TEN resulted in pure-tone thresholds that were relatively independent of frequency across the tested range. The audiometric thresholds, ages, unmasked thresholds for the pure-tone frequencies (500, 1500, and 4000 Hz) for each of the 24 listeners, as well as the masked thresholds for each listeners for the noted level of TEN are shown in Table I. Note that most, but not all, of the masked thresholds in TEN are similar across the three frequencies tested, as expected. The cause of some deviations across frequency (e.g., NMNH3 and HI11) is not clear, but it may be related to our use of brief tones – the TEN was designed to produce equal masked thresholds across frequency for long-duration tones (Moore *et al.*, 2000). Overall, the difference between the average TEN threshold for each HI listener and their NMNH counterpart was 3 dB or less. These TEN levels were used for all experiments described in this paper, whenever TEN was added. A new sample of TEN was generated for each stimulus presentation.

C. Stimuli and procedure

Intelligibility of IEEE sentences (IEEE, 1969) was measured in a masking noise that was spectrally shaped to match the long-term power spectrum of the sentences and was either unmodulated or 100% modulated with a 10-Hz square wave. The modulated noise was gated after scaling, so that its level during the "on" period was the same as that of the unmodulated noise, but the overall RMS of the modulated noise was 3 dB lower than that of the unmodulated noise. The choice of a 10-Hz modulation rate was based on results from previous studies that showed a peak in the MR function at this rate (e.g., Miller and Licklider, 1950; Gustafsson and Arlinger, 1994). Both the speech and the masking noise were bandpass filtered between 500 and 4000 Hz. The masking noise duration was 2000 ms (slightly longer than the longest duration sentence), gated on and off with 5-ms raised-cosine ramps, and the sentence was temporally centered within the masking noise. The additional TEN was gated on and off synchronously with the masking noise with 5-ms raised-cosine ramps. New noise samples were generated for each presentation. The speech level was fixed across conditions and the masker level was varied to measure the proportion of correct keywords in the sentences over a range of SNRs. The TEN level was held constant, as shown in Table I, and was not included in the calculation of SNR. The same RMS speech level (after bandpass filtering) was used for each HI/NMNH pair, as shown after the subject identifier in each panel of Figs. 1 and 2. The level of the speech was determined as part of the pilot testing for each HI listener and was set to yield 80% or better performance in

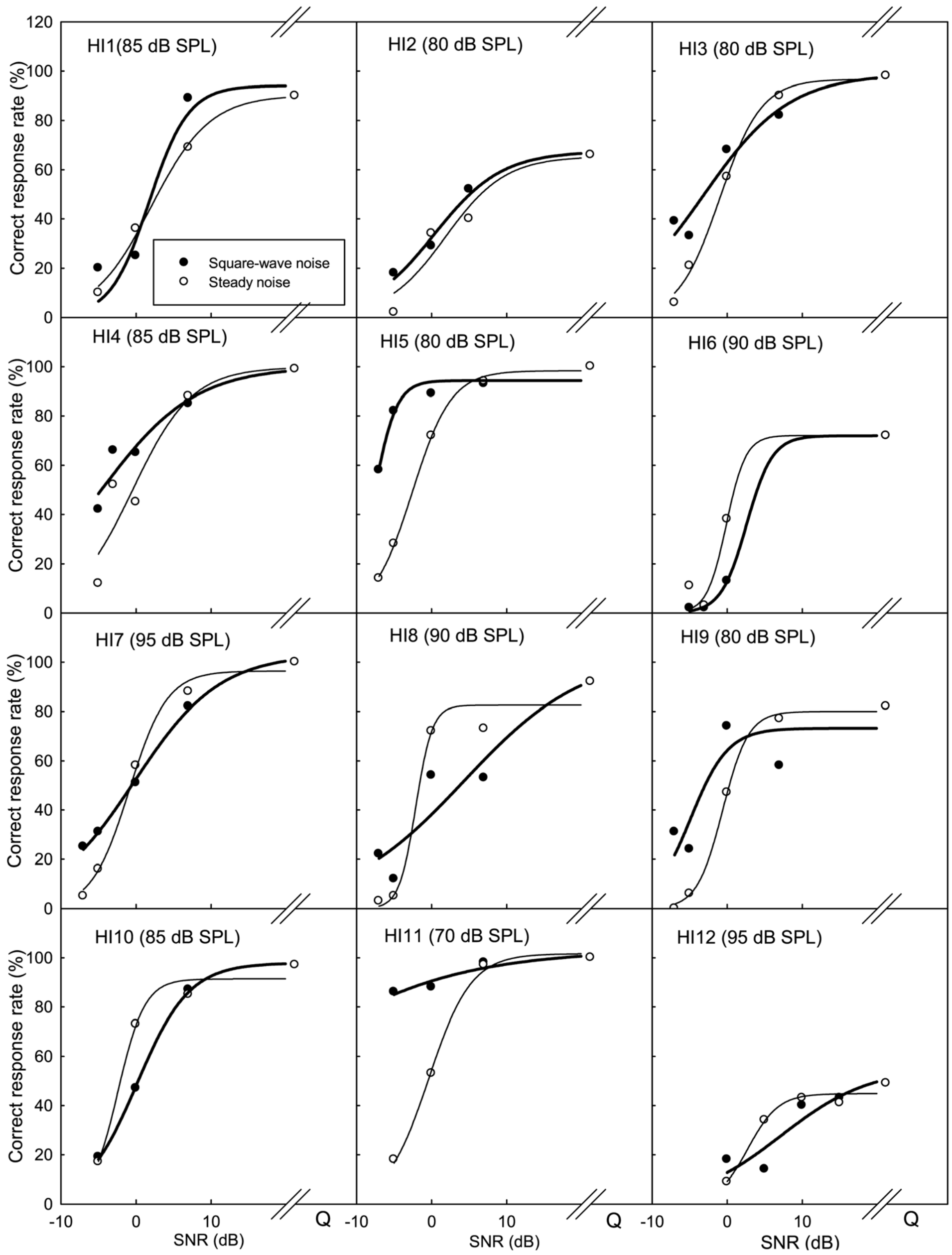

FIG. 1. Individual correct response rate (percent) scores from HI listeners for word identification in IEEE sentences as a function of SNR. The sentences were presented in a steady-state noise background (open circles) or a square-wave gated noise background (filled circles). The thin and thick curves represent best fits (least-squares criterion) to the steady-state and gated noise conditions, respectively, using a three-parameter sigmoidal function. The number in parentheses in each panel shows the sound pressure level at which the speech was presented.

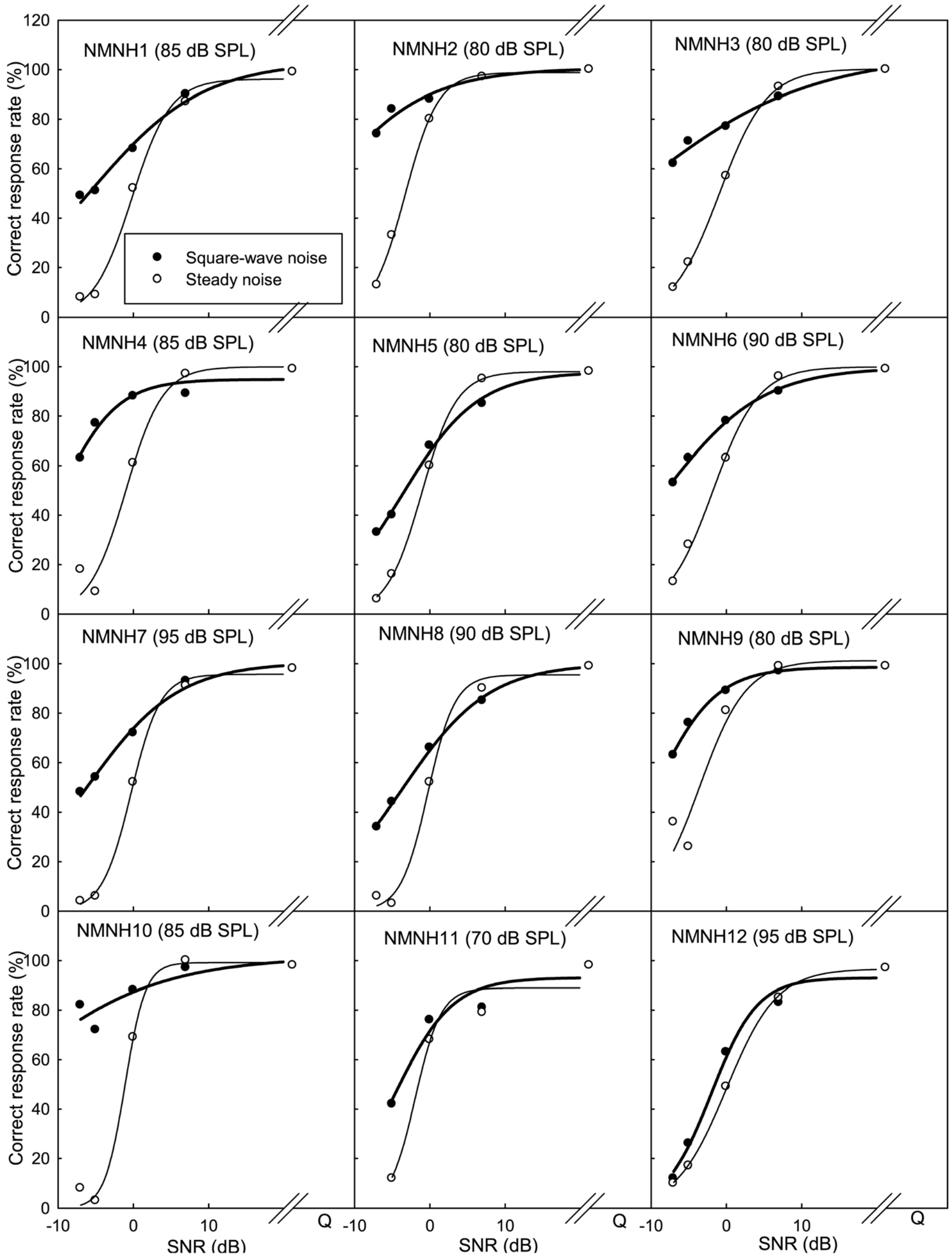

FIG. 2. Results are plotted as in Fig. 1, but for noise-masked normal-hearing (NMNH) listeners. Note that data from the age- and audibility-matched pairs of HI and NMNH listeners are shown in the same panel in the two figures.

quiet while not exceeding tolerable loudness levels. This was possible for all HI listeners except HI12, for whom quiet scores did not exceed 50% correct, even with speech presented at 95 dB SPL in quiet. Two lists of IEEE sentences were run for each SNR and masker condition. Listeners were instructed to type what they heard using a computer keyboard. Practice was provided prior to data collection to familiarize listeners with the experimental procedure. Feedback was provided during training but not during testing. Listeners were informed that the sentences they would hear may not make sense, and were instructed to report as many words as they could. The speech MR experiment took between 1.5 and 3 h to complete.

