Acknowledgements

Sincere gratitude to my co-advisors: Dr. Peggy Nelson and Dr. Andrew Oxenham. Without their support, guidance, funding, and painstaking editing, this project would not have reached completion. Special thanks to Peggy for serving as therapist in addition to advisor—I would not have made it through this without you.

Also thank you to my committee: Drs. Robert Schlauch, Dianne Van Tasell, and Neal Viemeister for each of their unique contributions to this venture.

Additional thanks to my friends and colleagues who have supported me along the way: Magda Wojtczak, Bess Borchert, Liz Anderson, Sarah Angerman, Jane Carlstrom, and Sharon Miller.

Finally, thank you to all of the students I have had the pleasure of teaching (and who have, in turn, taught me much as well).


This work was supported by NIH grant R01 DC03909. We thank Magdalena Wojtczak, Brian Moore, Enrique Lopez-Poveda, and an anonymous reviewer for their helpful comments on earlier versions of this manuscript.


This work was supported by the NIH Grant R01 DC 03909. The authors thank Chris Plack, Magdalena Wojtczak, and two anonymous reviewers for their thoughtful comments on an earlier version of this manuscript.

Chapter 4: This work was supported by NIH Grants R01 DC 03909 and R01 DC 008306, as well as a Bryngelson grant from the Department of Speech, Language, Hearing sciences at the University of Minnesota.
Dedication

My little sister…Melissa “Missy” JoAnne Gregan
4/5/72-3/17/87

Also to the furry purrballs…Fechner, Catapotamus, and Noodle…who do not care whether or not I have a Ph.D., as long as I give them snuggles and noms.
Abstract

Hearing-impaired (HI) listeners often show poorer performance on psychoacoustic tasks than do normal-hearing (NH) listeners. For example, listeners with cochlear hearing loss often show less benefit, or masking release (MR), than do listeners with NH when a steady-state masker is replaced by a temporally fluctuating one, even when overall audibility is equated. It is possible that some of this degraded performance may be due to reduced cochlear compression. To compare performance of HI and NH listeners, the sensation level (SL) and overall SPL of the stimuli must be equated. Masking noise can be used to accomplish this by raising the thresholds of NH to equal the thresholds in quiet of HI listeners. However, such noise may have other effects, including changing peripheral response characteristics, such as the compressive input-output function of the basilar membrane in the normal cochlea. To test this, the first study estimated compression behaviorally across a range of background-noise levels in six NH listeners at a 4-kHz signal frequency, using a growth-of-forward-masking (GOM) paradigm. For signals 5 dB or more above threshold in noise, no significant effect of broadband noise level was found on estimates of compression.

Due to limitations with the use of forward GOM for low-frequency signals, and the desire to estimate compression across a wide range of frequencies, additivity of forward masking (AFM) was evaluated in the second study. AFM was measured in six NH listeners for signal frequencies of 500, 1500, and 4000 Hz in the presence of the same TEN background noise used in the first study. Results consistent with compressive BM responses were found for all listeners for the 500 Hz signal, 5 listeners at 1500 Hz, and
only 2 listeners at 4000 Hz. Further measurements in the absence of background noise also indicate a lack of consistent compression at 4000 Hz at higher signal levels, in contrast to earlier results collected at lower levels. Better understanding of this issue will be required before AFM can be used as a general behavioral estimate of BM compression.

The final study tests whether changes in cochlear compression can predict the degree of MR loss. Behavioral estimates of compression, using temporal masking curves (TMCs), were compared with MR for speech and pure tones in hearing-impaired (HI) individuals and age-matched, noise-masked normal-hearing (NMNH) listeners.

Compression estimates were made at 500, 1500, and 4000 Hz. Pure-tone masking period patterns and MR for band-limited (500-4000 Hz) IEEE sentences were measured in a 10-Hz square-wave-gated noise. In addition, an estimate of temporal resolution was calculated using the off-frequency curve from the TMC measurements. No strong relationship was found between estimates of cochlear compression and MR for either speech or pure tones. There was a nonsignificant correlational trend between temporal resolution estimates and MR for speech overall, and NMNH had significantly steeper recovery slopes than did HI. 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 compression.
# Table of Contents

List of Tables .......................................................................................................................... vii

List of Figures .......................................................................................................................... viii

CHAPTER 1: INTRODUCTION ................................................................................................. 1

CHAPTER 2: FORWARD GROWTH OF MASKING ................................................................. 12

I. Introduction .......................................................................................................................... 12

II. Estimating compression as a function of background noise level ......................... 16
   A. Listeners ......................................................................................................................... 16
   B. Procedure ....................................................................................................................... 16
   C. Stimuli ............................................................................................................................. 17
   D. Results ............................................................................................................................ 20

III. Discussion ......................................................................................................................... 28

CHAPTER 3: ADDITIVITY OF FORWARD MASKING ......................................................... 35

I. Introduction .......................................................................................................................... 35

II. Experiment 1 ....................................................................................................................... 39
   A. Listeners ......................................................................................................................... 39
   B. Procedure ....................................................................................................................... 40
   C. Stimuli ............................................................................................................................. 43
   D. Results and discussion ................................................................................................. 45

III. Experiment 2 ....................................................................................................................... 55
   A. Listeners and procedure ............................................................................................... 56
   B. Results and discussion ................................................................................................. 56

IV. Discussion .......................................................................................................................... 62

CHAPTER 4: MASKING RELEASE ....................................................................................... 66

I. Introduction .......................................................................................................................... 66

II. Experiment 1: Masking release with speech ................................................................. 71
   A. Listeners ......................................................................................................................... 71
   B. Threshold elevation ........................................................................................................ 73
   C. Stimuli and procedure ................................................................................................. 74
List of Tables

Table 2.1: Thresholds in various TEN levels..........................................................20
Table 2.2: GOM slope estimates..............................................................................28
Table 3.1: M1 and M2 thresholds for S2 and S3 .................................................41
Table 3.2: Compression estimates using AFM......................................................51
Table 3.3: Compression estimates at 4000 Hz using GOM..............................53
Table 3.4: Compression estimates for high-level stimuli in quiet and low-level stimuli in quiet .................................................................59
Table 4.1: Demographic and threshold data for listeners ...............................72
Table 4.2: $R^2$ values for sigmoidal fits to speech functions.........................79
Table 4.3: Masking release for pure tones ..............................................................93
Table 4.4: Linear fits to TMC data .......................................................................105
List of Figures

