Behavioral estimates of basilar-membrane compression: Additivity of forward masking in noise-masked normal-hearing 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 9 March 2011; revised 31 August 2011; accepted 1 September 2011)

Cochlear hearing loss is often associated with a loss of basilar-membrane (BM) compression, which in turn may contribute to degraded processing of suprathreshold stimuli. Behavioral estimates of compression may therefore be useful as long as they are valid over a wide range of levels and frequencies. Additivity of forward masking (AFM) may provide such a measure, but research to date lacks normative data from normal-hearing (NH) listeners at high sound levels, which is necessary to evaluate data from hearing-impaired (HI) listeners. The present study measured AFM in six NH listeners for signal frequencies of 500, 1500, and 4000 Hz in the presence of background noise, designed to elevate signal thresholds to levels similar to those experienced by HI listeners. Results consistent with compressive BM responses were found for all six listeners at 500 Hz, five listeners at 1500 Hz, but only two listeners at 4000 Hz. Further measurements in the absence of background noise also indicated a lack of consistent compression at 4000 Hz at higher signal levels, in contrast to earlier results collected at lower levels. A better understanding of this issue will be required before AFM can be used as a general behavioral estimate of BM compression.

© 2011 Acoustical Society of America. [DOI: 10.1121/1.3643817]

PACS number(s): 43.66.Ba, 43.66.Dc [CJP]

Pages: 2835–2844

I. INTRODUCTION

The nonlinear gain that is found in measurements of the cochlea's basilar membrane (BM) results in an input–output function that is highly compressive, with compression ratios as high as 5:1 (Rhode, 1971; Sellick *et al.*, 1982; Ruggero, 1992). It is believed that a loss of nonlinear cochlear gain, presumably resulting from reduced function of the outer hair cells (e.g., Ruggero *et al.*, 1997; Robles and Ruggero, 2001), can contribute to sensorineural hearing loss in humans, leading to higher absolute thresholds and a decreased dynamic range (Oxenham and Bacon, 2003). Further, the loss of cochlear compression may affect the processing of suprathreshold stimuli, such as speech in temporally and/or spectrally varying background noise (Nelson and Jin, 2004). Behavioral estimates of cochlear compression in humans have suggested that moderate-to-severe hearing loss is associated with a reduction or elimination of compression (Oxenham and Moore, 1995; Oxenham and Plack, 1997; Lopez-Poveda *et al.*, 2005). For listeners with milder hearing losses, Plack *et al.* (2004) found that the range of levels over which compression occurred decreased with increasing hearing loss, whereas the maximum compression ratio did not. In addition, some forms of hearing loss might not be expected to affect BM compression at all. For instance, disorders involving only the inner hair cells and/or the auditory nerve would

not necessarily affect the outer hair cells or BM mechanics, and so would not be expected to affect the BM input–output function. A reliable and efficient measure of BM compression may therefore provide a useful clinical diagnostic tool, if it can be applied over a sufficiently wide range of levels and frequencies.

Many studies comparing normal-hearing (NH) with hearing-impaired (HI) listeners have added a background masking noise to the stimuli presented to NH listeners, in order to simulate the decreased audibility experienced by HI listeners, and in order to present stimuli at both the same physical level (SPL) and the same sensation level (SL). This noise may be used to elevate thresholds in a particular frequency region (e.g., Sommers and Humes, 1993) or over a wider range of frequencies (e.g., Dubno and Schaefer, 1992; Bacon *et al.*, 1998). The noise may also be modified in terms of spectral shape and amplitude to specifically match the thresholds of individual HI listeners (e.g., Dubno and Schaefer, 1992) or it may be adjusted to approximate the average thresholds of a group of HI listeners (e.g., Humes and Jesteadt, 1989). Some studies have shown that psychoacoustic and speech perception results from HI listeners and noise-masked normal-hearing (NMNH) listeners have been comparable, as in the case of loudness recruitment (e.g., Zurek and Delhorne, 1987; Dubno and Schaefer, 1992; Hellman and Meiselman, 1993; Hall and Grose, 1997). In other studies, some HI listeners exhibit results similar to those found for NMNH listeners, whereas other HI listeners perform more poorly, such as in tasks involving

^{a)}Author to whom correspondence should be addressed. Electronic mail: grega005@umn.edu

speech understanding or tone detection in a modulated background (e.g., Bacon *et al.*, 1998; Oxenham and Dau, 2004). Reed *et al.* (2009) provide a thorough review of studies involving comparisons between HI and NMNH listeners.

One potential problem of using NMNH listeners is that the addition of the background noise may have effects beyond simply elevating thresholds. For instance, it is possible that the noise may partially linearize the response of the BM (Recio-Spinoso and Lopez-Poveda, 2010). To test this concern, Gregan *et al.* (2010) used the growth-of-masking (GOM) technique (Oxenham and Plack, 1997) and found that the addition of background noise did not appear to affect estimates of BM compression. Unfortunately, both the GOM method and an alternative method, known as the temporal masking curve (TMC) method (Nelson *et al.*, 2001), do not provide a straightforward way to estimate BM compression over a wide range of frequencies. Both GOM and TMC methods require the use of an off-frequency masker as a “linear reference,” with the assumption that the response to the much lower off-frequency masker will not be compressed at the basilar membrane location tuned to the signal frequency. This assumption is supported by physiological studies of BM mechanics in the base of the cochlea (e.g., Ruggero *et al.*, 1997); however, the mechanics of the low-frequency apical portion of the cochlea are not well understood, and existing data suggest that compression may not be as frequency specific as at the base of the cochlea. Thus, in the apex the response to maskers well below the signal frequency may still be compressive, thereby violating the assumption of the linear reference in the GOM and TMC methods (e.g., Rhode and Cooper, 1996; Zinn *et al.*, 2000; Plack and Arifanto, 2010). The TMC method allows the use of a high-frequency linear reference in an attempt to circumvent this issue (e.g., Lopez-Poveda *et al.*, 2003), but this assumes that the linear reference curve, obtained at a relatively high frequency, is appropriate as a comparison for low-frequency signals. Recent studies have suggested that this may not be a valid assumption, due to the possibility that decay of forward masking may vary with signal frequency (e.g., Stainsby and Moore, 2006) and/or with masker frequency (Wojtczak and Oxenham, 2009).