No frequency-shaping or amplification was applied to the speech stimuli. Therefore, it is possible that the speech spectrum was not audible over the entire 500–4000 Hz range for all listeners. However, the audibility should have been the same for both HI and NMNH listeners in each pair, due to the presence of the TEN. Thus, although lack of full audibility may affect overall performance in both types of masking noise, it should not differentially affect the performance across listening conditions, nor should it affect the performance within each age- and audibility-matched HI and NMNH pair. Therefore, any differences in MR within a given HI and NMNH listener pair cannot be readily explained by audibility.

The stimuli were generated digitally and converted to an analog signal via a 24-bit Lynx22 (LynxStudio) soundcard at a sampling rate of 22.05 kHz. Sounds were presented to one ear via Sennheiser HD580 headphones. The test ear was either the ear with hearing loss (for unilaterally HI listeners) or the ear with thresholds closer to the 40-50 dB HL range (for listeners with asymmetrical hearing losses). The ear tested for the NH listeners was either their preferred ear (if neither ear had significantly better hearing) or their better hearing ear. This same ear was used as the "test ear" for all remaining experiments described in this paper. A 2/3-octave wide

Gaussian noise masker centered at the signal frequency and presented an overall level 40 dB below the level of the signal was presented to the non-test ear to avoid audible electrical and acoustical crosstalk. It was gated on and off with each of the three listening intervals. The subjects were tested in a double-walled sound-attenuating booth.

D. Results and discussion

Results for the 12 HI listeners and the 12 NMNH listeners are shown in Figs. 1 and 2, respectively. Note that the position of each listener in the figure coincides with their control (i.e., HI1 in the upper left hand corner of Fig. 1 is the HI counterpart for NMNH1 in the upper left hand corner of Fig. 2). The raw data points are indicated by circles (filled for results in the square-wave-gated noise and open for the results in the steady-state noise). The data are plotted as correct response rate (in percent) as a function of SNR. Note that not all SNRs were tested for each listener, due to time and individual performance constraints. In particular, HI12 was not able to perform the task at SNRs lower than 0 dB. Also note that the SNR designated "Q" on the *x* axis is the correct response rate (percent) in "quiet" (with just the TEN present). It can be seen that when MR was observed, the SNR was typically less than 0 dB. In other words, the lower SNR conditions tended to yield the higher MR values. This observation has been made in previous studies (e.g., Takahashi and Bacon, 1992; Bernstein and Grant, 2009; Oxenham and Simonson, 2009), and is particularly apparent in the NMNH data (Fig. 2).

A 3-parameter sigmoidal function was fitted to the data, to allow for prediction of SNR required for 50% correct performance in each noise background. The fits are shown as solid lines in the figures. The equation used was:

$$y = a/\{1 + \exp[-(x - x_0)/b]\}, \qquad (1)$$

where y is the proportion of words correctly reported, x is the SNR in dB, a and b are free parameters that control the maximum (asymptotic) y value and the slope of the transition region, respectively, and x_0 is the free parameter that determines the SNR at 50% correct.

The proportions of variance accounted for (R^2), indicating the goodness of the sigmoidal fits, are shown in Table II. Generally, the fits were reasonable, with the majority of R^2 values at 0.80 or higher. To analyze the results, two summary measures were initially investigated. The first summary measure was based on the improvement in speech recognition performance at a fixed SNR. The SNR of −5 dB was selected because previous studies had shown most MR at negative SNRs, and because all but one listener (HI12, see above) were tested at this SNR. The difference in correct response rate scores (in percentage points) at −5 dB SNR between the square-wave gated and the steady-state noise conditions is shown in the third column of Table II [labeled "MR: PC at −5 dB SNR" where PC stands for "correct response rate (percent)"]. By comparing this value of MR across each HI and NMNH pair, it can be seen that almost all NMNH listeners showed more MR than their HI counterparts, with the exceptions of HI/NMNH5 and HI/NMNH11. A paired-samples *t*-test confirmed that the MR at

TABLE II. Proportion of variance (R^2) account for by sigmoidal fits to speech understanding performance-intensity (SNR) functions (columns 1 and 2). Summary values for speech MR (columns 3 and 4) as described in text. Change in the correct response rate, in terms of percentage points, is shown in column 3 at a SNR of −5 dB. *Calculated at 0 dB SNR because listener was unable to perform task at −5 dB SNR. SQ = square wave gated noise; SSN = steady state (unmodulated) noise.

	R^2 SQ fit	R^2 SSN fit	MR: PC at −5 dB SNR (SQ-SSN in %)	MR: SRT (SQ-SSN in dB)
HI 1	0.94	0.99	10	−0.28
HI 2	0.98	0.92	16	−1.46
HI 3	0.96	0.99	12	−2.19
HI 4	0.92	0.89	30	−4.14
HI 5	0.94	0.99	54	−5.16
HI 6	0.99	0.96	−9	2.70
HI 7	0.99	0.99	15	0.68
HI 8	0.88	0.97	7	6.42
HI 9	0.77	0.99	18	−5.44
HI 10	0.99	0.98	2	2.54
HI 11	0.92	0.99	68	−22.79
HI 12	0.86	0.97	9*	4.40
NMNH 1	0.99	0.99	42	−5.29
NMNH 2	0.96	0.99	51	−11.40
NMNH 3	0.99	0.99	49	−10.79
NMNH 4	0.93	0.98	68	−8.88
NMNH 5	0.99	1.00	24	−2.46
NMNH 6	0.99	0.99	35	−6.34
NMNH 7	0.99	0.99	48	−5.83
NMNH 8	0.99	0.99	41	−3.24
NMNH 9	0.99	0.94	50	−5.96
NMNH 10	0.77	0.99	69	−16.81
NMNH 11	0.93	0.96	30	−2.62
NMNH 12	0.99	1.00	9	−1.99

−5 dB SNR was significantly different for NMNH than for HI listeners [$t(11) = -2.669$; $p = 0.022$].

The second summary measure was the difference between the SNR required for 50% correct for performance in square-wave gated noise and the SNR for 50% correct in steady-state noise. This speech reception threshold (SRT) was derived using the x_0 values from the sigmoidal fits to the data. These values are shown in the last column of Table II (labeled "MR: SRT"). Note that in many cases (2 HI and 6 NMNH listeners) these values had to be extrapolated from the existing data as the scores in gated noise did not drop to 50% correct, even at the lowest SNRs tested. Because of this potential flaw, the SRT measure is not analyzed further in this paper; however, it should be noted that it was highly correlated with the improvement in word recognition rate at −5 dB SNR (Pearson product moment correlation, $R^2 = 0.728$; $p < 0.001$).

Recently Bernstein and Grant (2009) and Bernstein and Brungart (2011) have explored the hypothesis that HI listeners exhibit less MR in part because they require more positive SNRs to achieve a given correct response rate in steady-state noise, as most masking release is observed at negative SNRs. This hypothesis was evaluated further with our data by plotting the SRT difference as a function of the SRT in steady-state noise. According to Bernstein and colleagues' hypothesis, the measures should be correlated;

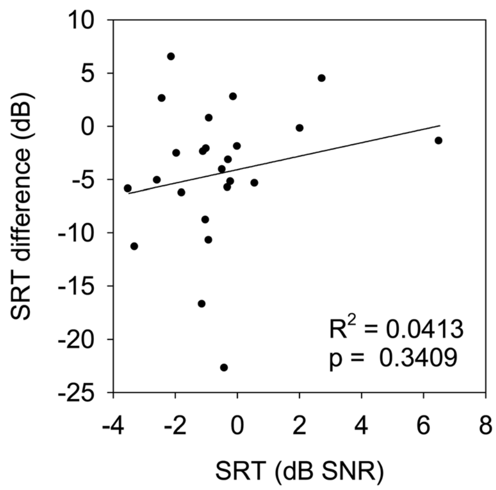

FIG. 3. Difference in SRT between the steady-state and square-wave gated noise as a function of the SRT in steady-state noise for HI and NMNH listeners.

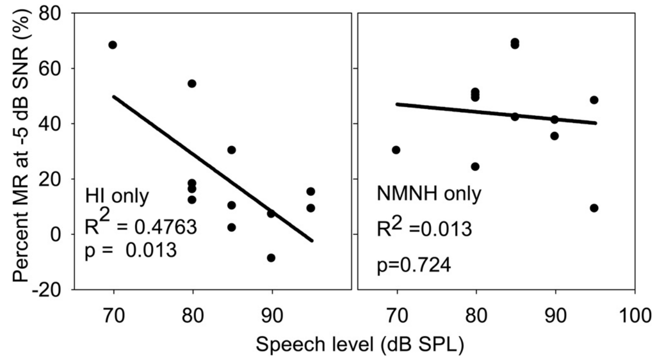

FIG. 4. Masking release, in terms of improvement in word recognition rate for sentences in square-wave gated noise compared with steady-state noise at −5 dB SNR, is plotted as a function of speech level. The left panel shows data from the HI group; the right panel shows data from the NMNH group. The proportion of variance accounted for (R^2) and statistical significance of the linear regression (p value) are shown in each panel.

greater MR should be seen for lower SRT in steady-state noise. The data are shown in Fig. 3. A linear regression analysis failed to find a significant relationship for this comparison for the HI listeners alone (not shown) or the combined NMNH and HI listener groups. This outcome suggests that the decrease in MR seen for our HI listeners compared to their NMNH counterparts cannot be accounted for fully by differences in SRT in steady-state noise.

Rhebergen and colleagues (Rhebergen and Versfeld, 2005; Rhebergen *et al.*, 2006, 2009) have provided a quantitative model for predicting speech intelligibility in fluctuating maskers by extending the Speech Intelligibility Index (SII; ANSI, 1997). Their Extended SII (ESII) operates by calculating the SII within short time windows and then averaging the resulting short-term SII values. Because the original ESII is based on audibility within short-term windows, and because audibility was equated between the HI and NMNH groups, we do not expect that the original model will account for the observed difference in MR between the HI and NMNH groups. However, it is possible that a further extension of the model, which incorporates forward masking (Rhebergen *et al.*, 2006), may predict a difference, providing that differences in temporal resolution are observed between the two groups. The possibility of differences in temporal resolution is tested in experiment 3.