Figure 2.1: Individual growth of masking (GOM) functions ...............21
Figure 2.2: Mean GOM functions..................................................23
Figure 2.3: Individual I/O functions ..............................................25
Figure 2.4: Mean I/O functions......................................................25
Figure 3.1: Additivity stimuli schematic .......................................41
Figure 3.2: Additivity of forward masking (AFM) at 500 Hz..............47
Figure 3.3: AFM at 1500 Hz...........................................................48
Figure 3.4: AFM at 4000 Hz............................................................49
Figure 3.5: AFM: high level stimuli without TEN............................57
Figure 3.6: AFM: low level stimuli without TEN.............................58
Figure 4.1: IEEE P/I function for HI listeners ..................................77
Figure 4.2: IEEE P/I function for NMNH listeners..........................78
Figure 4.3A: Masking release (MR) at -5 dB SNR vs. SNR for 50% in SSN ............................................................82
Figure 4.3B: SRT difference vs. SNR for 50% in SSN ......................83
Figure 4.4A: MR at -5 dB SNR vs. speech level (HI)......................85
Figure 4.4B SRT difference vs. speech level (HI) ............................85
Figure 4.5C MR at -5 dB SNR vs. speech level (NMNH) .................86
Figure 4.5D SRT difference vs. speech level (NMNH) .....................86
Figure 4.5A: MR at -5 dB SNR vs. speech level (NMNH and HI) .......87
Figure 4.5B: SRT difference vs. speech level (NMNH and HI) ...........87
Figure 4.6A MR at -5 dB SNR vs. average quiet threshold (HI) .........88
Figure 4.6B: SRT difference vs. average quiet threshold (HI).............. 89

Figure 7: Schematic for pure tone masking release stimuli....................93

Figure 4.8A: Average masking release for pure tones (MRPT) vs. MR for speech at -5 dBSNR.........................................................94

Figure 4.8B: Average MRPT vs. SRT difference..................................95

Figure 4.9A: TMC data for HI listeners .............................................101

Figure 4.9B: TMC data for NMNH listeners .....................................102

Figure 4.10A: Derived I/O functions for HI listeners .........................103

Figure 4.10B: Derived I/O functions for NMNH listeners .................104

Figure 4.11: Compression estimates vs. quiet threshold (HI).............108

Figure 4.12: Compression estimates vs. quiet threshold (NMNH)........ 108

Figure 4.13A: MR at -5 dB SNR vs. compression estimates (HI)....... 111

Figure 4.13B: SRT difference vs. compression estimates (HI)......... 111

Figure 4.14A: MR at -5 dB SNR vs. compression estimates (NMNH)...112

Figure 4.14B: SRT difference vs. compression estimates (NMNH)...... 112

Figure 4.15A: MRPT vs. compression estimates (HI)..........................113

Figure 4.15B: MRPT vs. compression estimates (NMNH)...................114

Figure 4.16A: Off-frequency TMC slope vs. MR at -5 dB SNR...........115

Figure 4.16B: Off-frequency TMC slope vs. SRT difference ...............115

Figure 4.17: Off-frequency TMC slope vs. average quiet threshold ......116
CHAPTER 1: INTRODUCTION

Hearing loss has been classified as the most common deficit involving sensory systems, with an estimated 250 million affected individuals worldwide in 2000 as reported by Mathers et al. (2000) a World Health Organization report. In the United States, there were approximately 31 million hearing-impaired (HI) people in 2007, and this number is expected to increase as people live longer lives and more are affected by presbycusis (age-related hearing loss) (Kochkin, 2007). Individuals with sensorineural hearing loss experience difficulty hearing low-level sounds due to their hearing loss. However, even when sounds are fully audible (when presented a high levels, or through amplification via a hearing aid), individuals with hearing impairment still show residual difficulty, especially in complex listening situations, such as those with background noise. In fact, a recent survey of 3174 hearing-aid users showed that while 79% were satisfied overall with their hearing instruments, 25% were dissatisfied with their ability to hear in noisy situations (Kochkin, 2010). This is a common theme for clinicians working with the HI population: the most difficult listening situation for these individuals involves hearing in background noise. Hearing aid performance in noisy situations, for the aforementioned survey population, was strongly correlated with overall satisfaction with hearing aid use (Kochkin, 2010). Clearly, while our ability to amplify speech in quiet or low-noise situations with hearing aids proves beneficial, there is still work to be done to improve the ability of HI listeners to understand speech in difficult listening situations.

In everyday listening situations, the sounds that typically interfere with speech understanding (such as other talkers, clattering of dishware during a meal, etc.) tend to
vary over time and in spectral composition. They are not steady-state signals. Research has shown that these types of signals seem more problematic for HI listeners than for normal-hearing (NH) individuals. In fact, whereas more similar performance is often seen between NH and HI listeners in a steady-state background, NH listeners tend to show improved performance in a modulated background, whereas many HI listeners either maintain their steady-state performance or show decreased performance. NH listeners are apparently able to take advantage of the temporal gaps in the masking noise to retrieve additional information and thereby improve their ability to understand the speech. This improvement has been termed masking release (MR). The findings of increased difficulty in modulated background noise for HI compared to NH listeners have been supported by numerous research studies (e.g., Miller and Licklider, 1950; Wilson and Carhart, 1969; Duquesnoy, 1983; Festen and Plomp, 1990; Gustafsson and Arlinger, 1994; Eisenberg et al., 1995; Bacon et al., 1998; Peters et al., 1998; Dubno et al., 2002; Nelson and Jin, 2004; Summers and Molis, 2004; Lorenzi et al., 2006).

One of the most obvious differences between NH and HI listeners that could contribute to the reduced MR seen for the HI group is the difference in absolute thresholds. Specifically, it is possible that the lack of MR for HI listeners is due to speech information in temporal dips of the masker falling below absolute threshold due to the hearing loss. If reduced hearing sensitivity, which limits access to speech information in the masker dips, is the primary reason MR is limited or absent in some HI listeners, then both groups should show similar results when absolute thresholds are equated. Trine (1995) showed that performance for a group of mild-to-moderately HI listeners was
comparable to that of NH listeners once the poorer thresholds and associated limited bandwidth of the HI listeners was taken into account. However, other studies have shown that some HI listeners do not show performance equal to NH listeners, even when absolute threshold has been taken into account (e.g., Eisenberg et al., 1995; Bacon et al., 1998; Dubno et al., 2002; 2003). In Bacon et al. (1998), the noise-masked NH (NMNH) listeners showed reduced MR compared to their unmasked results. With quiet thresholds equated across NH and HI groups, MR was similar to that seen for the NMNH listeners for some HI subjects, indicating lack of access to speech information was indeed the primary issue. However, some HI listeners still showed less MR than their noise-masked counterparts. Therefore, it appears that even when the proportion of speech information above threshold is the same for HI and NMNH listeners in a modulated background, some HI individuals are not able to utilize this information as well as NMNH listeners.