Additivity of nonsimultaneous masking is another method that has been used to estimate cochlear compression (Oxenham and Moore, 1995; Plack and O’Hanlon, 2003; Plack *et al.*, 2008; Plack and Arifanto, 2010). This technique has possible advantages over GOM and TMC methods in that it does not require a comparison of on- and off-frequency responses, and so allows estimates of cochlear compression to be obtained across a wider frequency range than the GOM or TMC techniques. One assumption underlying the use of additivity of masking to estimate compression is that the effects of nonoverlapping nonsimultaneous maskers combine linearly (Penner, 1980; Plack *et al.*, 2006). As a result, the combining of two equally effective maskers should result in a doubling of masking. If the signal is processed linearly (with respect to intensity) the doubling of masking should lead to a 3-dB increase in masked threshold. On the other hand, if the signal is compressed, the signal level will need to be increased by

more than 3 dB in the presence of the two maskers. This so-called “excess masking” has been observed in many studies with NH listeners using either two forward maskers and/or a combination of forward and backward maskers (e.g., Wilson and Carhart, 1971; Penner, 1980; Cokely and Humes, 1993; Oxenham and Moore, 1995; Plack and O’Hanlon, 2003; Plack *et al.*, 2008; Plack and Arifanto, 2010). In HI listeners, studies have tended to show less excess masking, consistent with the idea of reduced peripheral compression (e.g., Oxenham and Moore, 1995; Plack *et al.*, 2008).

Although masking additivity has been used in a number of studies, the frequencies tested have tended to be either very low [250 or 500 Hz (Plack and O’Hanlon, 2003; Plack *et al.*, 2008)] or high [4000 or 6000 Hz (Oxenham and Moore, 1995; Plack and O’Hanlon, 2003; Plack *et al.*, 2008)]. Also, the range of levels tested has not been very extensive, particularly regarding comparisons of NH and HI listeners at similar SLs and SPLs. For instance, Oxenham and Moore (1995) compared NH and HI listeners at the same SL, but not at the same SPL. To our knowledge, no studies of masking additivity have been carried out in a background noise to equalize SPL and SL when comparing NH and HI listeners. As mentioned previously, Gregan *et al.* (2010) found that compression estimates were not affected by the presence of a background noise when using the GOM method, suggesting that similar results should also be observed when using methods of masking additivity; however, this has not yet been tested.

The aim of this study was to estimate compression in NH listeners using additivity of forward masking (AFM) at low, medium, and high frequencies relevant for speech perception (500, 1500, and 4000 Hz), and at high stimulus levels (signal and masker) in the presence of background noise to enable a future comparison with data from HI listeners.

II. EXPERIMENT 1

A. Listeners

Six adult listeners (all female) participated. The listeners had audiometric thresholds of 20 dB HL or better at octave test frequencies, as defined by ANSI S3.6-2004 (ANSI, 2004) and their ages ranged from 19 to 37. Five listeners were paid an hourly rate for their participation; S2 was the first author.

B. Procedure

Masked thresholds for pure-tone signals of 500, 1500, and 4000 Hz were measured in the presence of two contiguous Gaussian noise forward maskers, M1 and M2, presented either together or singly. The levels of the two maskers were set so as to produce approximately the same amount of masking of the signal when each masker was presented alone. This was accomplished in Phase 1 of the experiment by presenting each masker (M1 or M2) alone with the signal and adaptively varying the masker level until it just masked a fixed-level signal for several signal levels [see Fig. 1(A)]. The signal levels were based on each listener’s thresholds for the signal stimuli in a background noise [threshold

FIG. 1. Depiction of stimuli for the two measurement phases of the additivity task. (a) Phase 1, where the level of each masker (M1 or M2) in isolation was adapted to just mask the fixed signal level. Note that the temporal relationships between the stimuli were maintained even when only one masker was presented. (b) Phase 2, where both M1 and M2 are presented contiguously (at the levels determined in Phase 1) and the signal level is varied adaptively.

equalizing noise, TEN (Moore *et al.*, 2000)]. The average thresholds for the brief signals with frequencies of 500, 1500, and 4000 Hz were 58, 59.3, and 57.5 dB SPL, respectively. The standard deviation across listeners was less than 2 dB. The signal levels are specified in dB SL re: these thresholds in TEN. The resulting M1 and M2 levels for two NH listeners (S2 and S3) were averaged and these levels were used in Phase 2 for all listeners. Given that all listeners were NH and listening in the same TEN, it was decided to

measure Phase 1 for only these two listeners. As described in Plack and O'Hanlon (2003), as long as these masker levels result in relatively equivalent, not necessarily identical, masked thresholds in Phase 2, the underlying compression exponent can be calculated. Therefore, it was decided the experiment could be shortened without compromising data by only completing Phase 1 for two listeners. The resulting averaged Phase 1 M1 and M2 spectrum level values for each signal frequency are shown in Table I.