The TEN was designed to equate audibility within each pair of NH and HI listeners, so it is unlikely that differences in audibility can explain differences in MR between the HI listeners and their matched NMNH listeners. However, the TEN level and overall speech level differed between pairs of listeners, depending on the absolute thresholds of the HI listener. It may be that MR is related to overall presentation level. To test this, the overall speech levels were compared to MR in both the HI and NMNH groups, as shown in Fig. 4. The correlation between MR and speech level was significant for the HI group, indicating that MR decreased significantly as the speech presentation level increased. The correlation was not significant for the NMNH group, however, suggesting that it was not the speech level *per se* that determined MR but rather something related to the degree of hearing loss.

To test this conjecture more directly, the MR measure for HI listeners was also plotted as a function of quiet threshold, averaged across the three test frequencies, which were selected to span the range of the speech bandwidth (Fig. 5). These plots show a significant correlation between average threshold in quiet and degree of MR (Pearson product-moment correlation; $R^2 = 0.5$, $p = 0.01$). In other words, listeners with less hearing loss tended to show more MR. However, absolute thresholds accounted for only around half the variance of the MR measure, suggesting other factors may also play a role.

Our finding of more MR for NH than for HI listeners, even when equated for audibility, age, and sound level, is consistent with previous studies (e.g., Eisenberg *et al.*, 1995; Bacon *et al.*, 1998). The goal of the remaining experiments in the current study was to attempt to determine what additional factors might explain this difference in performance.

III. EXPERIMENT 2: MASKING RELEASE WITH PURE TONES

In this experiment, masking release for pure tones (MRPT) was measured to determine if audibility of simple

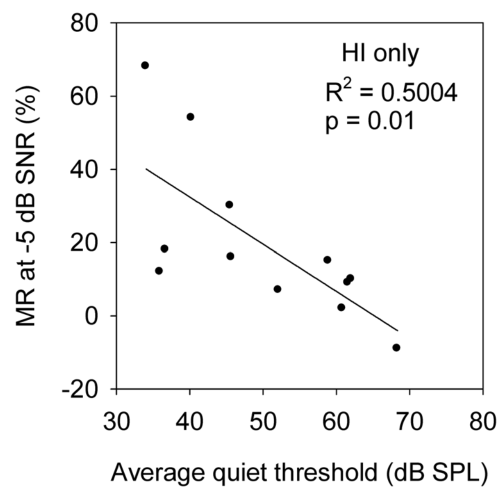

FIG. 5. Masking release in HI listeners only as a function of average pure-tone threshold in quiet across 500, 1500, and 4000 Hz. Panel shows increase in correct responses (in percentage points) at −5 dB SNR.

signals in fluctuating noise predicts MR for speech. The pure-tone frequencies of 500, 1500, and 4000 Hz were selected to span the frequency range of the speech used in experiment 1. Similar to speech in a modulated background, it has been shown that listeners are better able to detect a brief tone when it coincides with a temporal valley in the masker than when it coincides with a temporal peak (e.g., Egan and Hake, 1950; Zwicker, 1976; Buus, 1985; Glasberg and Moore, 1994; Kohlrausch and Sander, 1995; Nelson and Swain, 1996). If MR found with speech stimuli is due primarily to the increased audibility of the speech within masker valleys, then MR in speech should be correlated with MR found with simpler (e.g., pure-tone) stimuli. Oxenham and Dau (2004) compared pure-tone masked thresholds in positive and negative Schroeder-phase (Schroeder, 1970) complex-tone maskers in HI and NMNH listeners, and found smaller differences between the two maskers with the HI listeners than with the NMNH listeners, implying less MR in the HI listeners. The difference between the NMNH and HI listeners could not be attributed to audibility, due to the presence of background noise in both groups. Instead, Oxenham and Dau (2004) suggested that the difference may be due to reduced basilar-membrane compression in the HI listeners, which would result in a loss of amplification of the signal during low-level portions of the masker, as well as less difference in effective overall excitation produced by modulated and unmodulated maskers. Similarly, a study examining masking period patterns in NH and HI listeners showed masking period patterns that were similar for on- versus off-frequency maskers for the HI listeners, consistent with a reduction in basilar membrane compression (Wojtczak *et al.*, 2001).

A. Listeners and stimuli

All 24 listeners from experiment 1 also participated in this experiment. Signal tones were fixed in level at 8 dB SL (1500 and 4000 Hz) or 10 dB SL (500 Hz), where SL is referenced to detection threshold in the presence of the TEN. The 500-Hz tone was presented at 10 dB SL because several listeners were not consistently able to detect the 8 dB SL tone in pilot studies. The signal duration was 4 ms, including 2-ms raised-cosine rise/fall ramps (no steady state) for 1500 and 4000 Hz. The signal duration at 500 Hz was 10 ms (including 5-ms raised-cosine onset and offset ramps) to avoid audible "spectral splatter" and possible physical overlap between the BM response to the masker and signal at this low frequency (e.g., Shailer and Moore, 1987). The long-term power spectrum of the noise masker was matched to that of the IEEE sentences, with the exception that the noise was not subsequently bandpass-filtered between 500 and 4000 Hz. The noise was either unmodulated or was 100% modulated with a 10-Hz square wave and a 50% duty cycle (meaning the noise was on for 50% of the time and off for 50% of the time). The masker duration was 500 ms (i.e., 5 periods of the gated masker, beginning with an on-period). The signal was placed at the temporal center of the third masker on-period (225 ms after the start of the masker), or at one of three locations in the following masker off-period (12.5, 25, or 37.5 ms after the end of the third on-period, i.e., 262.5, 275, or 287.5 ms after the start of the masker), as depicted in Fig. 6. For the unmodulated masker, the signal was presented 225 ms after masker onset, at the same location as the signal in the on-period of the gated masker. As in Experiment 1, all the stimuli were embedded in TEN, selected for each subject pair individually, to equate audibility in the absence of the masker. These levels are listed in Table I. The TEN was gated on 300 ms before the beginning of the first interval and was gated off 300 ms after the end of the third interval in each trial.

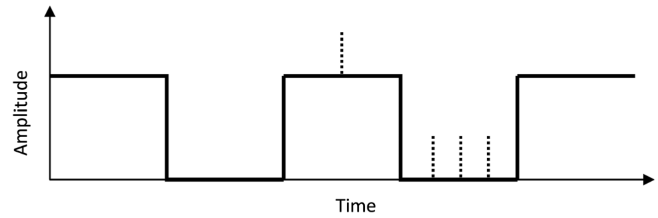

FIG. 6. Schematic diagram showing the temporal location of signals (vertical dashed lines) in relation to the gated masker for the pure-tone masking release experiment. Note that the modulation depth is 100% for the square wave gated masker.

B. Procedure

The masker levels at signal detection threshold were measured using a three-interval three-alternative forced-choice method, with a fixed signal level and a masker level that was adaptively varied with a two-up, one-down rule to track the 70.7% correct point on the psychometric function (Levitt, 1971). The three intervals in each trial were separated by 300-ms interstimulus intervals. The masker noise was presented in all three intervals, and the signal was presented in one, chosen at random in each trial with uniform probability. The listener's task was to select the interval that contained the signal. The masker level was initially varied with a step size of 8 dB, which was reduced to 4 dB after the first two reversals, reduced to 2 dB after two more reversals, and was then held constant for the remaining six reversals in each adaptive run. Threshold calculation for each run was based on the average masker level at the last six reversal points and threshold for each condition and subject was taken as the average of three runs. The presentation order of temporal position and masker type was randomized within and across subjects and repetitions. Each listening session was 2 h in length, including frequent breaks, and the experiment took a total of 1 to 2 sessions per subject to complete.

In the adaptive procedure, the maximum RMS level of the masker was not allowed to exceed 110 dB SPL for the HI listeners and four of the NH listeners, and was not permitted to exceed 103 dB SPL for the remaining NH listeners. Runs in which the adaptive procedure called for a masker level that exceeded these limits were aborted and the maximum value was used in lieu of an actual threshold value for that condition. This means that for listeners who required a higher masker level than the allowed maximum output, the difference in level required to mask the signal at a masker peak versus a masker valley may be underestimated.

TABLE III. Amount of masking release (in dB) for pure-tone signals (MRPT) for the three frequencies measured. The last column shows the mean MRPT, averaged across the three frequencies.

	500 Hz	1.5 kHz	4 kHz	Average
HI1	30.6	22.2	22.1	25.0
HI2	16.8	43.5	8.6	23.0
HI3	0.1	34.6	6.7	13.8
HI4	4.8	8.5	15.6	9.6
HI5	5.3	26.8	9.9	14.0
HI6	0.4	4.9	8.5	4.6
HI7	11.8	16.1	14.1	14.0
HI8	20.4	26.7	11.1	19.4
HI9	30.1	21.0	17.2	22.8
HI10	3.8	−0.6	3.2	2.1
HI11	49.6	33.2	10.1	31.0
HI12	0.6	9.1	6.2	5.3
NMNH1	12.4	8.5	5.3	8.7
NMNH2	34.9	30.4	26.3	30.5
NMNH3	35.6	40.9	26.7	34.4
NMNH4	15.2	29.8	17.6	20.9
NMNH5	8.6	18.8	23.0	16.8
NMNH6	24.2	11.6	11.1	15.6
NMNH7	16.8	14.3	14.9	15.3
NMNH8	10.7	7.0	13.5	10.4
NMNH9	3.0	27.4	16.9	15.8
NMNH10	25.2	20.2	19.5	21.6
NMNH11	28.3	36.3	31.5	32.0
NMNH12	11.9	11.4	7.9	10.4

C. Results

Masker levels at threshold for all listeners were highest when the signal was in either temporal position 3 or 4. Therefore, the threshold masker levels in these two conditions were averaged and then subtracted from the threshold level in the steady-state (unmodulated) masker to provide the summary measure of MRPT for each subject. These values are listed in Table III for each of the test signals, along with an average MRPT value across test frequencies. Positive values indicate that a higher masker level was required to mask the signal in the modulated than in the unmodulated masker condition.

There was no significant difference in the magnitude of MRPT between the HI and NMNH groups at any of the three test frequencies, as evaluated using paired *t*-tests [500 Hz: $t(11) = -0.770$, $p = 0.458$; 1500 Hz: $t(11) = -0.237$, $p = 0.817$; 4000 Hz: $t(11) = -2.097$, $p = 0.060$]. In addition, as can be seen from Table III, there was a great deal of variability in the amount of MRPT within each group of listeners. There was essentially no negative MR seen in either group of listeners, suggesting that impaired performance in modulated, compared with unmodulated noise does not occur for pure tones and may be specific to more complex stimuli, such as speech, presumably due to modulation interference between the masker and speech envelopes (e.g., Jorgensen and Dau, 2011). In other words, the modulation introduced by the fluctuating masker may lead to masking of fluctuations in the speech envelope. In contrast, envelope modulation of a broadband carrier is unlikely to interfere with (and often helps in) the detection of pure tones.