Another possible explanation for the reduced MR seen in some HI listeners is a loss of the active cochlear process. The active process (or cochlear amplifier) is believed to be primarily a result of the motility of outer hair cells (OHCs). This motion results in amplified motion of the basilar membrane, with the amount of amplification varying as a function of input level. Specifically, the most amplification is applied at low input levels that excite a relatively frequency-specific region of the basilar membrane (e.g., Ruggero and Rich, 1991; Robles and Ruggero, 2001). This results in compression that gradually increases for input levels of approximately 15-50 dB SPL, and reaches a maximum for input levels of approximately 40-80 dB SPL. Perceptual consequences of this level- and frequency-dependent amplification include low thresholds for the detection of pure tones,
good frequency selectivity (narrow tuning of the auditory filters), and compressive loudness growth (Moore et al., 1999b). This compressive non-linearity may also allow listeners to better detect and understand speech in the gaps of a fluctuating masker. The cochlear amplifier would presumably apply gain to the moderate level speech in the masker dips, thereby increasing its effective level. This would not be the case for HI listeners with reduced or absent compression due to OHC damage associated with certain types of sensorineural hearing loss. The effects of cochlear damage (as may occur with certain types of sensorineural hearing loss) on basilar membrane nonlinearity have been measured directly in animal models. Following extreme damage to the OHCs (due to organism death, furosemide application, etc.), the basilar-membrane response becomes linear (e.g., Ruggero and Rich, 1991; Ruggero et al., 1997). Perceptually, this means that even if the speech information is presented at the same sensation level across the two (threshold-equated) groups; the effective internal level of the speech would be higher for the NMNH listeners due to amplification of speech in the masker dips. Several studies have shown that HI listeners have varying degrees of cochlear compression that may not be strongly associated with their pure-tone thresholds (Plack et al., 2004; Lopez-Poveda et al., 2005; Poling et al., 2011). This may be due to different patterns of underlying sensory cell damage in the cochlea that involves the OHCs to varying degrees and could be responsible for the range of MR seen across HI listeners. Therefore, it is not currently possible for clinicians to predict which HI listeners will experience the most difficulty in complex and fluctuating background sounds from audiogram review alone.
The goal of this thesis is to determine the relationship between MR in speech for HI listeners and behavioral estimates of basilar membrane compression, temporal resolution, and masking release for pure tones. In order to control for audibility effects between the HI listeners and their NH controls, TEN will be used to equate thresholds across listener pairs. Therefore, preliminary experiments will explore how two measures used to estimate compression are affected by the addition of a background noise in NH listeners. The experiments designed to address each of these goals are described in more detail in the following sections.

In order to better understand the range of degraded performance in complex listening conditions for HI listeners, and the possible contribution of the underlying basilar-membrane responses, it is necessary to differentiate results that are due to elevated absolute thresholds, versus supra-threshold performance deficits. Performance on some tasks varies within NH listeners as a function of overall stimulus level as well as stimulus level relative to the threshold of hearing (sensation level; SL). Therefore, it is also important to make comparisons between damaged and normal auditory systems using stimuli that are as similar as possible in terms of overall sound pressure level (SPL) and SL. One way to achieve this goal is to use steady background noise to elevate thresholds in NH listeners: both NH and HI listeners would then be listening at the same overall absolute and sensation levels and, therefore, if differences were seen in performance between the two groups, it would need to be attributed to something other than audibility and relative stimulus levels.
Performance of NMNH listeners has been found to be similar to that of HI listeners, at least in certain psychoacoustic tasks. In addition to threshold elevation, loudness growth functions become steeper (Steinberg and Gardner, 1937; Hellman and Meiselman, 1990; Hellman and Meiselman, 1993) and auditory filters become broader (Leek and Summers, 1993; Sommers and Humes, 1993). All of these effects are consistent with a reduction in the cochlear active process (Bacon and Oxenham, 2004). However, the effect of background noise on estimates of cochlear compression in NH listeners has not been examined. In order to explore the hypothesis that reduced cochlear compression results in reduced MR for HI listeners compared to their noise-masked NH counterparts, it is first necessary to test whether adding noise to NH listeners does not result in changes to estimates of the underlying compression. Therefore, Chapter 2 explores the effects of adding a background noise to simulate hearing loss on behavioral estimates of cochlear response growth in six NH listeners.

Due to the invasiveness of the procedure, it is not possible to measure the response growth of the basilar membrane directly in humans. However, there are behavioral techniques that appear to provide a reasonable estimate of cochlear compression, such as growth of masking (GOM, Oxenham and Plack, 1997) and temporal masking curves (TMC, Nelson et al., 2001). These behavioral estimates of compression have both been shown to provide results that are in relatively good agreement with direct physiological measures undertaken in other species (e.g., Lopez-Poveda et al., 2003). While both measures appear to yield similar estimates, GOM may
result in more stable estimates of compression for a given subject, at least for frequencies above approximately 1000 Hz (Rosengard et al., 2005).

Despite their relative strengths, both GOM and TMC techniques suffer from some limitations. In particular, it is difficult to obtain compression estimates at lower signal frequencies. This is primarily because both methods rely on the assumption that the response at a given place along the BM to frequencies well below the characteristic frequency (CF) of the BM place is linear. Although this assumption is justified for the basal (high-frequency) portion of the cochlea (Rhode and Recio, 2000), less is known about the mechanics of the apical (low-frequency) portion. For technical reasons, access to the apex of the cochlea is much more difficult to achieve without damaging the structure (Rhode and Cooper, 1996). Nevertheless, the available data suggest that the compressive response of the cochlea at the apical end may extend to frequencies well below (and above) CF (Cooper and Rhode, 1995). Because HI listeners often have varying amounts of hearing loss across frequencies, they may also have varying amounts of compression across frequencies. Therefore, it is important to be able to estimate cochlear response growth for low as well as high frequencies. An alternative behavioral measure involves measuring the additivity of forward masking to estimate compression for various frequencies. Briefly, two forward maskers (that produce equal masking of the signal in isolation) are presented contiguously and threshold for the signal is determined. If the signal is compressed, masked threshold in this condition should be greater than the 3-dB that would be predicted when summing the effective energy of two equally effective maskers (Green and Swets, 1966). This “excess masking” in the presence of
two equally effective maskers has been shown in NH listeners (Wilson and Carhart, 1969; Cokely and Humes, 1993; Oxenham and Moore, 1995; Plack and O'Hanlon, 2003; Plack et al., 2008; Plack and Arifianto, 2010). This technique has been shown to produce compression estimates in NH listeners that are similar to those obtained using other behavioral measures such as GOM (e.g., Plack and O'Hanlon, 2003), but this approach has not been fully evaluated in NMNH or HI listeners. Chapter 3 explores estimates of cochlear compression using additivity of forward masking in six NH listeners in quiet and in a masking noise designed to elevate their thresholds.

Chapter 4 directly explores the relationship between MR and estimates of cochlear compression. It is likely that MR for a broadband signal (i.e. speech) does not depend just on compression in one frequency region. Rosengard (2004) looked at correlations between MR and various suprathreshold measures at frequencies between 1 and 4 kHz. She found that amount of MR in her six HI listeners was correlated with the lowest slope ratio (or highest amount of compression) for a given subject—not an average of the slope ratios across frequencies. She suggested that this showed MR for band-limited speech (approximately 1000-4000 Hz) required residual compression in at least one frequency region encompassed by the stimuli, not necessarily all frequency regions. Jin and Nelson (2006) found that poorer low-frequency thresholds were associated with reduced MR in their HI listeners. They postulated that this reflected a lack of fundamental frequency information available to the HI listeners that could be used to help connect the speech information in the masker dips. However, it also is possible that the reduced MR seen for subjects with less low-frequency hearing was due to the fact
that this resulted in less compression across the entire frequency region of the speech signal. Perhaps the HI listeners with better low-frequency hearing were able to make use of residual compression in that region to gain information from speech in the masker dips.