In Phase 2, M1 and M2 were presented together or alone at the levels determined in Phase 1, and the signal level was varied adaptively until it was just masked [see Fig. 1(B)]. Thresholds were measured using a three-interval three-alternative forced-choice method, with a two-down one-up (or two-up one-down, when masker level was varied) adaptive rule that tracks 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 masking stimulus (either M1 alone, M2 alone, or M1 and M2 presented contiguously) was presented in all three intervals. The listener's task was to select the interval that contained the signal, which was presented in one interval chosen at random with uniform probability. The initial step size for adaptively varying the masker or signal level was 4 dB. The step size was then 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 level at the last six reversals. Subjects completed three practice runs per condition and the threshold for each condition and each subject was taken as the average of the three subsequent runs. For each subject, any threshold estimate from a single run that was not within 4 standard deviations of the mean of the remaining thresholds for that subject and condition was considered an outlier and was excluded from the analysis. In that case, an additional data point was measured. Less than 10 threshold estimates fell into this category (over all conditions and listeners). The presentation order of conditions was randomized within and across subjects and repetitions. Listening sessions were 2 h in length, including frequent breaks. The entire experiment took an average of five sessions per listener to complete.

C. Stimuli

The stimuli were generated digitally and converted to an analog signal via a 24-bit Lynx22 (LynxStudio) soundcard at a sampling rate of 48 kHz. Sounds were presented to the left

TABLE I. Spectrum levels of M1 and M2 (dB SPL), averaged across S2 and S3, as derived in Phase 1. Values in parentheses are standard deviations.

Frequency (Hz)	Masker	5 SL	10 SL	15 SL	20 SL	25 SL
500	M1	22.1 (0.9)	33.5 (0.04)	38.9 (4.6)	48.1 (0.5)	54.8 (4.8)
	M2	28.6 (2.5)	36.3 (8.4)	46.0 (11.6)	51.1 (2.8)	53.9 (8.3)
1500	M1	22.2 (6.9)	28.9 (2.4)	32.9 (0.04)	41.8 (3.8)	45.3 (4.3)
	M2	27.0 (0.7)	31.5 (5.7)	37.0 (11.1)	43.6 (9.9)	44.4 (7.5)
4000	M1	17.1 (4.4)	26.5 (3.7)	34.0 (2.8)	37.9 (1.8)	41.4 (1.3)
	M2	24.0 (3.6)	29.6 (8.4)	32.2 (6.6)	37.0 (6.6)	39.7 (10.8)

ear of subjects via Sennheiser HD580 headphones. The subjects were tested in a double-walled sound-attenuating booth.

The signal was a brief pure tone with a frequency of 500, 1500, or 4000 Hz. The total duration of the 500 Hz signal was 10 ms, including 5 ms raised-cosine onset and offset ramps (no steady state). The total duration for the 1500 and 4000 Hz signals was 6 ms, including 2 ms raised-cosine onset and offset ramps. For the 1500 and 4000 Hz signals, M1 and M2 were both one-octave-wide bands of noise, geometrically centered around the signal frequency. For the 500 Hz signal, the masker bandwidth extended from 20 to 1000 Hz. The duration of M1 was always 200 ms, including 5-ms (500-Hz signal) or 2-ms (1500- and 4000-Hz signals) raised-cosine ramps. The duration of M2 was the same as the signal (10 ms at 500 Hz, including 5-ms raised-cosine ramps, and 6 ms for the 1500- and 4000-Hz signals, including 2-ms ramps). The longer ramps and duration were used at the lowest signal frequency to avoid potential problems of “spectral splatter” and temporal overlap of the stimuli due to “ringing” of the cochlear filters (Shailer and Moore, 1987). When both maskers were present, all three stimuli (M1, M2, and the signal) were presented contiguously with no temporal gaps. When only M1 was present, the gap between M1 and the signal was equal to the total duration of M2. A 2/3 octave-wide Gaussian noise masker, generated and filtered in the spectral domain and centered geometrically around the signal frequency, was presented to the contralateral ear at an overall level 40 dB below that of the signal, to prevent the detection of the signal in the contralateral ear via acoustic and/or electric crosstalk.

TEN was used to elevate signal thresholds to approximate the thresholds in quiet for listeners with moderate hearing loss. The TEN was set at a level of 40 dB SPL within the estimated equivalent rectangular bandwidth (ERB) of the auditory filter centered at 1 kHz (Glasberg and Moore, 1990). In all the experiments the TEN and the contralateral noise were gated on together 300 ms before the beginning of the first interval and gated off at the end of the third interval. The level of 40 dB SPL per ERB was selected because this was the highest TEN level that resulted in a sufficient number of usable signal levels above the masked threshold, and so allowed for estimates of compression over a reasonable level range. A lower TEN level would have resulted in milder hearing losses that were approximated by the TEN thresholds; a higher TEN level would have resulted in relatively few masker levels that were high enough to produce elevations of signal threshold, but not so high as to be uncomfortably loud.

Thresholds for M1 and M2 were evaluated for signal levels ranging from 5 to 25 dB above masked thresholds in the TEN, in 5 dB steps. Because the TEN was always present, and limited the audibility of the signal, we will refer to these signal levels in terms of sensation level (SL), relative to the masked threshold of the signal in the TEN alone.