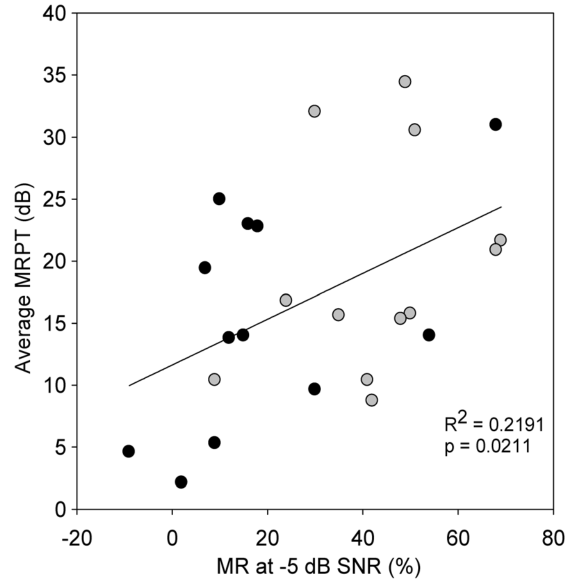

FIG. 7. Average MRPT across all 3 signal frequencies as a function of the MR for speech for both the HI (black symbols) and NMNH (gray symbols) listeners. The speech MR is defined as the improvement in performance (percentage points) at −5 dB SNR.

The possible relationship between MRPT and speech MR was explored. Figure 7 shows the MRPT for each subject from both HI and NMNH groups, averaged across the three signal frequencies, plotted against the MR for speech. Linear regression analysis showed that this relationship was significant (Pearson product-moment correlation, $R^2 = 0.2191$, $p = 0.021$). This significant correlation shows a relationship between the audibility of pure tones and the intelligibility of speech in modulated vs steady-state noise. Thus, MR for speech may reflect in part simple audibility differences between steady-state and modulated maskers, as reflected in pure-tone thresholds. Nevertheless, the percentage of variance accounted for is relatively low, suggesting that other factors also play a role.

IV. EXPERIMENT 3: ESTIMATING COMPRESSION AND TEMPORAL RESOLUTION

The relationship between cochlear compression and MR for speech and pure tones was explored here by estimating cochlear compression for frequencies across the speech bandwidth (500, 1500, and 4000 Hz) using the TMC technique (Nelson *et al.*, 2001). A recent study has shown that compression estimates using distortion product otoacoustic emissions and TMCs were in good agreement, suggesting that both may be effective at measuring the non-linear functioning of the human cochlea (Lopez-Poveda and Johannesen, 2009). Because hearing loss can vary across frequency, it is important to have a means of estimating compression for several frequencies across the range important for speech. A recent study (Jürgens *et al.*, 2011) compared compression estimated via TMCs with compression estimated via categorical loudness judgments and found similar results for the two techniques, but with the loudness judgments being less time-consuming. However, it is known that loudness judgments can be affected by the presence of background noise (Scharf, 1964), which may make it less suitable for our purposes. Growth of masking (GOM) has been used in the past to estimate cochlear compression (Oxenham and Plack, 1997a), but it is not an ideal method for estimating compression at lower signal frequencies, because it relies on the assumption that frequencies well below the test frequency are processed linearly at the place along the BM with a CF corresponding to the test frequency. At lower frequencies (e.g., 500 Hz), it is not clear from behavioral (e.g., Plack and Drga, 2003) or physiological (e.g., Cooper and Yates, 1994; Robles and Ruggero, 2001) data whether this assumption holds. Indeed, some studies have suggested that although the apex of the cochlea responds compressively to sound, the compression may not be as frequency-specific as the response at the (high-frequency) base of the cochlea, meaning that the response remains compressive over a much wider range of frequencies (e.g., Cooper and Rhode, 1995; Lopez-Poveda *et al.*, 2005; Rosengard *et al.*, 2005). Although the TMC method, in its original implementation, has similar issues with regard to estimating compression at the low frequencies, several researchers have attempted to circumvent this problem by using an off-frequency linear reference for the highest signal frequency as the linear reference for all signal frequencies (e.g., Lopez-Poveda *et al.*, 2003; Lopez-Poveda *et al.*, 2005), making the explicit assumption that the decay of forward masking is relatively frequency-independent. This approach is also used in the present study to estimate compression at CFs between 500 and 4000 Hz, by using the off-frequency masking curve with a 4000-Hz signal as the linear reference for all three signal frequencies.

A. Listeners and stimuli

The same 24 listeners from experiments 1 and 2 took part in this experiment. The stimuli were generated digitally at a sampling rate of 48 kHz and were presented as described in experiment 1. For all listeners, the signal levels were fixed at a low sensation level in the TEN background (8 dB SL for 1500 and 4000 Hz and 10 dB SL for 500 Hz), with signal parameters as described in experiment 2. The masker duration was 200 ms (including 5-ms raised-cosine onset and offset ramps), and the gap between the masker and signal (measured at the 0-V points of the envelopes) varied from 0 to 60 ms (actual delays used varied across subjects). The TEN was present to elevate thresholds and equate them across the HI and NMNH listeners in each subject pair, as described in Sec. II B. For the on-frequency conditions, the signal and masker frequencies were equal and were either 500, 1500, or 4000 Hz. For the 4000-Hz signal, masker thresholds were also measured in an off-frequency condition, where the masker frequency was 1800 Hz ($0.45f_s$, where f_s is the signal frequency). A frequency ratio of more than an octave was selected based on data from Plack and Arifianto (2010) and Lopez-Poveda and Alves-Pinto (2008), which suggest that the signal and masker should be separated by more than an octave to ensure a linear response to the masker at the location with a CF corresponding to the signal frequency. The off-frequency condition was only tested for the 4000-Hz signal, because earlier studies have suggested that the off-frequency-masker conditions at lower signal frequencies may not reflect truly linear processing. For the purposes of analysis, it was assumed that the function relating off-frequency masker level to masker-signal gap reflected temporal resolution, without influence of peripheral compression, and that this temporal decay was the same for all signal frequencies.

B. Procedure

The TMC functions were measured using a three-interval three-alternative forced-choice method, with a fixed signal level and a masker level that was adaptively varied with a two-up, one-down rule to track the 70.7% correct point on the psychometric function (Levitt, 1971). Various masker-signal delays were employed, which varied across listeners. The three intervals in each trial were separated by 300-ms interstimulus intervals. The TEN noise was gated on 300 ms before the first interval and was gated off 300 ms after the third interval. The pure-tone masker was presented in all three intervals and the signal was presented in one, chosen at random in each trial with uniform probability. The listener's task was to select the interval that contained the signal. The initial step size for the adaptively varying masker level was 4 dB, which was reduced to 2 dB after the second reversal point, and was held constant for the remaining six reversals in each adaptive run. Threshold calculation for each run was based on the average masker level at the last six reversals and threshold for each condition was taken as the average of three runs. The presentation order of delay was randomized across subjects and repetitions. Listening sessions were 2 h in length, including frequent breaks. The TMC experiment took approximately 4 to 5 sessions per subject to complete.

C. Results

The TMC data for HI and NMNH listeners are shown in Figs. 8 and 9, respectively. As described in Nelson *et al.* (2001), these curves were used to derive estimates of BM input–output functions by plotting the on-frequency TMC for a given masker-signal delay on the x axis and the off-frequency TMC for 4000 Hz at the same masker-signal on the y axis. For the on-frequency functions, the increase in masker level required to mask the signal with increasing masker-signal delay is assumed to be due to the recovery from forward masking as well as compression of the masker at the BM location with a CF corresponding to the signal frequency. The increase in off-frequency masker level with increasing masker-signal delay is assumed to be due to recovery from forward masking alone, as it is assumed that response to the relatively low-frequency masker is linear at the BM location tuned to the signal frequency (e.g., Nelson *et al.*, 2001). To reduce the effects of measurement variability at single points on the function, the off-frequency 4000-Hz TMC was fitted with a straight line. The fit to the off-frequency masker-level thresholds as a function of masker-signal delay was then used as the "linear" reference to derive estimates of BM input-output function for each

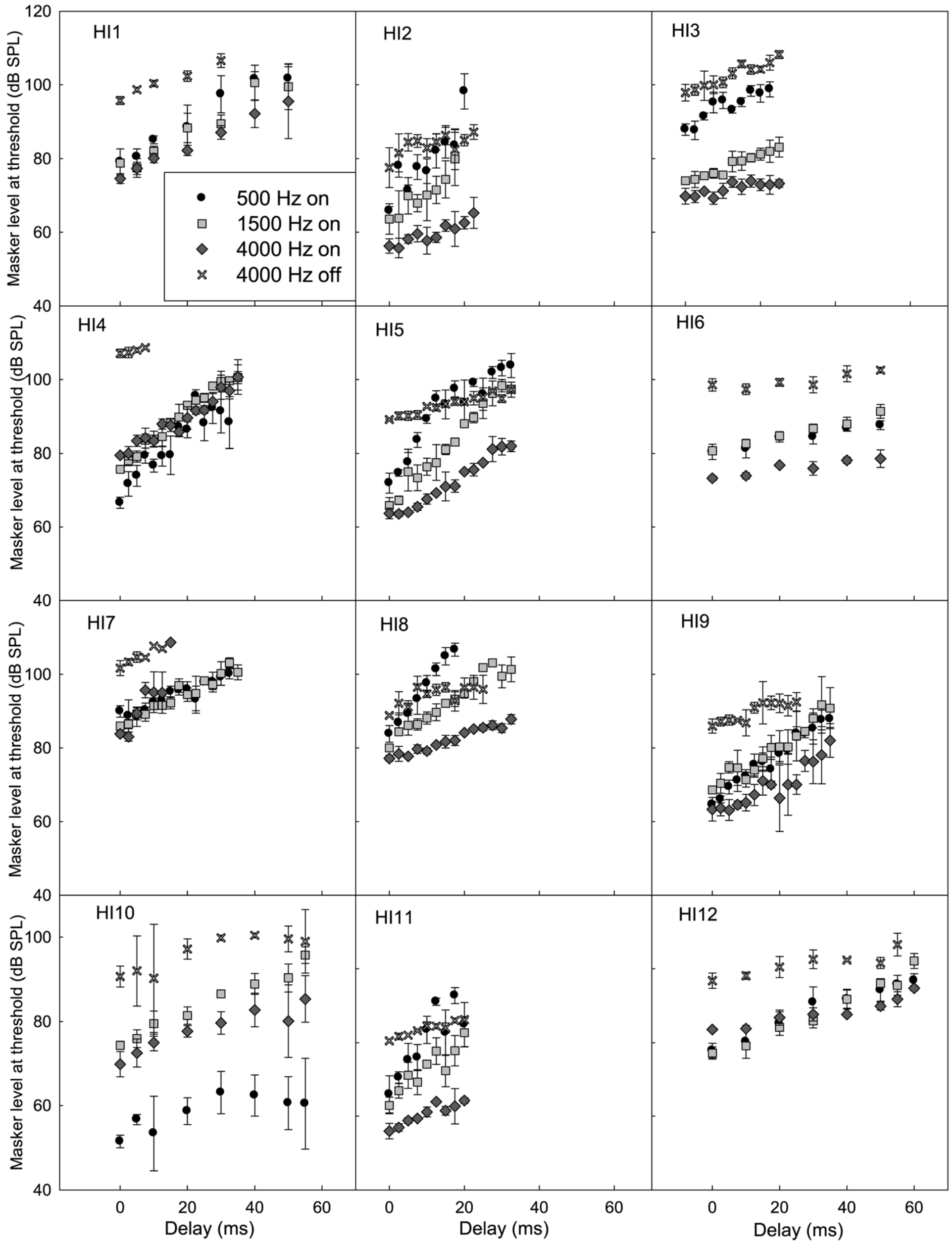

FIG. 8. Masker levels at threshold as a function of the delay between the forward masker and the signal for the three on-frequency masker conditions and the one off-frequency masker condition for HI listeners. Error bars indicate standard deviations across individual runs.