These results suggest that a measure of cochlear compression at just one frequency may not be sufficient to predict MR. If a HI listener did not show compression at 4000 Hz, but they showed MR, it is possible that compression in another frequency region may be contributing to the results. A further potential factor contributing to reduced MR in HI listeners is an impairment in temporal resolution (e.g., George et al., 2006). Some studies have measured recovery from forward masking as an estimate of temporal resolution (e.g., Dubno et al., 2002; 2003). However, recovery from forward masking is also affected by cochlear compression making it difficult to dissociate the two potential factors (Oxenham and Moore, 1997).

A possible method for measuring temporal resolution separately from compression is suggested by Wotjczak and Oxenham’s (2009) data. In this abbreviated TMC method the signal level would be held constant and the off-frequency masker level varied for several masker-signal delays. The constant signal level would presumably result in a constant amount of compression across the various delays and, therefore, the slope of the recovery function could be viewed as a measure of temporal resolution that is independent of cochlear compression. However, it appears this assumption may not be accurate for high presentation levels (above approximately 85 dB SPL), which may limit the use of this technique for individuals with significant hearing loss.
The degree of MR for pure tone signals will also be measured. Similar MR phenomena have also been seen when presenting tonal stimuli in a modulated background and this has been attributed to underlying cochlear compression (e.g., Oxenham and Dau, 2004). This sub-experiment will explore the relationship between cochlear compression and MR for pure tones in a broadband modulated background, as well as the correlation between the degree of MR for pure tones and MR for speech stimuli. Listeners with a range of degrees of cochlear compression are tested in Chapter 4. The hypothesis is that degree of cochlear compression will be a significant predictor of MR. Subject age was limited to 21-70 years and each HI listener was assigned an age-matched NH control. This is to control for potential age effects that may affect MR (e.g., Dubno et al., 2002; 2003).

In summary, the goal of this thesis is to better understand the relationship between MR and behavioral estimates of basilar membrane compression. To accomplish this, the effects of background noise on cochlear compression estimates in NH listeners was examined (Chapters 2 and 3) to allow for interpretation of the results in the following experiment (Chapter 4) that involves comparisons between noise-masked NH listeners and age-matched HI counterparts. The hypothesis is that differences in underlying cochlear compression seen in HI listeners with sensorineural hearing loss will be significantly correlated with behavioral estimates of compression. A positive finding would suggest development of a clinical tool to estimate cochlear compression would allow for more precise counseling of HI clients as to prediction of their ability to hear well in noisy listening conditions. In addition, better understanding of the mechanisms
underlying reduced MR in the HI population may allow for more appropriate hearing aid design in an attempt to at least partially alleviate these difficulties.
CHAPTER 2: FORWARD GROWTH OF MASKING

I. INTRODUCTION

Cochlear hearing loss leads to a number of changes in auditory perception and performance. Most obvious is a loss of sensitivity to low-level sounds, as evidenced by the audiogram. Beyond absolute threshold, psychoacoustic studies have identified a number of tests in which hearing-impaired (HI) listeners perform differently from normal-hearing (NH) listeners. To rule out effects of signal level and audibility from true supra-threshold deficits, it is necessary to somehow provide equivalent conditions for both NH and HI listeners. Some studies have attempted to equate audibility by presenting the stimuli at the same sensation level (equal SL) for the two groups (Moore and Glasberg, 1988; Lentz and Leek, 2003). Because of the higher absolute thresholds in the HI group, this results in higher sound pressure levels (SPL) for the HI group than for the NH group, meaning that any difference in results could potentially be due to the higher overall sound level. For instance, frequency selectivity in NH listeners becomes poorer at high levels (e.g. Moore and Glasberg, 1987; Nelson et al., 1990), meaning that a difference between NH and HI listeners in frequency selectivity when measured at the same SL may have less to do with an auditory deficit than it does with a level effect that is also found in normal hearing. Other studies have equated overall sound pressure levels (equal SPL) for the two groups (e.g., Buus et al., 1999; Summers, 2001). The potential problem with this approach is that the stimuli are at a lower SL for the HI group than for the NH group, making it difficult to rule out the possibility that differences in performance are due to differences in audibility, rather than any genuine supra-threshold deficit due to the hearing loss.
One way of equating both overall level and audibility has been to present the stimuli in a background of noise that elevates the thresholds of NH listeners to approximate those of HI listeners in quiet. This noise may be designed to raise thresholds in a particular frequency region (e.g. Sommers and Humes, 1993) or over a wide frequency range (e.g. Dubno and Schaefer, 1992; Bacon et al., 1998). The noise may also be spectrally shaped and adjusted in level to specifically match the individual thresholds of certain 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, 1990}. In some situations the 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 situations, some HI listeners exhibit results similar to those found for NMNH listeners, whereas other HI listeners perform more poorly, such as in tasks involving speech understanding in a modulated background (e.g., Bacon et al., 1998; Oxenham and Dau, 2004). For a recent review of studies involving comparisons between HI and NMNH listeners, see (Reed et al., 2009). As pointed out by Reed et al. (2009), when comparing performance of NH and HI listeners, it is also important to control for age differences, as many studies have shown that ageing, independent of hearing loss, also affects many aspects of auditory processing, including temporal processing (e.g. Abel et al., 1990; Moore et al., 1992; Fitzgibbons and Gordon-Salant, 2004; Lister and Roberts, 2005; Dubno et al., 2008).
Aside from the issues of audibility and age, many of the deficits experienced by listeners with cochlear hearing loss are consistent with the changes in basilar-membrane mechanics that occur with outer hair cell damage or dysfunction. The physiological phenomena observed at the level of the basilar membrane (BM) include a loss of sensitivity to low-level sounds, a linearization of the normally compressive input-output function in response to tones near the characteristic frequency (CF), and a broadening of tuning (e.g., Ruggero et al., 1997; Robles and Ruggero, 2001). Potential psychoacoustic correlates of these effects include higher absolute thresholds, loudness recruitment, and poorer frequency selectivity (for a review, see Oxenham and Bacon, 2003).

In general, the effect of using a background masking noise with NH listeners has been interpreted in terms of a reduction in the audibility of the stimuli. An alternative possibility is that the background noise changes the mechanical response of the BM to effectively linearize it so that it resembles that of an impaired cochlea. The particular response changes would depend on the specific underlying cochlear changes. As described in Plack et al. (2004), changes in cochlear compression may be manifested as a reduction in the compression exponent over the entire dynamic range, a reduced level range over which normal compression exponents are found, or a combination of the two. The BM response to wideband noise at a given level has been described as “quasi-linear” (de Boer and Nuttall, 1997; 2000), and it is known that a low-frequency suppressor tone can linearize the response to a higher-frequency probe tone at the characteristic frequency of the point of measurement (e.g., Ruggero et al., 1992). Another possible mechanism by which cochlear response may be linearized is via efferent fibers originating in the medial
olivary complex. Efferent activation is hypothesized to lead to a reduction in the gain of the cochlear amplifier, particularly in response to longer duration stimuli. It is possible that the presence of background noise may initiate gain reduction in this manner, which would result in a less compressive response (Guinan, 2006; Krull and Strickland, 2008).