D. Results and discussion

Results are shown in Figs. 2–4 for signal frequencies of 500, 1500, and 4000 Hz, respectively. The data from each

FIG. 2. Masked thresholds with a 500 Hz signal. The plot depicts the signal level at threshold as a function of the sensation level of the signal (re: the masked threshold for the 500 Hz signal in the presence of 40 dB TEN, indicated by the value in square brackets in the upper left hand corner next to the listener number). The upward- and downward-pointing triangles show the threshold in the presence of M1 alone and M2 alone, respectively. The shaded diamonds show the threshold in the presence of both M1 and M2. The bold solid curves depict the predicted thresholds in the presence of M1 and M2 from the best fit using a power function, as described in the text. The dotted lines indicate predictions based on a linear summation of the effects from the two maskers. Error bars indicate standard deviations.

subject are shown in the individual panels, along with the mean data in the last panel of each figure. Signal levels at threshold (in dB SPL), as determined in Phase 2 of the experiment, are plotted as a function of nominal signal level (in dB SL) in the presence of the single maskers, as determined in Phase 1. If the correspondence between Phase 1 and Phase 2 were perfect, and no individual differences existed, signal thresholds in the presence of M1 alone and M2 alone (upward- and downward-pointing open triangles, respectively) would fall on top of each other, and follow a line that is parallel to the major diagonal with an offset (y intercept) equal to the signal threshold (dB SPL) in the presence of the TEN

FIG. 3. As for Fig. 2, but for a 1500 Hz signal.

FIG. 4. As for Fig. 2, but for a 4000 Hz signal.

alone. As can be seen for some listeners, such as S2 and S3 in Fig. 2, S2 in Fig. 3, and S1, S2, and S6 in Fig. 4, the M1 and M2 levels set in Phase 1 did not always result in exactly equivalent masked thresholds for those same maskers in Phase 2. Note that this occurred at times even for the listeners (S2 and S3) whose average data were used to set the M1 and M2 levels.

The degree to which the thresholds in the presence of the combined M1 + M2 masker (shown as shaded diamonds) exceed thresholds for either masker alone indicates the amount of compression. In many cases (but not all; see Fig. 4) for the individual and average data, thresholds for the combined masker exceeded those for the same maskers presented individually. Also note that unequal effectiveness for M1 and M2 in Phase 2 data does not preclude the finding of excess masking (e.g., see S3 in Fig. 2). Assuming that the masking effects are linearly additive, the increase in masking in the presence of the combined maskers can be used to determine the (compressive) nonlinearity of the response to the signal (Penner, 1980; Humes and Jesteadt, 1989), and

this nonlinearity has been associated with the compressive nonlinearity of the BM, which can be approximated as a power law over a limited level range (Oxenham and Moore, 1995; Plack and O'Hanlon, 2003; Plack *et al.*, 2008).

The calculation of the exponents associated with the compressive power law was accomplished by finding the best-fitting modified power function for each subject's data in the manner described in Humes and Jesteadt (1989). This method assumes that the ratio of the internal masker and signal intensities at threshold is a constant, and that the effect of combining M1 and M2 is a linear sum of their separate effects. The modified power function includes the listener's unmasked threshold (modeled as an internal noise separate from the forward maskers) and has been shown to accurately reflect forward masking additivity results when there is little to no forward masking (i.e., at low masker sensation levels when the thresholds are influenced more by absolute threshold than by the masking). In our case, the "unmasked" threshold refers to the threshold in the presence of the

background TEN masker alone. The following equation was used:

$$S_{M1+M2}^c = S_{M1}^c + S_{M2}^c - S_T^c, \tag{1}$$

where the first term represents the signal intensity at threshold, raised to a power c (where $c < 1$ for compressive response) in the presence of both maskers, the next two terms represent the compressed signal intensity at threshold for each masker in isolation, and the final term represents the compressed threshold in the absence of either forward masker. Based on the above-mentioned assumptions, all quantities can be specified except for c . A one-dimensional minimization routine (“fminbnd” in MATLAB) was used to determine the c value that minimized the squared error between the predicted and empirical values of $S_{M1 + M2}$, expressed in dB. The range of possible c values was limited to between 0 and 2. A constant value of c was assumed, rather than a polynomial function of signal level (as done in Plack *et al.*, 2008) because the range of levels over which c was estimated was typically small, and a polynomial function did not yield significantly better predictions.

The predicted thresholds in the presence of both maskers (shown as the bold solid line in Figs. 2–4) agreed quite well with the actual measured data for each of the individual subjects (gray diamonds), as well as the data averaged across subjects. The compression exponents that resulted in these fits are shown in Table II. The dotted line indicates the predicted thresholds for the combined masker condition that would be expected for linear summation in the absence of compression.

As can be seen from the first column of Table II, all listeners showed compressive responses to the 500 Hz signal, with exponent values ranging from 0.04 to 0.47. The second column of Table II shows compression exponents that best fitted the data from the 1500 Hz signal. There was also variability across the individual listeners, with exponents ranging from 0.11 to 1.82 (the latter indicating an expansive response), with the average compression less than that found at 500 Hz. The third column of Table II shows the compression exponents for the 4000 Hz signal frequency. Only data from two listeners (S1 and S2) resulted in compressive estimates ($c < 1$); the remaining exponents are consistent with essentially linear responses (with respect to intensity), or

TABLE II. Compression exponents for 500, 1500, and 4000 Hz signals in 40 dB/ERB TEN. The compression exponents fit to the average data are shown as well as the averaged compression exponents.