subject at all three signal frequencies. Note that this procedure resulted in some extrapolation of off-frequency data points to derive full input-output functions. Although caution has been recommended in using extrapolation of the off-frequency linear reference (Lopez-Poveda and Alves-Pinto, 2008), it was necessary in order to obtain derived input–output functions for the relatively high stimulus levels tested.

In addition, because of the high stimulus levels required, due to the threshold elevation from TEN, some listeners required off-frequency masker levels in excess of 92 dB SPL. As described in Wojtczak and Oxenham (2009), the slope of the off-frequency masking function may become shallower when levels in excess of 92 dB SPL are used, and this can lead to overestimates of compression. However, this appears to be more of an issue for NH than for HI listeners (Wojtczak and Oxenham, 2010).

The raw TMC data from Figs. 8 and 9 were used to derive the estimated input-output functions for the HI and NMNH listeners shown in Figs. 10 and 11, respectively. The slope of the linear regression of the input-output function was taken as the compression estimate. These compression estimates (along with the R^2 values for the linear regressions) are shown in Table IV. The dashed lines in Figs. 10 and 11 depict a linear response function with a slope of unity (consistent with no BM compression). Given the relatively small range of input levels that the response functions cover, it was decided not to use the more complex fitting routines (e.g., a third order polynomial) that have been used in previous studies (e.g., Plack *et al.*, 2004). In addition, since the measure of interest in this study involves a speech stimulus that is by nature broadband in both frequency and amplitude, it was decided to obtain an overall estimate of compression rather than seek a minimum compression estimate that may only cover a small range of input levels (e.g., Plack *et al.*, 2004). As can be seen from the proportion of variance accounted for (R^2), this simple fitting process generally describes the

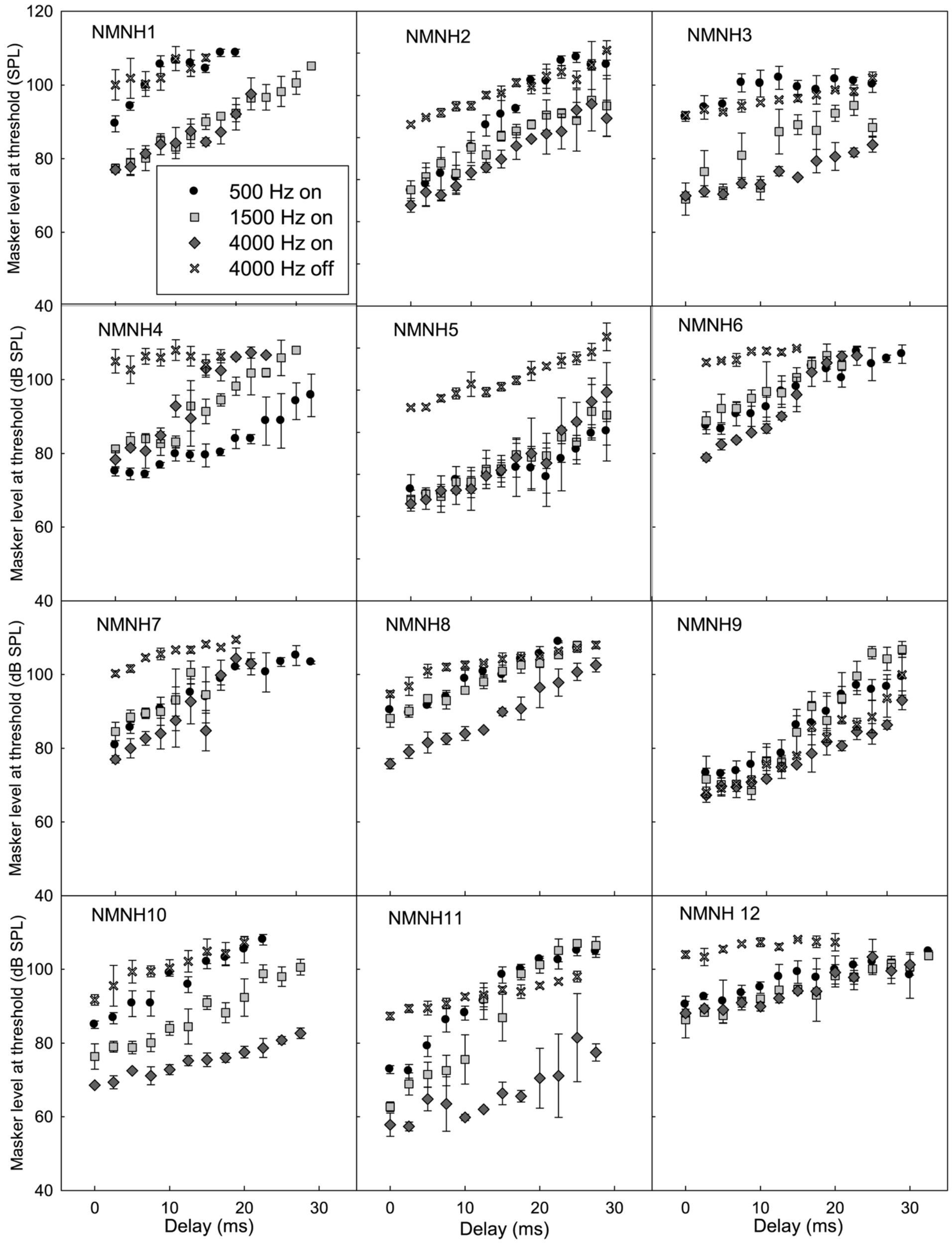

FIG. 9. As for Fig. 8, but for the NMNH listeners.

underlying functions well, with a few exceptions (i.e., NMNH3 and HI10 at 500 Hz, and HI3 at 4000 Hz).

Our estimates of cochlear compression were compared with absolute thresholds for the test signals. For a given HI listener, the estimated compression exponent at a particular signal frequency was plotted as a function of the quiet threshold at that same frequency. There was a slight trend towards less compression with increasing absolute threshold, but there was also a great deal of variability in the data, with some HI listeners with normal thresholds showing compression exponents of around 1.0 (linear) and other HI listeners with thresholds of 55 dB SPL showing compression exponents of around 0.3—close to standard estimates for NH listeners, which are typically around 0.2–0.3 (e.g., Oxenham and Plack, 1997; Nelson *et al.*, 2001). There was a significant relationship between compression estimate and absolute threshold only for the 500-Hz tones ($R^2 = 0.5271$; $p = 0.008$). None of the remaining relationships were significant. Clearly the quiet threshold is not a strong predictor of the underlying compression exponent, at least for this group of HI listeners. As expected, no significant relationships were found between compression exponents and absolute thresholds in the NMNH group. Surprisingly, paired t-tests showed that there were no significant differences in estimated compression slope between the HI and NMNH listeners at any of the test frequencies.

D. Discussion

The estimates of compression were quite variable for both NH and HI groups. At first glance, it is interesting to note that HI listeners with higher degrees of hearing loss (such as HI6) do not necessarily display linear input–output functions. In contrast, other HI listeners (such as HI2), with less audiometric hearing loss, display compression estimates consistent with linear processing for some test