A recent physiological study addressed the question of whether broadband noise linearizes the BM response to simultaneously presented tones. Recio-Spinoso and Lopez-Poveda (2010) measured BM displacement in a chinchilla cochlea in response to a tone at the characteristic frequency (CF) in the presence of various levels of white noise. They found that the white noise suppressed the response to the tone, especially for low tone levels and reduced the range of levels over which compression was observed.

The aim of the present study was to estimate the effects of background noise on behavioral estimates of compression, using the growth of forward masking (Oxenham and Plack, 1997) in NH listeners in the presence of various levels of threshold-equalizing noise (TEN, Moore et al., 2000). The question was whether the presence of a background noise would result in a more linear BM input-output function for tones, as estimated behaviorally. If so, then similarities between HI and NMNH listeners found in previous studies in tasks such as loudness judgments might be due to changes in BM compression, rather than audibility per se. If background noise does not make the BM input-output function more linear, then any similarities between HI and NMNH listeners are unlikely to be due to noise-induced changes in BM response for the NMNH group.
II. ESTIMATING COMPRESSION AS A FUNCTION OF BACKGROUND NOISE LEVEL

A. Listeners

Six adult listeners (all female) completed the task. 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. Training was provided until stable performance was achieved. No subject required more than 2 hours of training. Five listeners were paid an hourly rate for their participation; S1 was the first author.

B. Procedure

Behavioral techniques that appear to provide a reasonable estimate of cochlear compression include growth of masking (GOM; Oxenham and Plack, 1997) and temporal masking curves (TMC; Nelson et al., 2001), both of which use forward masking to avoid the potentially confounding effects of suppression due to the masker. These behavioral estimates of compression have both been shown to provide results that are in relatively good agreement with direct physiological measurements in other species (e.g., Robles and Ruggero, 2001). Here we used the GOM method.

Forward GOM 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). The three intervals in each trial were separated by 300-ms interstimulus intervals. The pure-tone masker was presented in all three intervals,
whereas the signal was presented in one, chosen at random in each trial. The listener’s
task was to select the interval that contained the signal. The initial step size for adaptively
varying the 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. Subjects completed practice runs until their thresholds showed stability and the
threshold for each condition and each subject was taken as the average from three
subsequent runs. The presentation order of conditions was randomized within and across
subjects and repetitions. Listening sessions were 2 hours in length, including frequent
breaks. The entire experiment took an average of 5 sessions to complete.

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

C. Stimuli

The signal was a 4-ms, 4000-Hz pure tone, gated on and off with 2-ms raised-
cosine ramps (no steady state). The pure-tone masker frequency was either 4000 Hz (on-
frequency masker) or 1800 Hz (off-frequency masker). The masker duration was 200 ms,
including 5-ms raised-cosine onset and offset ramps. The delay between the masker
offset and signal onset was fixed at 0 ms. The relatively high signal frequency allowed
the use of a short signal, while still ensuring that multiple cycles of the stimulus could be
presented, thereby avoiding audible “spectral splatter” (e.g., Leshowitz and Wightman, 1972). Additionally, the short signal made it more likely that the off-frequency masker would provide sufficient masking even at high signal levels. With a 4-kHz signal, the response to the off-frequency masker, more than 1 octave below the signal, is thought to produce a linear response at the BM location with a CF corresponding to the signal frequency (Lopez-Poveda et al., 2003; Rosengard et al., 2005).

A 2/3-octave wide Gaussian noise masker, generated and filtered in the spectral domain and centered at 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 via acoustic and/or electric crosstalk. Additionally, a low-level high-pass noise masker (off-frequency condition) or spectrally notched noise masker (on-frequency condition) was presented to the test ear to reduce “off-frequency listening” (e.g., O’Loughlin and Moore, 1981). For the high-pass noise, the cut-off frequency was set to 1.117 times the signal frequency (4468 Hz). The notched noise incorporated this high-pass noise, as well as a low-pass noise with the cut-off set to 0.883 times the signal frequency (3532 Hz). The noises were generated in the spectral domain, so that only the onset and offset ramps of the noise limited the slopes of the spectral edges. To determine the appropriate levels for the notched and high-pass noise maskers, the levels of these maskers required to just mask the signal at different levels were determined during pilot testing. The notched and high-pass masker levels required to just mask the signal with a given level were obtained from a straight-line fit to these pilot data. The spectrum level of the maskers was then set 25 dB below the levels at which they would mask the signal for both the on-frequency
and off-frequency conditions. No attempt was made to reduce potential “confusion” effects due to having a forward masker and signal at the same frequency (e.g., Neff, 1986). Using the TMC paradigm, Nelson et al. (2001) showed that confusion may result in somewhat lower masker levels, but that the derived compression exponents were essentially not affected.

Forward GOM functions were measured in quiet and in the presence of the TEN at levels of 20, 30, 40, and 50 dB SPL within the equivalent rectangular bandwidth of an auditory filter (ERB_N) centered around 1 kHz (Moore et al., 2000). Because TEN is designed to produce equal masked thresholds (in dB SPL) for pure tones at all frequencies, the noise can be thought of as simulating a relatively flat hearing loss (when measured in SPL). The passband of the TEN extended from 20 to 20,000 Hz. The TEN was presented throughout each trial by gating it on 300 ms before the beginning of the first interval and gating it off 300 ms after the end of the third interval. The notched, high-pass, and contralateral maskers, when present, were gated on and off with the TEN, and all had raised-cosine onset and offset ramps of 5 ms. Signal levels ranged from 35 to 85 dB SPL (but were fixed within a given run). Only thresholds for signal levels where the tonal masker produced adequate masking (i.e. masked thresholds that were greater than 5 dB above the threshold in the TEN, as estimated from separate threshold measurements for the signal in the presence of the background TEN alone), were included in the plots.
Individual and mean thresholds for the signal measured in quiet and various levels of TEN are shown in Table 2.1. When the TEN level was raised from 20 to 30 dB SPL/ERB<sub>N</sub>, thresholds increased by approximately 10 dB. When the TEN level was increased from 40 to 50 dB, the thresholds increased by approximately 13 dB. The greater increase in masked threshold between and 40 and 50 dB/ERB<sub>N</sub> may in part reflect the broadening of auditory filters at high levels, leading to greater noise power within the auditory filter. However, wider auditory filters are unlikely to account for the entire 3 dB increase in signal-to-noise ratio at threshold, as this would imply a doubling in filter bandwidth with a 10-dB in noise level. Individual forward GOM data for each of the six subjects are shown in Fig. 2.1. The masker level required to just mask the signal is plotted as a function of the signal level. Open and filled symbols represent thresholds in the presence of the on- and off-frequency pure-tone maskers, respectively. Different symbols represent thresholds in different levels of TEN, as indicated in the legend. The vertical lines represent thresholds for the signal in the different levels of TEN (from left to right, the lines represent threshold in quiet, in 20 dB TEN, in 30 dB TEN, etc.).
FIG. 2.1. Individual forward GOM functions for 6 NH listeners. Vertical lines represent signal threshold in various levels of TEN alone (from left to right: no TEN, and 20, 30, 40, and 50 dB/ERB50). Each line pattern represents a given TEN level. Filled symbols represent threshold off-frequency masker levels; open symbols represent on-frequency masker levels at threshold. Error bars represent ±1 standard deviation. For reference, the dashed line on the major diagonal has a slope of 1.0, indicating linear growth of masking.