Listener	500 Hz signal	1500 Hz signal	4000 Hz signal
S1	0.18	0.40	0.18
S2	0.18	0.11	0.40
S3	0.04	0.47	1.44
S4	0.26	1.82	0.99
S5	0.47	0.56	1.94
S6	0.04	0.47	1.71
Fit to mean data	0.11	0.41	0.73
Fitted exponent averaged across listeners	0.20	0.54	1.11

even expansive responses. However, the expansive exponents are likely to be artifactual, due to some thresholds for M2 alone that were higher than for M1 + M2 (see S4, S5, and S6)—this apparent “release from masking” cannot be explained within the framework of the modified power law model. It may be that M1 produces adaptation in the response to M2, thereby reducing its effectiveness, but it remains unclear why such adaptation would not also reduce the response to the signal. Adaptation effects are not incorporated within this linear model of additivity and in most cases are not required to account for AFM data (e.g., Plack *et al.*, 2006).

Examination of the error bars in Figs. 2–4 shows that results across subjects were quite variable. However, overall, the results for the 4000 Hz signal appear to be the least compressive in this study, whereas the results for the 500 Hz signal appear to be the most compressive. The Shapiro–Wilk W test was used to evaluate whether or not the distribution of estimated compression exponents across subjects and across the three signal frequencies varied significantly from a normal distribution. The results indicated there was not a significant deviation from a normal distribution [$W(18)=0.933$; $p=0.218$]; therefore, standard parametric statistical techniques were used to evaluate the data. A one-way repeated-measures analysis of variance, with the compression exponent as the dependent variable, revealed a significant effect of signal frequency [$F(2,10)=4.698$, $p=0.036$], with a significant linear trend [$F(1,5)=10.178$, $p=0.024$], consistent with decreasing estimated compression with increasing signal frequency. *Post hoc* paired comparisons showed that there was a significant difference in the compression exponents between the 500 and the 4000 Hz signals [$t(5)=-3.190$; $p=0.024$]. There were no significant differences in compression estimates between the remaining possible pairs. (The same statistical conclusions were reached with nonparametric tests.)

Gregan *et al.* (2010) evaluated estimated BM compression using GOM functions for the same six listeners that were used in the present study. In contrast to the present results, the estimated exponents from Gregan *et al.* (2010), shown in Table III, were compressive for all listeners at 4 000 Hz. As can be seen by comparing the compression estimates from column 3 of Table II with those in Table III, the results obtained using the AFM technique indicate less compression than those obtained using forward GOM, even

TABLE III. Compression estimates for a 4000 Hz signal, taken from Gregan *et al.* (2010), using a different technique (GOM) but presented in the same level of background TEN.

Listener	Compression estimate for 4 000 Hz signal
S1	0.23
S2	0.16
S3	0.39
S4	0.44
S5	0.26
S6	0.30
Mean data	0.30

though both estimates were obtained from measurements in the presence of 40 dB TEN in the same listeners. This difference was confirmed using a paired t -test, which indicated a significant difference between the compression estimates from the two techniques [$t(5) = -2.933$, $p = 0.033$]. It is not clear what accounts for this difference in estimates. To our knowledge, there have been no previous attempts to compare estimates of compression in the same NH listeners using these two behavioral techniques. The level ranges over which the compression exponents were estimated are also comparable, with the masked signal threshold levels ranging from ~ 70 to 95 dB SPL in the off-frequency masker portion of the GOM experiment, and from 65 to 90 dB SPL in the current AFM experiment (4000 Hz signal).

The finding of reduced compression at the highest frequency is also in apparent contrast to earlier AFM data. Plack and O'Hanlon (2003) found compressive responses for all listeners at the higher signal frequency (4000 Hz) and variable responses (some compressive and some linear) for the lower signal frequencies (250 and 500 Hz), and Plack *et al.* (2008) showed similar degrees of compression for the lower frequency (250 Hz) and the higher frequency (4000 Hz) signals. Two possible reasons for the difference are (i) the relatively high signal levels used in the present study and (ii) the presence of a broadband TEN in the present study. Our study included the range from 40 to 90 dB SPL, whereas Plack and O'Hanlon (2003) tested levels that extended only to ~ 65 dB SPL. Plack *et al.* (2008) did test levels that extended to 90 dB SPL and their results were consistent with a highly compressive response for all four of their NH listeners; however, Plack and Arifianto (2010) also tested at high stimulus levels and their results did show some return to linearity at higher levels, for some listeners. Therefore, although level effects may at least partly explain the difference in results between Plack and O'Hanlon (2003) and the present study, they do not seem to account for the apparent difference with the Plack *et al.* (2008) study. The other potential factor is the presence of the TEN. Although Gregan *et al.* (2010) concluded that background noise did not affect estimates of compression for the GOM method, it is possible that background noise affects compression estimates more for the AFM method. Experiment 2 examined the effects of removing the background noise and lowering the signal presentation level on estimates of compression to determine if either of these factors could account for the lack of compression seen in Experiment 1 at 4000 Hz.

III. EXPERIMENT 2

Given that the compression estimates for the NMNH listeners using AFM for a 4000-Hz signal did not agree well with the compression estimates in other studies, or for the same six listeners in a previous study using GOM, it was decided to repeat the AFM experiment with the same subjects but in the absence of TEN. In addition, to evaluate the possibility that the high stimulus levels necessitated by the use of TEN affected compression estimates, the AFM was also measured at lower signal levels. These manipula-

tions also allowed for more direct comparison with results from Plack and O'Hanlon (2003).