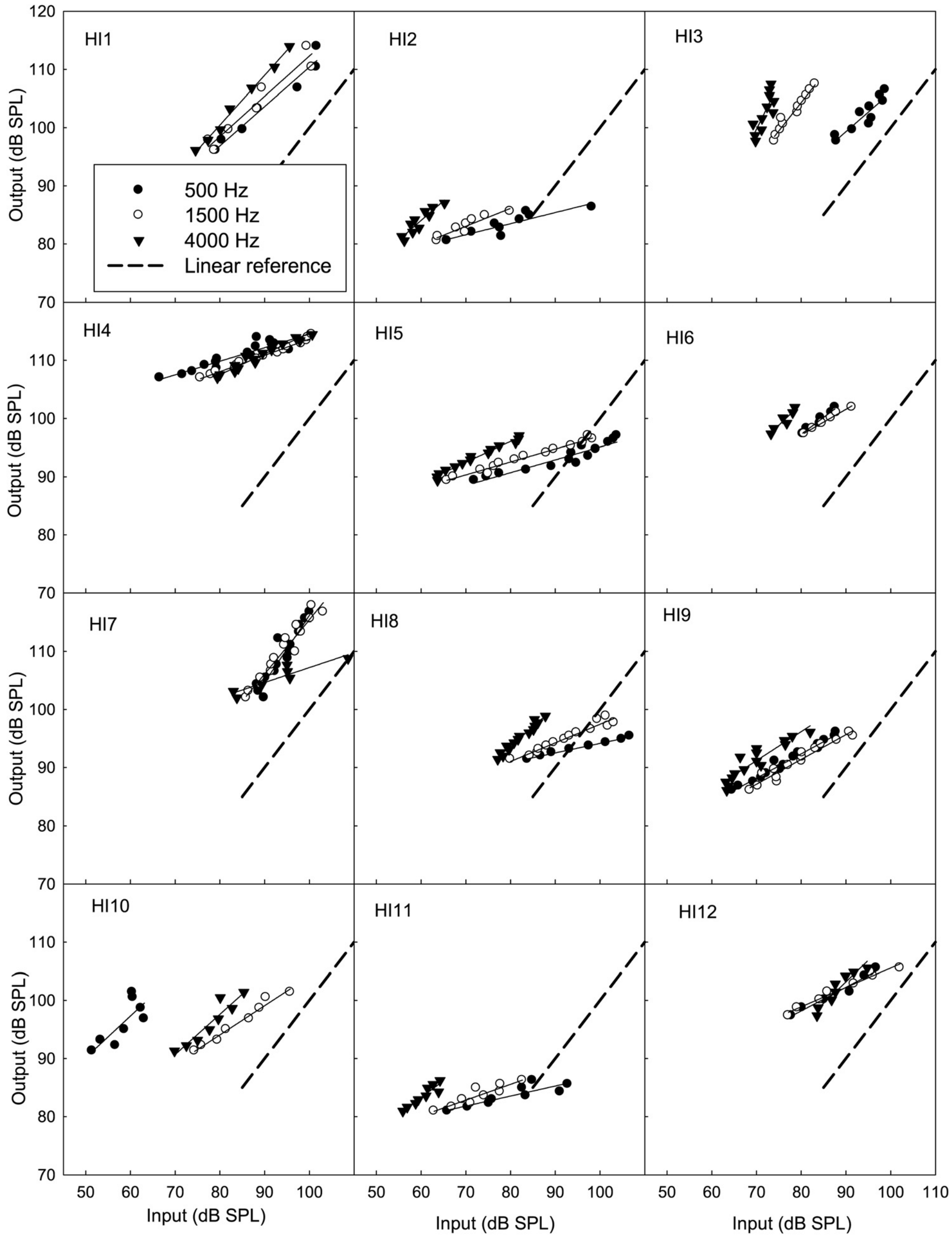

FIG. 10. Derived input-output functions from the TMC data in Fig. 8. Symbols indicate data points while solid lines indicate best linear fits to the data. The dashed line indicates a response with a slope of 1.0 (no compression).

signals. Previous studies have noted that estimates of maximum compression do not appear to be strongly correlated with the underlying audiometric thresholds (Plack *et al.*, 2004; Lopez-Poveda *et al.*, 2005; Lopez-Poveda and Johannesen, 2012), and the present results support that conclusion. Also interesting is the finding that some NMNH listeners, with presumably normal underlying cochlear function, show exponents consistent with little to no BM compression (see NMNH9). A similar finding has also been reported in a recent study by Poling *et al.* (2011). In their study, compression estimates were derived from TMC data for listeners with thresholds in the 0 to 20 dB HL range for a signal frequency of 1000 Hz. The resulting I/O curve slopes for these individuals with audiometrically normal hearing at the test frequency ranged from 0.083 dB/dB up to 1.75 dB/dB. The fact that Poling *et al.* (2011) showed similar variability in I/O curve slopes using the TMC method in quiet suggests that the variability in the present results are not due solely to our use of TEN to elevate thresholds. In addition, an earlier study using GOM to estimate cochlear compression in NH listeners found no effect of background noise on estimates of compression exponents (Gregan *et al.*, 2010).

According to the assumptions of the TMC method, the on-frequency TMC curves are thought to reflect the influence of BM compression of the masker, as well as temporal resolution (in terms of recovery from forward masking). In contrast, the reference off-frequency TMC curve is assumed to reflect only the recovery from forward masking and, therefore, it can be used as an estimate of temporal resolution. As stated previously, there have been conflicting views in the literature to date as to whether or not temporal resolution *per se* affects MR for speech in a temporally varying background. It seems plausible that listeners with slower recovery from forward masking (as indicated by a shallower slope for the off-frequency TMC function) may

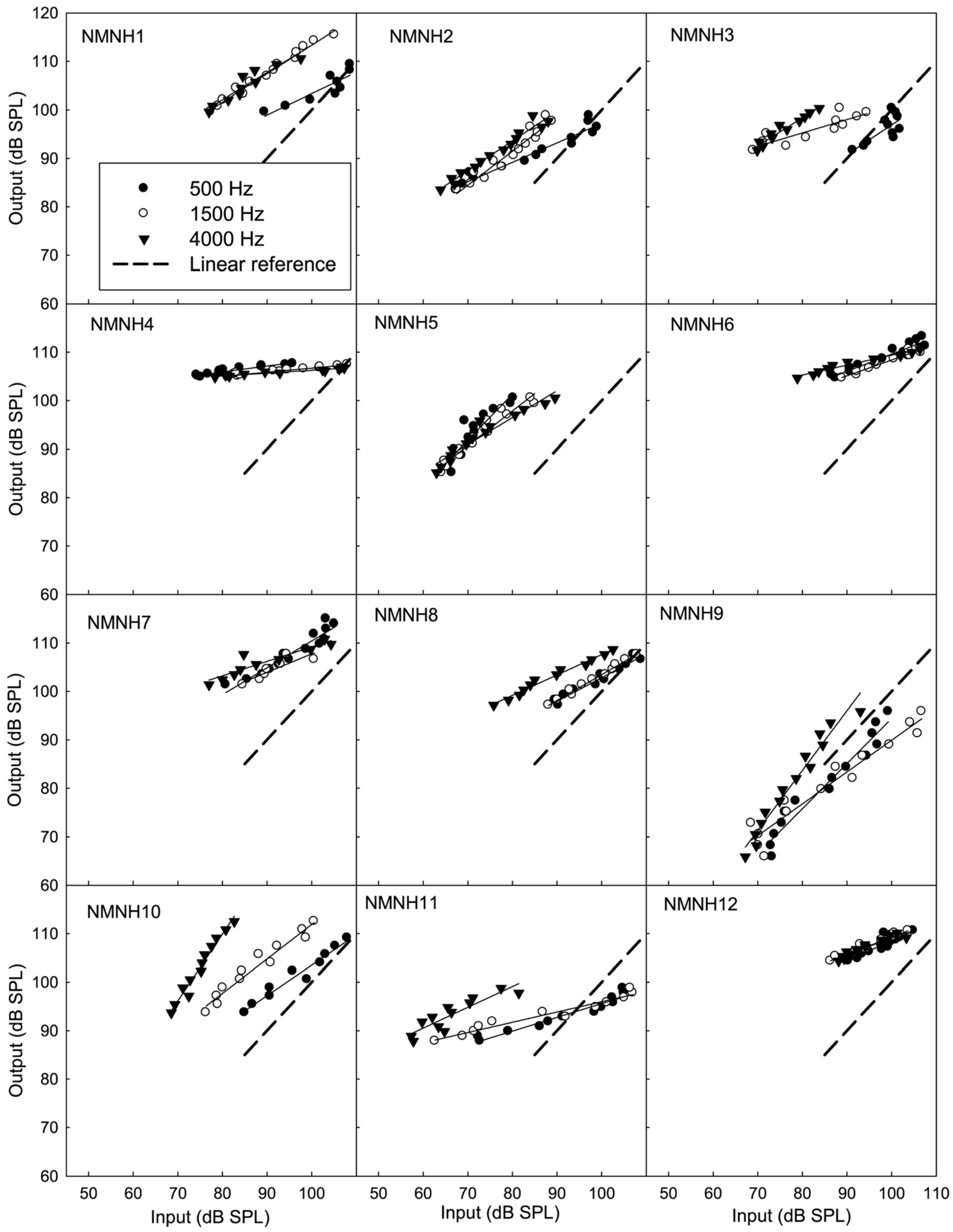

FIG. 11. As for Fig. 10, but for the NMNH listeners.

be less able to make use of brief temporal gaps in a masker. Linear fits were made to the off-frequency TMC curve of each listener (see Table IV). As can be seen from Table IV, there was considerable variability in the slopes, even within the NMNH group. However, a paired *t*-test showed that the slopes of the off-frequency TMC curves were significantly shallower for the HI than for the NMNH group [$t(11) = -2.78$; $p = 0.018$] and, consistent with earlier studies (e.g., Derleth *et al.*, 2001; Rhebergen *et al.*, 2006), the slope of the off-frequency function decreases as hearing loss increases ($R^2 = 0.2141$; $p = 0.0228$; Fig. 12). This is, however, inconsistent with the results of a previous study (Plack *et al.*, 2004). Because the ESII of Rhebergen *et al.* (2006) can include a forward-masking function, the difference in forward masking observed between the HI and NMNH groups may be used in the framework of the ESII to account for at least part of the difference in speech MR observed between the two groups.

V. RELATIONS BETWEEN TMC MEASURES AND MASKING RELEASE

A. Compression estimates and MR for speech

To test our hypothesis that underlying compression estimates are correlated with MR for speech, the summary measure of MR (percent-point improvement at -5 dB SNR in gated versus steady noise) was examined as a function of several compression summary values. Three summary measures of compression were selected: the average compression exponent across the three test frequencies (500, 1500, and 4000 Hz), the least compressive exponent of the three, and the most compressive exponent of the three. None of the three summary measures showed a significant linear relation with the measure of speech MR ($p > 0.05$ in all cases, even with no correction for multiple comparisons). This was true when considering the HI data alone, the NMNH data alone, or both groups combined. Thus, no significant relationship was observed between MR

TABLE IV. Columns 1–3: Straight-line fits (i.e., slope) to TMC derived input–output data for the three signal frequencies. These values represent the estimated cochlear compression exponent (dB/dB). Column 4: Straight-line fits to the off-frequency TMC at 4 kHz showing the growth of masker level, as a function of masker-signal gap (dB/ms). Numbers in parentheses are R^2 values indicating the goodness of the linear fits to the data.