When the masker and signal were at same frequency, and similar levels, both stimuli would presumably be subjected to the same amount of compression at the signal-frequency location on the BM and the resulting GOM function would be expected to have a slope close to 1.0 (Plack and Oxenham, 1998). In the off-frequency condition, it is assumed that the masker is not subject to compression at the place with a CF corresponding to the signal frequency, while the signal continues to be compressed.
Therefore the relationship between the slopes of the on- and off-frequency functions is assumed to reflect the underlying BM compression function (Oxenham and Plack, 1997). If the addition of the TEN resulted in less compression applied at the signal location on the BM, then we would expect to see the slopes of the off-frequency functions becoming gradually more linear (steeper) as the TEN level increased. We would not expect to see a change in slope for the on-frequency functions if the TEN resulted in less compression, because both the signal and the masker would be equally subjected to less compression. On the other hand, the TEN might be expected to have some effects on masker level at threshold that are not related to BM compression. For instance, when the signal is presented only a few dB above its threshold in TEN, then the TEN is likely to contribute to the masking of the signal, leading to a lower (tone) masker level at threshold. In order to reduce the effects of this “additivity of masking” (e.g., Penner, 1980; Humes and Jesteadt, 1989; Oxenham and Moore, 1995), only signal levels that were at least 5 dB above their threshold in the TEN were used in the analysis.

The data in Fig. 2.1 display some differences between subjects. For instance, the absolute threshold for the signal (solid vertical line in each panel) was lower for some subjects (e.g., S1) than for others (e.g., S5); however, thresholds in the presence of TEN alone (dashed or dotted vertical lines) were relatively similar for all six subjects. On-frequency masker levels at threshold (open symbols) also differ between subjects, with some subjects (e.g., S1) showing less susceptibility to masking (i.e. higher masker levels at threshold) than others (e.g., S6). Overall, though, the patterns of results are reasonably similar across subjects, with the GOM functions much shallower for the off-frequency
masker than for the on-frequency masker, as found in earlier studies (e.g., Oxenham and Plack, 1997). In particular, based on the individual data, it is difficult to discern any systematic effects of TEN level on either the on- or off-frequency masker levels at threshold.

FIG. 2.2. As Fig. 2.1 but showing GOM functions averaged across 6 NH listeners.

To determine whether more noticeable trends emerged for the listener group as a whole, the data from the six listeners were averaged, and are plotted in Fig. 2.2. For the off-frequency masker functions (filled symbols), where changes in compression would be most likely to have an effect, the data from all the functions lie essentially on top of each other, with no visible effect of TEN level. For the on-frequency functions (open symbols), the data lie close to the main diagonal, suggesting near-linear GOM. Again, for the most part the functions overlap with each other. An exception involves the left-most
data point in each function, where the masker level at threshold is typically somewhat lower than for the masker levels in the presence of lower TEN levels. This effect may involve the additivity of masking, as mentioned above, and suggests that our limit of including only signal levels that were at least 5 dB above masked threshold in the TEN may not have been stringent enough to eliminate such effects.

Plotting the off-frequency masker levels as a function of the on-frequency masker level at each signal level provides a behavioral estimate of the BM input-output function, if it is assumed that the BM response to the off-frequency masker is linear (e.g. Plack et al., 2004). Fig. 2.3 shows these derived input-output functions for the six subjects individually. The input-output functions derived from the mean data across the six subjects are shown in Fig. 2.4. As in Fig. 2.1, the different symbols represent different levels of TEN. Note that, as the TEN level increases, the total number of signal levels available decreases (due to the increase in signal threshold in TEN). In fact, the data for the 50-dB/ERB_N TEN were not used to construct the I/O function in Figs. 2.3 and 2.4 because there were too few data points to permit further analysis.
FIG. 2.3. Individual input-output functions derived using data from Figure 2.1. Off-frequency masker levels at threshold are plotted as a function of on-frequency masker level at threshold for the same signal levels at each TEN masker level separately. The different symbols denote different TEN levels, as in Fig. 2.1. Only data shown as open symbols were used to calculate the function slopes shown in Table 2.1. For reference, the dashed diagonal line in each panel has a slope of 1.0, indicating linear growth.

FIG. 2.4. As Fig. 2.3 but showing average input-output functions derived using data from Figure 2.2.
To compare compression estimates across the different TEN levels, a straight line was fitted to the data points for the five highest levels in each function. These points were chosen because they were measurable for all levels of TEN below 50 dB/ERB$_N$. allowing a fair comparison of slopes across different levels of TEN. The truncated functions used to obtain these slopes are shown as open symbols in Figs. 2.3 and 2.4. Bootstrapping was used to determine the individual slope values for each listener, as well as the associated variability. For each TEN level and tonal masker (i.e. on and off-frequency) individual threshold estimates were selected randomly for each of the signal levels and a slope was derived using orthogonal regression. This was repeated 10,000 times for each condition and the resulting values were used to construct a histogram. The mean of the histogram was used as the slope value for that particular condition and the deviation from the mean was used to construct the 95% confidence intervals. The decision to use orthogonal regression instead of simple linear regression was based on the fact that the data to be fitted incorporated variability along both the x- and the y-axes. Linear regression assumes the data vary only in one dimension (the y-axis) and is designed is to minimize deviation between the fitted line and the data in this dimension only. Orthogonal regression minimizes the deviation between the data and the fitted line in both dimensions. Linear regression on data that have variability along both axes can be shown to underestimate the underlying slope, resulting in slight overestimates of compression. The derived individual slope values are shown in Table 2.2, along with the 95% confidence intervals. The range of individual slope values for these 6 listeners is consistent with that from previous studies using similar measurement methods (e.g.,
Gifford and Bacon, 2005; Rosengard et al., 2005). Table 2.2 also includes the slope values obtained by averaging across the slope values from the individual subjects (Mean slope), as well as the slope values derived by fitting lines to the average data shown in Fig. 2.4 (Slope of mean data). Reassuringly, these two approaches yielded very similar slope values, which are highly compressive and are broadly consistent with those found in other experiments using NH listeners tested in quiet (e.g., Oxenham and Plack, 1997).

The aim of this study was to test whether background noise affected BM compression, as estimated behaviorally. This question was addressed using a repeated-measures analysis of variance (RMANOVA) with estimated input-output slope value as the dependent variable and TEN level as the independent variable, with values of no noise, and TEN levels of 20, 30 and 40 dB SPL/ERB_N. Overall, no significant effect of TEN on the estimated compression exponent was observed \[ F(1,3) = 0.417, p = 0.743 \]. Of course, average data may obscure consistent trends within individual. However, in each case, the confidence intervals around each individual slope estimate are large, such they generally overlap across all TEN levels. Thus, there is no evidence of a trend towards systematic changes in slope with increasing TEN level.