A. Listeners and procedure

The AFM data were obtained in the same listeners who participated in Experiment 1, using identical stimuli and methods, with the only exceptions being the lack of TEN and the wider range of signal levels tested. Two level ranges were tested. The first level range was taken from Experiment 1, based on sensation levels in the presence of the TEN, with the levels of M1 and M2 presented at the same levels that were used in Experiment 1; the second level range (low-SPL condition) was determined based on thresholds in quiet, such that SL in this case refers to the more traditional measure relative to absolute threshold. Appropriate levels of M1 and M2 for the lower signal levels were determined as in Experiment 1, using the same two listeners (S2 and S3).

B. Results and discussion

Results for the high-level stimuli presented without TEN (4000 Hz signal) are shown in Fig. 5. Time constraints did not allow the measurement of all conditions for all listeners; therefore, some listeners do not have results for either 5 or 10 dB SL conditions. Figure 6 shows the data for low-level stimuli measured in quiet. Note that S4 was not available to run the low-level condition.

Signal levels for Phase 1 for the low-level stimuli measured in quiet were comparable with those used in Plack and O'Hanlon (2003) for a 4000-Hz signal. As in Experiment 1, these data were fitted using modified power law functions and the resulting exponents provide an estimate of compression. For the new data, the threshold parameter in the modified power law function was set to threshold in quiet (rather than threshold in 40 dB TEN). However, for the higher-level signals in Fig. 5, the abscissa shows the nominal signal levels in dB SL, relative to thresholds in the presence of TEN, for ease of comparison with Fig. 4. The derived compression exponents are shown in Table IV.

To determine whether the presence of the background noise affected estimates of compression, the fitted compression estimates for the 4000-Hz high-level stimuli in the absence of TEN were compared with those from experiment 1. A paired-samples t -test did not reveal a significant difference between the compression estimates in TEN versus in quiet [$t(5) = 2.479$, $p = 0.056$], in line with our earlier study that showed no significant effect of TEN on compression exponents derived using the GOM technique (Gregan *et al.*, 2010). However, it is important to note that there is a trend for the compression estimates to become more compressive in the absence of TEN for all listeners except S1, even though this trend did not quite reach statistical significance. Recall that in Experiment 1, there was a significant difference between compression estimates for a 500 versus 4000 Hz signal. We used a paired t -test to evaluate whether or not the compression estimates for the 4000-Hz high-level stimuli without TEN were still significantly different than the estimates for the 500-Hz signal in TEN. This comparison confirmed that the compression exponents for 4000 Hz, even in

FIG. 5. Additivity results for a 4000 Hz signal using the same M1 and M2 levels as for Fig. 4, but in the absence of TEN. Note that all listeners completed the 20 and 25 dB SL conditions, but not all listeners completed both the 10 and 15 dB SL conditions. Therefore, the data points on the average panel have different n values for the two lower level conditions. See the text for further details.

the absence of TEN, were still significantly higher (i.e., less compressive) than those for the 500-Hz signal in TEN [$t(5) = -2.758, p = 0.040$].

It is possible that the relatively high stimulus levels required due to the presence of the TEN resulted in decreased compression estimates due to a return to linearity at high intensities, as has been found in physiological (e.g., Recio *et al.*, 1998) and psychophysical (e.g., Oxenham and Plack, 1997) studies. The right column of Table III shows the compression exponents for a subset of listeners using lower level stimuli in quiet, comparable to those from Plack and O'Hanlon (2003). These exponents are compressive for all listeners, except S5. A paired t -test was used to evaluate whether these compression exponents were significantly different from those derived from the high-level 4000-Hz signal

FIG. 6. Thresholds for a 4000 Hz signal using lower signal levels in the absence of TEN. Symbol legend same as for previous figures.

in the absence of TEN. Note that S4 was excluded from this analysis as she did not have results for both conditions. The difference was not significant [$t(4) = -0.214, p = 0.841$], indicating that compression estimates for the 4000 Hz low-level quiet stimuli were not significantly more compressive than those for the 4000 Hz high-level quiet stimuli. Therefore, although the trend toward more compressive results for the lower level stimuli suggests that some of the reduction in compression may have been due to the high overall levels used, this trend did not reach statistical significance. In addition, it is also not clear why similar effects were not observed at the lower frequencies (500 and 1500 Hz), which had similarly high stimulus levels, or why earlier studies using higher signal levels did not observe as linear responses at these levels (e.g., Plack *et al.*, 2008).

TABLE IV. Compression exponents from fits to data in the absence of TEN. The missing value (—) indicates the listener was not available to run the specified condition.

Listener	4 000 Hz signal no TEN	4 000 Hz signal low SPL/no TEN
S1	0.34	0.26
S2	0.29	0.11
S3	0.33	0.26
S4	0.58	—
S5	1.09	1.96
S6	1.20	0.37
Average fit to mean data	0.44	0.26
Compression exponent averaged across listeners	0.64	0.59

One possible explanation for the reduced compression estimates seen at 4000 Hz, besides level, is that the signal representation was affected by the medial olivocochlear (MOC) efferents. These fibers synapse with the outer hair cells and have been shown to reduce the effects of the cochlear amplifier, particularly for longer duration, high-frequency stimuli (e.g., Guinan, 2006; Krull and Strickland, 2008). It is possible that the longer duration of M1 may have resulted in a reduction in cochlear gain via this efferent system, the effect of which may have been seen in the present data as a reduction in the compression exponent.