	500 Hz (On)	1500 Hz (On)	4000 Hz (On)	4000 Hz (Off)
HI 1	0.68 (0.96)	0.69 (0.94)	0.85 (0.99)	0.33 (0.97)
HI 2	1.9 (0.78)	0.31 (0.89)	0.67 (0.85)	0.26 (0.51)
HI 3	0.67 (0.83)	0.97 (0.96)	1.57 (0.67)	0.39 (0.92)
HI 4	0.23 (0.86)	0.28 (0.98)	0.36 (0.95)	0.21 (0.96)
HI 5	0.22 (0.89)	0.22 (0.98)	0.36 (0.98)	0.24 (0.93)
HI 6	0.58 (0.96)	0.44 (0.99)	0.74 (0.89)	0.09 (0.73)
HI 7	1.1 (0.89)	0.93 (0.94)	0.26 (0.85)	0.45 (0.89)
HI 8	0.16 (0.99)	0.32 (0.95)	0.67 (0.96)	0.23 (0.51)
HI 9	0.42 (0.98)	0.43 (0.94)	0.50 (0.86)	0.29 (0.81)
HI 10	0.72 (0.62)	0.50 (0.97)	0.69 (0.89)	0.18 (0.76)
HI 11	0.17 (0.76)	0.27 (0.85)	0.57 (0.87)	0.28 (0.92)
HI 12	0.41 (0.97)	0.33 (0.97)	0.74 (0.90)	0.14 (0.81)
NMNH 1	0.44 (0.78)	0.59 (0.98)	0.58 (0.94)	0.49 (0.73)
NMNH 2	0.40 (0.94)	0.70 (0.94)	0.62 (0.97)	0.47 (0.96)
NMNH 3	0.61 (0.56)	0.28 (0.76)	0.56 (0.94)	0.35 (0.94)
NMNH 4	0.12 (0.92)	0.09 (0.94)	0.06 (0.93)	0.09 (0.10)
NMNH 5	0.95 (0.85)	0.72 (0.94)	0.56 (0.93)	0.47 (0.96)
NMNH 6	0.36 (0.91)	0.33 (0.94)	0.21 (0.96)	0.26 (0.86)
NMNH 7	0.56 (0.91)	0.38 (0.75)	0.30 (0.86)	0.42 (0.89)
NMNH 8	0.49 (0.95)	0.53 (0.99)	0.42 (0.98)	0.42 (0.90)
NMNH 9	0.93 (0.95)	0.66 (0.94)	1.23 (0.96)	0.92 (0.95)
NMNH 10	0.63 (0.96)	0.71 (0.95)	1.38 (0.97)	0.68 (0.93)
NMNH 11	0.29 (0.94)	0.21 (0.95)	0.43 (0.81)	0.40 (0.98)
NMNH 12	0.43 (0.83)	0.36 (0.96)	0.34 (0.90)	0.20 (0.70)

for speech and compression as estimated by the TMC method.

This outcome does not support our original hypothesis, that peripheral compression increases MR and improves the overall speech-to-noise ratio by amplifying low-level speech in the temporal valleys of the masker. Such effects of compression have also been predicted using broadband dynamic-range compression, such as that found in some hearing-aid algorithms (e.g., Rhebergen *et al.*, 2009). It is important to note, however, that the lack of a significant correlation does not completely rule out the potential effects of compression. First, as discussed above, it remains unclear how accurate or reliable a measure of compression the TMC method provides. Both our results and those of Poling *et al.* (2011) suggest much more variable estimates of compression than would be expected among subjects with normal hearing or mild hearing loss. Thus, our negative result may be more a reflection of the reliability of the TMC measure than evidence against the importance of compression. Second, the HI group was relatively heterogeneous and, although large when compared to many previous studies, still represents a relatively small group of listeners. Although it is possible that a much larger group would have yielded a significant result, the fact that the relationship did not even show a trend towards significance, despite multiple measures and comparisons, makes this outcome unlikely.

Another possibility is that the sound levels were too high for compression to be effective, even in the NH listeners. It is thought that the BM input-output function becomes more linear again at high sound input levels, for pure tones exceeding about 80 dB SPL (e.g., Ruggero, 1992; Oxenham and Plack, 1997). The sound level of the (broadband) speech ranged from 70 to 95 dB SPL, resulting in average levels per third-octave band of roughly 60 to 85 dB SPL. Thus, based on our current understanding, some effects of compression should have been measurable; indeed, most of the NH group did show slope estimates that were compressive. As mentioned before, Plack *et al.* (2004) found no correlation between the maximum compression value measured in individual subjects and their degree of hearing loss; instead, the range over which compression was found was reduced. Our method, of estimating the average compression over the measurable range should have resulted in an estimate that was sensitive not only to the maximum amount of compression, but also the range over which compression was maximal.

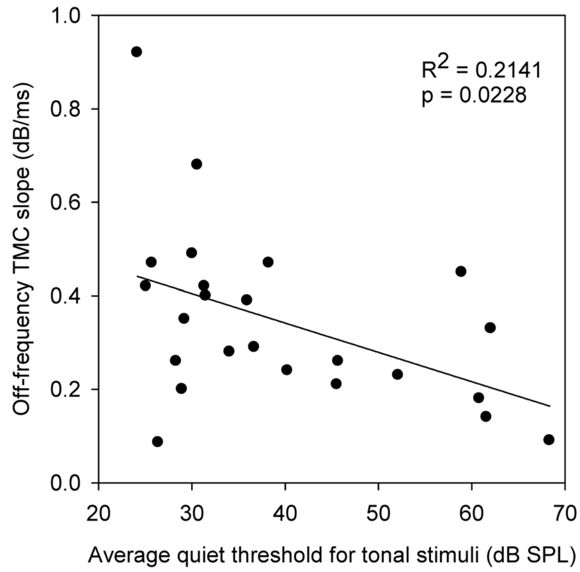

FIG. 12. Slope of the off-frequency TMC function for both HI and NMNH listeners as a function of the average threshold for the tonal stimuli in quiet.

Although we cannot rule out the importance of peripheral compression in determining speech MR, our negative results do suggest that it will not be possible to predict the performance of individual HI listeners in complex, fluctuating backgrounds, based on the current measures of peripheral compression, at least at levels that are relevant for listeners with mild-to-moderate hearing loss.

B. Compression estimates and MRPT

To determine if the degree of MRPT varied systematically with estimates of compression at the same frequency, MRPT was examined as a function of the estimated compression for each of the signal frequencies. As with speech MR, there was no significant relationship for any of these comparisons for either group or for the two groups combined. Therefore, for the present results, it does not appear that the degree of MRPT can be predicted based on the underlying compression estimates from the TMC method. This outcome is inconsistent with results from Oxenham and Dau (2004), who found a weak but significant correlation

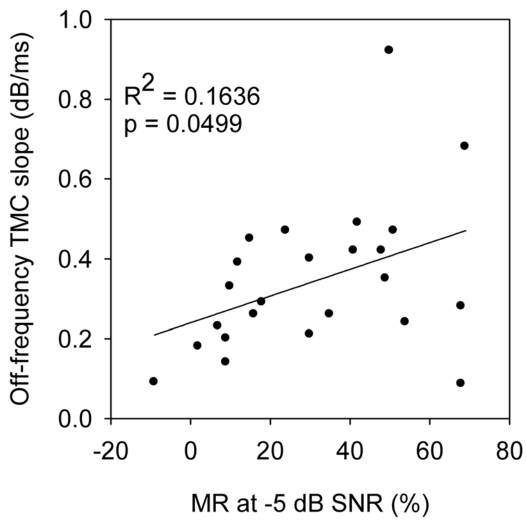

FIG. 13. Slope of the off-frequency TMC function as a function of speech MR for both HI and NMNH listeners. The MR is defined in terms of improvement in speech understanding (percentage points) at −5 dB SNR.

between the degree of masking difference (between flat and modulated Schroeder-phase maskers) and auditory filter bandwidth, which was used as a proxy measure of BM compression, based on the strong correlations reported by a previous study (Moore *et al.*, 1999a). Other than the large variability associated with the TMC estimates of compression in the present study, it is not clear what accounts for this difference.

C. Relationship of temporal resolution to MR

To examine the relationship between temporal resolution and speech MR, the slopes of the off-frequency TMC curves are plotted as a function of speech MR in Fig. 13. The correlation, using data from both groups, just reached statistical significance (Pearson product-moment correlation; $R^2 = 0.1636$; $p = 0.0499$), with the trend in the expected direction of better temporal resolution being associated with more masking release. However, the proportion of variance accounted for is very small.

VI. SUMMARY AND CONCLUSIONS

Masking release in speech was measured by comparing word identification in sentences in steady-state noise with that in square-wave-gated noise. Age effects were controlled by using pairs of HI and NMNH listeners who were similar in age; audibility and overall level effects were controlled by embedding the test stimuli in a background noise (TEN) that equated audibility for each pair of listeners. Cochlear compression and temporal resolution were estimated using the TMC method, and pure-tone MR was estimated using the same masker as was used with the speech measures. The findings can be summarized as follows:

(1) The HI listeners showed less speech MR than their NH counterparts in 10 of the 12 pairs of subjects, despite matching age, stimulus level, and stimulus audibility.
(2) Pure-tone MR was significantly correlated with speech MR, suggesting a common underlying mechanism. However, as pure-tone MR accounted for only 23%–33% of the variance of speech MR, other factors may also play a role.
(3) Overall, NMNH listeners had significantly steeper off-frequency forward-masking recovery slopes than did the HI listeners. The trend for a correlation between this measure of temporal resolution and speech MR requires further verification.
(4) Most importantly, estimates of peripheral compression, obtained with the TMC method, were not correlated with either speech or pure-tone MR. This outcome does not support the initial hypothesis that a reduction in peripheral compression underlies the reduction in MR observed with hearing loss. However, such a relationship cannot be completely ruled out, as the outcome may be due in part to the highly variable estimates of compression derived from the TMC method, even in the NH group, as well as the relatively small sample size and heterogeneity of the HI group.

ACKNOWLEDGMENTS

This work was supported by NIH grants R01 DC03909 (AJO) and R01 DC008306 (PBN). The associate editor, Enrique Lopez-Poveda, provided helpful comments on an earlier version of this paper.

ANSI (**1997**). *Methods for Calculation of the Speech Intelligibility Index* (American National Standards Institute, New York).

Bacon, S. P., Opie, J. M., and Montoya, D. Y. (**1998**). "The effects of hearing loss and noise masking on the masking release for speech in temporally complex backgrounds," J. Speech Lang. Hear. Res. **41**, 549–563.

Bernstein, J. G., and Brungart, D. S. (**2011**). "Effects of spectral smearing and temporal fine-structure distortion on the fluctuating-masker benefit for speech at a fixed signal-to-noise ratio," J. Acoust. Soc. Am. **130**, 473–488.

Bernstein, J. G., and Grant, K. W. (**2009**). "Auditory and auditory-visual intelligibility of speech in fluctuating maskers for normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am. **125**, 3358–3372.

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

Buus, S. (**1985**). "Release from masking caused by envelope fluctuations," J. Acoust. Soc. Am. **78**, 1958–1965.

Cooper, N. P., and Rhode, W. S. (**1995**). "Nonlinear mechanics at the apex of the guinea-pig cochlea," Hear. Res. **82**, 225–243.

Cooper, N. P., and Yates, G. K. (**1994**). "Nonlinear input-output functions derived from the responses of guinea-pig cochlear nerve fibres: Variations with characteristic frequency," Hear. Res. **78**, 221–234.

Derleth, R. P., Dau, T., and Kollmeier, B. (**2001**). "Modeling temporal and compressive properties of the normal and impaired auditory system," Hear. Res. **159**, 132–149.

Desloge, J. G., Reed, C. M., Braida, L. D., Perez, Z. D., and Delhorne, L. A. (**2010**). "Speech reception by listeners with real and simulated hearing impairment: effects of continuous and interrupted noise," J. Acoust. Soc. Am. **128**, 342–359.

Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (**2002**). "Benefit of modulated maskers for speech recognition by younger and older adults with normal hearing," J. Acoust. Soc. Am. **111**, 2897–2907.

Dubno, J. R., Horwitz, A. R., and Ahlstrom, J. B. (**2003**). "Recovery from prior stimulation: masking of speech by interrupted noise for younger and older adults with normal hearing," J. Acoust. Soc. Am. **113**, 2084–2094.

Duquesnoy, A. J. (**1983**). "Effect of a single interfering noise or speech source on the binaural sentence intelligibility of aged persons," J. Acoust. Soc. Am. **74**, 739–743.

Egan, J. P., and Hake, H. W. (**1950**). "On the masking pattern of a simple auditory stimulus," J. Acoust. Soc. Am. **22**, 622–630.

Eisenberg, L. S., Dirks, D. D., and Bell, T. S. (**1995**). "Speech recognition in amplitude-modulated noise of listeners with normal and listeners with impaired hearing," J. Speech Hear. Res. **38**, 222–233.

Festen, J. M., and Plomp, R. (**1990**). "Effects of fluctuating noise and interfering speech on the speech-reception threshold for impaired and normal hearing," J. Acoust. Soc. Am. **88**, 1725–1736.

George, E. L., 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. (**1994**). "Growth-of-masking functions for several types of maskers," J. Acoust. Soc. Am. **96**, 134–144.

Gregan, M. J., Nelson, P. B., and Oxenham, A. J. (**2010**). "Effects of background noise level on behavioral estimates of basilar-membrane compression," J. Acoust. Soc. Am. **127**, 3018–3025.

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

IEEE (**1969**). "IEEE recommended practice for speech quality measurements," IEEE Trans. Audio Electroacoust. **17**, 225–246.

Jin, S. H., and Nelson, P. B. (**2006**). "Speech perception in gated noise: The effects of temporal resolution," J. Acoust. Soc. Am. **119**, 3097–3108.

Jin, S. H., and Nelson, P. B. (**2010**). "Interrupted speech perception: the effects of hearing sensitivity and frequency resolution," J. Acoust. Soc. Am. **128**, 881–889.

Jorgensen, S., and Dau, T. (**2011**). "Predicting speech intelligibility based on the signal-to-noise envelope power ratio after modulation-frequency selective processing," J. Acoust. Soc. Am. **130**, 1475–1487.

Jürgens, T., Kollmeier, B., Brand, T., and Ewert, S. D. (**2011**). "Assessment of auditory nonlinearity for listeners with different hearing losses using temporal masking and categorical loudness scaling," Hear Res. **280**, 177–191.

Kohlrausch, A., and Sander, A. (**1995**). "Phase effects in masking related to dispersion in the inner ear. II. Masking period patterns of short targets," J. Acoust. Soc. Am. **97**, 1817–1829.

Kwon, B. J., and Turner, C. W. (**2001**). "Consonant identification under maskers with sinusoidal modulation: Masking release or modulation interference?," J. Acoust. Soc. Am. **110**, 1130–1140.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am. **49**, 467–477.

Lopez-Poveda, E. A., and Alves-Pinto, A. (**2008**). "A variant temporal-masking-curve method for inferring peripheral auditory compression," J. Acoust. Soc. Am. **123**, 1544–1554.

Lopez-Poveda, E. A., and Johannesen, P. T. (**2009**). "Otoacoustic emission theories and behavioral estimates of human basilar membrane motion are mutually consistent," J. Assoc. Res. Otolaryngol. **10**, 511–523.

Lopez-Poveda, E. A., and Johannesen, P. T. (**2012**). "Behavioral estimates of the contribution of inner and outer hair cell dysfunction to individualized audiometric loss," J. Assoc. Res. Otolaryngol. **13**, 485–504.

Lopez-Poveda, E. A., Plack, C. J., and Meddis, R. (**2003**). "Cochlear nonlinearity between 500 and 8000 Hz in listeners with normal hearing," J. Acoust. Soc. Am. **113**, 951–960.

Lopez-Poveda, E. A., Plack, C. J., Meddis, R., and Blanco, J. L. (**2005**). "Cochlear compression in listeners with moderate sensorineural hearing loss," Hear Res. **205**, 172–183.

Lorenzi, C., Husson, M., Ardoint, M., and Debruille, X. (**2006**). "Speech masking release in listeners with flat hearing loss: effects of masker fluctuation rate on identification scores and phonetic feature reception," Int. J. Audiol. **45**, 487–495.

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

Moore, B. C. J., Huss, M., Vickers, D. A., Glasberg, B. R., and Alcantara, J. I. (**2000**). "A test for the diagnosis of dead regions in the cochlea," Br. J. Audiol. **34**, 205–224.

Moore, B. C. J., Vickers, D. A., Plack, C. J., and Oxenham, A. J. (**1999a**). "Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism," J. Acoust. Soc. Am. **106**, 2761–2778.

Moore, B. C. J., Peters, R. W., and Stone, M. A. (**1999b**). "Benefits of linear amplification and multichannel compression for speech comprehension in backgrounds with spectral and temporal dips," J. Acoust. Soc. Am. **105**, 400–411.

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.

Nelson, D. A., and Swain, A. C. (**1996**). "Temporal resolution within the upper accessory excitation of a masker," Acustica **82**, 328–334.

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

Nelson, P. B., Jin, S. H., Carney, A. E., and Nelson, D. A. (**2003**). "Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners," J. Acoust. Soc. Am. **113**, 961–968.

Oxenham, A. J., and Dau, T. (**2001**). "Reconciling frequency selectivity and phase effects in masking," J. Acoust. Soc. Am. **110**, 1525–1538.

Oxenham, A. J., and Dau, T. (**2004**). "Masker phase effects in normal-hearing and hearing-impaired listeners: Evidence for peripheral compression at low signal frequencies," J. Acoust. Soc. Am. **116**, 2248–2257.

Oxenham, A. J., and Moore, B. C. J. (**1997**). "Modeling the effects of peripheral nonlinearity in normal and impaired hearing," in *Modeling Sensorineural Hearing Loss*, edited by W. Jesteadt (Erlbaum, Hillsdale, NJ), pp. 273–288.

Oxenham, A. J., and Plack, C. J. (**1997**). "A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **101**, 3666–3675.

Oxenham, A. J., and Simonson, A. M. (**2009**). "Masking release for low- and high-pass-filtered speech in the presence of noise and single-talker interference," J. Acoust. Soc. Am. **125**, 457–468.

Peters, R. W., Moore, B. C. J., and Baer, T. (**1998**). "Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people," J. Acoust. Soc. Am. **103**, 577–587.

Plack, C. J., and Arifianto, D. (**2010**). "On- and off-frequency compression estimated using a new version of the additivity of forward masking technique," J. Acoust. Soc. Am. **128**, 771–786.

Plack, C. J., and Drga, V. (**2003**). "Psychophysical evidence for auditory compression at low characteristic frequencies," J. Acoust. Soc. Am. **113**, 1574–1586.

Plack, C. J., Drga, V., and Lopez-Poveda, E. A. (**2004**). "Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss," J. Acoust. Soc. Am. **115**, 1684–1695.

Poling, G. L., Horwitz, A. R., Ahlstrom, J. B., and Dubno, J. R. (**2011**). "Individual differences in behavioral estimates of cochlear nonlinearities," J. Assoc. Res. Otolaryngol. **13**, 91–108.

Rhebergen, K. S., and Versfeld, N. J. (**2005**). "A Speech Intelligibility Index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners," J. Acoust. Soc. Am. **117**, 2181–2192.

Rhebergen, K. S., Versfeld, N. J., and Dreschler, W. A. (**2006**). "Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise," J. Acoust. Soc. Am. **120**, 3988–3997.

Rhebergen, K. S., Versfeld, N. J., and Dreschler, W. A. (**2009**). "The dynamic range of speech, compression, and its effect on the speech reception threshold in stationary and interrupted noise," J. Acoust. Soc. Am. **126**, 3236–3245.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev. **81**, 1305–1352.

Rosengard, P. S., Oxenham, A. J., and Braida, L. D. (**2005**). "Comparing different estimates of cochlear compression in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **117**, 3028–3041.

Ruggero, M. A. (**1992**). "Responses to sound of the basilar membrane of the mammalian cochlea," Curr. Opin. Neurobiol. **2**, 449–456.

Ruggero, M. A., and Rich, N. C. (**1991**). "Furosemide alters organ of Corti mechanics: Evidence for feedback of outer hair cells upon the basilar membrane," J. Neurosci. **11**, 1057–1067.

Scharf, B. (**1964**). "Partial masking," Acustica **14**, 16–23.

Schroeder, M. R. (**1970**). "Synthesis of low peak-factor signals and binary sequences with low autocorrelation," IEEE Trans. Inf. Theory **16**, 85–89.

Shailer, M. J., and Moore, B. C. J. (**1987**). "Gap detection and the auditory filter: Phase effects using sinusoidal stimuli," J. Acoust. Soc. Am. **81**, 1110–1117.

Summers, V., and Molis, M. R. (**2004**). "Speech recognition in fluctuating and continuous maskers: Effects of hearing loss and presentation level," J. Speech Lang. Hear. Res. **47**, 245–256.

Takahashi, G. A., and Bacon, S. P. (**1992**). "Modulation detection, modulation masking, and speech understanding in noise in the elderly," J. Speech Hear. Res. **35**, 1410–1421.

Wagener, K. C., Brand, T., and Kollmeier, B. (**2006**). "The role of silent intervals for sentence intelligibility in fluctuating noise in hearing-impaired listeners," Int. J. Audiol. **45**, 26–33.

Wojtczak, M., and Oxenham, A. J. (**2009**). "Pitfalls in behavioral estimates of basilar-membrane compression in humans," J. Acoust. Soc. Am. **125**, 270–281.

Wojtczak, M., and Oxenham, A. J. (**2010**). "Recovery from on- and off-frequency forward masking in listeners with normal and impaired hearing," J. Acoust. Soc. Am. **128**, 247–256.

Wojtczak, M., Schroder, A. C., Kong, Y. Y., and Nelson, D. A. (**2001**). "The effect of basilar-membrane nonlinearity on the shapes of masking period patterns in normal and impaired hearing," J. Acoust. Soc. Am. **109**, 1571–1586.

Zheng, J., Shen, W., He, D. Z., Long, K. B., Madison, L. D., and Dallos, P. (**2000**). "Prestin is the motor protein of cochlear outer hair cells," Nature **405**, 149–155.

Zwicker, E. B. (**1976**). "Masking period patterns of harmonic complex tones," J. Acoust. Soc. Am. **60**, 429–439.