In summary, aside from restricting the range of signal levels that could be tested, the presence or level of the background noise did not significantly influence the compression exponent estimates obtained in a forward GOM paradigm.
### Table 2.2

<table>
<thead>
<tr>
<th></th>
<th>QUIET</th>
<th>20 TEN</th>
<th>30 TEN</th>
<th>40 TEN</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.597</td>
<td>0.176</td>
<td>0.536</td>
<td>0.283</td>
</tr>
<tr>
<td></td>
<td>(0.34; 0.85)</td>
<td>(0.07; 0.29)</td>
<td>(0.23; 0.84)</td>
<td>(0.19; 0.37)</td>
</tr>
<tr>
<td>S2</td>
<td>-0.008</td>
<td>0.032</td>
<td>-0.035</td>
<td>0.161</td>
</tr>
<tr>
<td></td>
<td>(-0.12; 0.10)</td>
<td>(-0.03; 0.10)</td>
<td>(-0.08; 0.01)</td>
<td>(0.05; 0.27)</td>
</tr>
<tr>
<td>S3</td>
<td>0.051</td>
<td>0.348</td>
<td>0.272</td>
<td>0.388</td>
</tr>
<tr>
<td></td>
<td>(-0.14; 0.24)</td>
<td>(0.18; 0.51)</td>
<td>(0.14; 0.41)</td>
<td>(0.34; 0.44)</td>
</tr>
<tr>
<td>S4</td>
<td>0.562</td>
<td>0.435</td>
<td>0.163</td>
<td>0.437</td>
</tr>
<tr>
<td></td>
<td>(0.14; 0.98)</td>
<td>(0.26; 0.61)</td>
<td>(-0.09; 0.41)</td>
<td>(0.23; 0.65)</td>
</tr>
<tr>
<td>S5</td>
<td>0.124</td>
<td>0.289</td>
<td>0.134</td>
<td>0.260</td>
</tr>
<tr>
<td></td>
<td>(-0.13; 0.38)</td>
<td>(0.03; 0.55)</td>
<td>(0.03; 0.23)</td>
<td>(0.13; 0.39)</td>
</tr>
<tr>
<td>S6</td>
<td>0.555</td>
<td>0.219</td>
<td>0.314</td>
<td>0.296</td>
</tr>
<tr>
<td></td>
<td>(-0.23; 0.88)</td>
<td>(-0.08; 0.51)</td>
<td>(-0.30; 0.93)</td>
<td>(0.13; 0.46)</td>
</tr>
<tr>
<td>Slope of mean data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.302</td>
<td>0.242</td>
<td>0.201</td>
<td>0.304</td>
</tr>
<tr>
<td>Mean slope</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.313</td>
<td>0.250</td>
<td>0.231</td>
<td>0.304</td>
</tr>
<tr>
<td></td>
<td>(-0.25; 0.87)</td>
<td>(-0.03; 0.53)</td>
<td>(-0.15; 0.61)</td>
<td>(0.11; 0.50)</td>
</tr>
</tbody>
</table>

The 95% confidence intervals are given in parentheses.

### III. DISCUSSION

The results indicate that estimates of compression for NH listeners in a forward GOM paradigm are not significantly affected by background noise. The average slope values for the data in Fig. 2.2, obtained over the range of signal levels from 65 to 85 dB SPL, fall between 0.20 and 0.30, which is in reasonably good agreement with earlier behavioral estimates of compression in quiet from NH listeners using a variety of techniques, which are typically around 0.2 (e.g., Oxenham and Moore, 1995; Oxenham and Plack, 1997; Nelson et al., 2001; Lopez-Poveda et al., 2003 etc.). Thus, the results suggest that BM compression exponents are not affected by noise, at least in the conditions that were measurable here, where the probe was at least 5 dB above its threshold in the noise.
As noted in the introduction, it has been proposed that efferent fibers may function to linearize the response of the cochlea, especially with longer duration stimuli (Guinan, 2006; Krull and Strickland, 2008). The present results do not demonstrate linearization of BM response in the presence of TEN and, therefore, do not provide any evidence for efferent activation. It is possible that either the TEN did not produce an efferent response that changed the BM response or that any efferent-driven response changes were not measurable in our data because they were restricted to lower stimulus levels than were used in the analysis.

Our conclusion can be compared with that of Recio-Spinosa and Lopez-Poveda (2010), based on direct measures of BM response in a chinchilla cochlea. They showed that the BM response to a tone was reduced (suppressed) by the presence of a broadband noise, which at first glance might appear to contradict our conclusion. However, a number of factors make a direct comparison difficult. First, their analysis of the responses involved a Fourier transform of the vibration, and the extraction of only the Fourier component corresponding to the tone’s frequency. While this is common practice in BM studies, it makes a direct comparison with auditory-nerve or psychophysical suppression studies difficult: auditory-nerve studies report changes in overall firing rate (rather than the response locked to a particular frequency); psychophysical studies use masking, where the noise also contributes to the masking, making the outcome more comparable to the overall-response measure used in auditory-nerve studies than the frequency-specific measure used in BM studies. Second, the noise levels required to show suppression in the Recio-Spinosa and Lopez-Poveda (2010) study were much higher than were used here.
Assuming the noise level was adjusted in 10-dB steps, the lowest noise level to produce measurable suppression in one animal (m31) was about 90 dB SPL. Assuming a noise bandwidth of 20 kHz, this corresponds to a level of more than 75 dB SPL within an ERB\textsuperscript{N} centered around the CF of 9.75 kHz. No suppression was observed for signal levels of 75 dB SPL or more, suggesting that the noise level within the ERB\textsuperscript{N} had to exceed the level of the signal for suppression to be observed. When viewed in this light, the results of Recio-Spinosa and Lopez-Poveda (2010) are consistent with ours in showing no suppression in conditions where the signal level exceeds the level of the background noise.

The conclusion that suppression is only observed when the noise level exceeds that of the signal may be understood in terms of the underlying mechanisms, as follows. When the signal level exceeds the noise level, it is likely that the signal, rather than the noise determines the effective gain of the cochlear amplifier in the region with a CF corresponding to the signal frequency. If the signal determines the gain, then the noise would not be expected to affect compression (see, e.g., Pang and Guinan, 1997). In contrast to studies of BM motion, it is not clear how to evaluate compression at lower signal levels psychophysically, because the signal then approaches its masked threshold in the noise alone, making it difficult to distinguish the effects of the sinusoidal masker from those of the TEN.

It therefore remains possible, indeed likely, that the noise dominates, and suppresses, the response to the signal at lower signal-to-noise ratios, where the gain is determined by the noise. Nevertheless, for the practical purposes of psychoacoustic and
speech experiments, as well as for real-world applications, our results suggest that the
addition of a broadband background noise does not change the underlying BM responses
to sounds that are presented at positive signal-to-noise ratios measured within one ERB_N.