Data from Krull and Strickland (2008) showed that the apparent contribution of efferent activation was most pronounced for lower signal levels, below ~ 60 dB SPL. This suggests that the present data, with minimum signal levels of ~ 60 dB SPL due to the presence of the TEN, would not have been influenced by efferent effects. A recent study by Plack and Arifanto (2010) examined possible MOC effects in an AFM paradigm. They evaluated results with a longer (200 ms) and shorter (20 ms) M1 duration and found only minor effects of the duration change in one of their experiments. The lack of systematic changes in the results suggests the present results (which also used an M1 duration of 200 ms) were likely not affected by a reduction in gain due to efferent activation.

IV. DISCUSSION

The present study examined AFM results for NH listeners in a background of TEN, thus requiring stimulus levels that were comparable to those used in previous studies for HI listeners (in terms of SL and SPL). Results were sampled at low, medium, and high signal frequencies. The data from Experiment 1 showed estimates of compression for a low (500 Hz) signal frequency that agreed well with those of previous studies (Plack and O'Hanlon, 2003; Plack *et al.*, 2008). Similar estimates of compression were seen for the 1500 Hz signal, although the average compression exponent was somewhat higher (less compressive) than for the 500-Hz signal. The data at 4000 Hz showed compressive results for two listeners, but essentially linear results for the remaining four listeners, resulting in no significant compression overall. Testing in the absence of TEN (at the same stimulus levels) did not result in a significant change in the estimated amount of compression, although there was a trend for the estimates to become more compressive in the absence of TEN. The lack of an effect of TEN on estimates of compression is consistent with the results of Gregan *et al.* (2010), who also found no effect of TEN on estimates of compression using a different technique (GOM). However, they found significant compression at a signal frequency of 4000 Hz, over a similar range of signal levels, both with and without TEN.

The fact that compressive responses were observed for all but one listener for low-level 4000 Hz signals suggests the reduction in compressive response may have been due to the high stimulus levels used in the main experiment, as found in some other studies for signal levels of 80 dB SPL or higher (e.g., Oxenham and Plack, 1997). However, this

trend did not reach statistical significance and it is not clear why similar level effects were not observed at the lower signal frequencies. One possible explanation for the less-compressive results at the higher signal frequency is inhibition due to activation of the MOCB. This activity has been shown to be more pronounced for higher signal frequencies. This was deemed unlikely because, previous research (Jennings *et al.*, 2009) showed that the effect of assumed MOC inhibition using psychoacoustic tasks was most apparent for moderate stimulus levels (below about 70 dB SPL) and became less effective at higher levels. In addition, results from Plack and Arifanto (2010) showed only minor efferent effects in an AFM paradigm very similar to the one used in the present study.

Another possible explanation for the trend toward reduced compression with increasing signal frequency is the activation of the acoustic reflex (AR). In response to high-level stimuli, the stapedius muscle contracts and this results in decreased transmission of sound through the middle ear. The AR is elicited more consistently for lower frequency stimuli, whereas it is more variable at 4000 Hz, even among NH listeners (e.g., Jerger *et al.*, 1972). In addition, the effects of the AR result primarily in attenuation of lower-frequency (below ~ 2000 Hz), regardless of the stimulating frequency (e.g., Humes, 1978). Therefore, even if the high levels of the 4000 Hz stimuli elicited the AR in the current study, it is unlikely that this would have resulted in attenuation of the 4000-Hz stimuli. Plack and Arifanto (2010) did not find an effect of AR using a similar AFM paradigm. However, they did not use stimulus frequencies below 1500 Hz. The time course of the AR is on the order of tens of milliseconds, with an onset latency of ~ 70 – 100 ms for a 500-Hz stimulus (Church and Cudahy, 1984). Therefore, it could have readily been invoked by the higher level masking stimuli, thereby decreasing the effective level of stimuli reaching the cochlea. This, in turn, may have resulted in the stimulus levels coinciding with a region of compression. If the AR were not elicited (or not as effective) at 4000 Hz, the level of the stimuli reaching the cochlea would have been higher than for the 500- and 1500-Hz signals, and this may have coincided with a more linear region of the basilar membrane response. It could be argued that the level of the TEN itself may have activated the AR, as it has been shown that broadband stimuli have a lower AR activation threshold than do more frequency-specific stimuli (Wilson, 1981). Although the AR could have been present for all frequencies, it would have still been least effective for the 4000 Hz signal, perhaps resulting in more linear cochlear processing at that high frequency. Nevertheless, it still remains difficult to reconcile the present data at 4000 Hz with those from earlier studies of GOM (Gregan *et al.*, 2010) and AFM (Plack and O'Hanlon, 2003).

Overall, it appears that AFM can result in estimates of compression that are consistent with those from physiological data and other psychophysical techniques, for some NMNH listeners. However, the lack of compression seen for some listeners at high stimulus levels and high signal frequencies, suggests that caution should be applied when interpreting noncompressive results for HI listeners when measured at similarly high stimulus levels.

ACKNOWLEDGMENTS

This work was supported by the NIH Grant No. R01 DC 03909. The authors thank Chris Plack, Magdalena Wojtczak, and two anonymous reviewers for their thoughtful comments on earlier versions of this manuscript.

- ANSI (2004). *Specifications for audiometers* (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.
- Church, G. T., and Cudahy, E. A. (1984). "The time course of the acoustic reflex," *Ear Hear.* **5**, 235–242.
- Cokely, C. G., and Humes, L. E. (1993). "Two experiments on the temporal boundaries for the nonlinear additivity of masking," *J. Acoust. Soc. Am.* **94**, 2553–2559.
- Dubno, J. R., and Schaefer, A. B. (1992). "Comparison of frequency selectivity and consonant recognition among hearing-impaired and masked normal-hearing listeners," *J. Acoust. Soc. Am.* **91**, 2110–2121.
- Glasberg, B. R., and Moore, B. C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res.* **47**, 103–138.
- 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.
- Guinan, J. J., Jr. (2006). "Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans," *Ear Hear.* **27**, 589–607.
- Hall, J. W., and Grose, J. H. (1997). "The relation between gap detection, loudness, and loudness growth in noise-masked normal-hearing listeners," *J. Acoust. Soc. Am.* **101**, 1044–1049.
- Hellman, R. P., and Meiselman, C. H. (1993). "Rate of loudness growth for pure tones in normal and impaired hearing," *J. Acoust. Soc. Am.* **93**, 966–975.
- Humes, L. E. (1978). "The effects of middle ear muscle contraction on auditory and overload thresholds," *Audiology* **17**, 360–367.
- Humes, L. E., and Jesteadt, W. (1989). "Models of the additivity of masking," *J. Acoust. Soc. Am.* **85**, 1285–1294.
- Jennings, S. G., Strickland, E. A., and Heinz, M. G. (2009). "Precursor effects on behavioral estimates of frequency selectivity and gain in forward masking," *J. Acoust. Soc. Am.* **125**, 2172–2181.
- Jerger, J., Jerger, S., and Mauldin, L. (1972). "Studies in impedance audiometry. I. Normal and sensorineural ears," *Arch. Otolaryngol.* **96**, 513–523.
- Krull, V., and Strickland, E. A. (2008). "The effect of a precursor on growth of forward masking," *J. Acoust. Soc. Am.* **123**, 4352–4357.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- 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.
- 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.
- 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, 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.
- Oxenham, A. J., and Bacon, S. P. (2003). "Cochlear compression: Perceptual measures and implications for normal and impaired hearing," *Ear Hear.* **24**, 352–366.
- 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. (1995). "Additivity of masking in normally hearing and hearing-impaired subjects," *J. Acoust. Soc. Am.* **98**, 1921–1934.
- 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.
- Penner, M. J. (1980). "The coding of intensity and the interaction of forward and backward masking," *J. Acoust. Soc. Am.* **67**, 608–616.
- Plack, C. J., and Arifanto, 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., 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.
- Plack, C. J., and O'Hanlon, C. G. (2003). "Forward masking additivity and auditory compression at low and high frequencies," *J. Assoc. Res. Otolaryngol.* **4**, 405–415.
- Plack, C. J., Oxenham, A. J., and Drga, V. (2006). "Masking by inaudible sounds and the linearity of temporal summation," *J. Neurosci.* **26**, 8767–8773.
- Plack, C. J., Oxenham, A. J., Simonson, A. M., O'Hanlon, C. G., Drga, V., and Arifanto, D. (2008). "Estimates of compression at low and high frequencies using masking additivity in normal and impaired ears," *J. Acoust. Soc. Am.* **123**, 4321–4330.
- Recio, A., Narayan, S. S., and Ruggero, M. A. (1998). "Basilar-membrane responses to clicks at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.* **103**, 1972–1989.
- Recio-Spinoso, A., and Lopez-Poveda, E. A. (2010). "Basilar membrane responses to simultaneous presentations of white noise and a single tone," in *The Neurophysiological Bases of Auditory Perception*, edited by E. A. Lopez-Poveda, A. R. Palmer, and R. Meddis (Springer, New York), pp. 15–23.
- Reed, C. M., Braid, L. D., and Zurek, P. M. (2009). "Review article: Review of the literature on temporal resolution in listeners with cochlear hearing impairment: A critical assessment of the role of suprathreshold deficits," *Trends Amplif.* **13**, 4–43.
- Rhode, W. S. (1971). "Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique," *J. Acoust. Soc. Am.* **49**, 1218–1231.
- Rhode, W. S., and Cooper, N. P. (1996). "Nonlinear mechanics in the apical turn of the chinchilla cochlea in vivo," *Auditory Neurosci.* **3**, 101–121.
- Robles, L., and Ruggero, M. A. (2001). "Mechanics of the mammalian cochlea," *Physiol. Rev.* **81**, 1305–1352.
- Ruggero, M. A. (1992). "Responses to sound of the basilar membrane of the mammalian cochlea," *Curr. Opin. Neurobiol.* **2**, 449–456.
- 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.
- Sellick, P. M., Patuzzi, R., and Johnstone, B. M. (1982). "Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique," *J. Acoust. Soc. Am.* **72**, 131–141.
- 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.
- 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.
- Stainsby, T. H., and Moore, B. C. (2006). "Temporal masking curves for hearing-impaired listeners," *Hear. Res.* **218**, 98–111.
- Wilson, R. H. (1981). "The effects of aging on the magnitude of the acoustic reflex," *J. Speech Hear. Res.* **24**, 406–414.
- Wilson, R. H., and Carhart, R. (1971). "Forward and backward masking: Interactions and additivity," *J. Acoust. Soc. Am.* **49**, 1254–1263.
- Wojtczak, M., and Oxenham, A. J. (2009). "Pitfalls in behavioral estimates of basilar-membrane compression in humans," *J. Acoust. Soc. Am.* **125**, 270–281.
- Zinn, C., Maier, H., Zenner, H., and Gummer, A. W. (2000). "Evidence for active, nonlinear, negative feedback in the vibration response of the apical region of the in-vivo guinea-pig cochlea," *Hear. Res.* **142**, 159–183.
- Zurek, P. M., and Delhorne, L. A. (1987). "Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment," *J. Acoust. Soc. Am.* **82**, 1548–1559.