One potential complication in our results involves the presence of the notched or
highpass noise, designed to limit the possibilities for off-frequency listening. This noise
was present in all our conditions, and could in principle have swamped any additional
effect of the TEN noise. This seems unlikely, however, as the level of the notched or
highpass noise was always much lower than that of the signal and in most cases was
lower than that of the TEN. For instance, at a signal level of 55 dB SPL, the average
spectrum level of the notched and highpass noise was around 0 dB (re 20 μPa). The
ERB_N at 4 kHz is estimated to be around 460 Hz. Thus, the level per ERB_N of the
notched or highpass noise in its passband near the signal frequency would be around 27
dB SPL, which is nearly 30 dB lower than the signal level, meaning it was unlikely to
have directly affected signal thresholds. It is also lower than both the 30 and 40 dB/ERB_N
TEN levels, meaning that it was unlikely to have affected, or even been audible in, the
critical conditions of high TEN levels.

Overall, our results suggest that any similarities in performance between NH and
NMNH listeners in most psychoacoustic and speech tasks are unlikely to be due to
changes in the underlying BM compression function caused by the presence of threshold
elevating background noise. Statistical analysis of the slopes derived from the individual
and mean data in various levels of TEN demonstrated that the presence of the TEN
masker did not significantly alter the estimates of the BM compression exponents. This
conclusion leaves open the possibility that any differences observed between HI and NMNH listeners are due to differences between these two groups in terms of the underlying BM mechanics (for a review, see Oxenham and Bacon, 2003).

At least one common feature of hearing impairment that can be recreated in NMNH listeners is loudness recruitment, which has been hypothesized to be due to lack of cochlear compression in HI listeners (e.g., Moore and Glasberg, 1997; Moore et al., 1999b). The results from the present study suggest that if loudness recruitment in HI listeners is indeed due to loss of compression, then the mechanisms underlying loudness recruitment in NMNH (e.g. Steinberg and Gardner, 1937) and HI listeners are different. This conclusion is consistent with the way in which partial loudness is explained in the model of Moore et al. (1997). It is also consistent with the physiological results of Phillips (1987). He measured neural response functions in an animal model (cat) with thresholds that were elevated either by permanent (noise-induced) cochlear hearing loss or by a simultaneously presented noise. Philips found that the effect on the neural responses was not similar across the two groups, and concluded that loudness recruitment in cochlear hearing loss did not have the same peripheral basis as loudness recruitment produced by partial masking (Scharf, 1964). A study by Heinz (2005), which evaluated neural responses to various stimuli in an animal model with cochlear hearing loss (due to noise trauma), has also questioned whether loudness recruitment in HI listeners is due to reduced BM compression. They found that auditory-nerve rate-intensity functions, which presumably reflect BM motion, were not consistent with expectations based on recruitment. However, the link between the physiology and human perception remains
tenuous, due to species differences and uncertainty regarding how responses in the auditory nerve relate to the percept of loudness (Doucet and Relkin, 1997).

As noted in the introduction, studies have shown mixed results when attempting to simulate the performance of HI listeners with NMNH listeners. For example, masking release in a temporally varying background, as measured by Bacon et al. (1998), was well-simulated by NMNH listeners for approximately half of the HI listeners, while the remaining HI listeners showed deficits beyond those of their NMNH counterparts. Another study, measuring effects of masker phase curvature on signal thresholds, found substantial differences between most of the HI listeners and their NMNH counterparts (Oxenham and Dau, 2004). Earlier psychophysical results suggested that moderate-to-severe cochlear hearing loss was often sufficient to eliminate cochlear compression, as measured psychophysically (Oxenham and Moore, 1995; Oxenham and Plack, 1997; Moore et al., 1999b). More recent results suggest that milder hearing loss can result in normal values of maximum compression, but over a more limited range of levels than normal and that the amount of compression can vary, even for the same amount of hearing loss (e.g., Moore et al., 1999b; Plack et al., 2004; Lopez-Poveda et al., 2005). These findings suggest that the residual differences between some HI and NMNH listeners may be indicative of loss of compression for the HI ears. In Bacon et al. (1998), the results from the HI listeners that were accounted for by the NMNH results may have been representative of HI ears with near-normal amounts of compression while the HI results that showed additional deficits may reflect a loss of cochlear compression.
Further research will be necessary to determine if degree of cochlear compression can indeed account for variation in HI listeners’ performance on these types of tasks.

In summary, the results from this study indicate that the addition of masking noise to elevate thresholds in NH listeners does not alter the compressive characteristics of the BM response for tones 5 dB or more above their masked threshold in the noise, as estimated behaviorally using a forward GOM paradigm. It therefore remains possible that differences in underlying cochlear compression account for at least some of the differences in performance seen between NMNH and HI listeners.
CHAPTER 3: ADDITIVITY OF FORWARD MASKING

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). Furthermore, 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 individual 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 speech understanding or tone detection in a modulated background (e.g., Bacon et al., 1998; Oxenham and Dau, 2004). Reed et al. (2009) provides 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 (e.g., 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 at the signal place 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 Arifianto, 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 non-simultaneous 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 Arifianto, 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 non-overlapping non-simultaneous 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; Cokely and Humes, 1993; Oxenham and Moore, 1995; Plack and O'Hanlon, 2003; Plack et al., 2008; Plack and Arifianto, 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 above, 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. 3.1, panel A). The signal levels were based on each listener’s thresholds for the signal stimuli in a background noise (threshold elevating 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), so 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 3.1.

![Fig. 3.1 Depiction of stimuli for the two measurement phases of the additivity task.](image)

**Table 3.1**: Spectrum levels of M1 and M2 (dB SPL), averaged across S2 and S3, as derived in Phase 1. Values in parentheses are standard deviations.
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. 3.1, panel 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 5 sessions 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 ear of subjects via Sennheiser HD580 earphones. 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 filtered 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. 3.2, 3.3, and 3.4 for signal frequencies of 500, 1500, and 4000 Hz, respectively. The data from each subject are shown in the individual panels, along with the mean data in 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 alone. As can be seen for some listeners, such as S2 and S3 in Fig. 3.2, S2 in Fig. 3.3, and S1, S2, and S6 in Fig. 3.4, 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. 3.4) 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 (for example, see S3 in Fig. 3.2).
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).
Fig. 3.2 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 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.
Fig. 3.3 As for Fig. 3.2, but for a 1500-Hz signal.
Fig. 3.4 As for Fig. 3.2, but for a 4000-Hz signal.

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 equation to be solved:

\[ S_{M1+M2}^c = S_{M1}^c + S_{M2}^c - S_T^c, \]  
(Eq. 3.1)

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 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 by 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. 3.2-3.4) 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 3.2. The dotted line indicates the predicted thresholds for the combined masker condition that would be expected for linear summation in the absence of compression.

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

**Table 3.2:** 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.

As can be seen from the first column of Table 3.2, 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 3.2 shows the compression exponents for the 4000-Hz signal frequency. Only data from 2 listeners (S1 and S2) resulted in compressive estimates ($c < 1$); the remaining exponents are consistent with essentially linear responses (with respect to intensity), or 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. 3.2-3.4 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 (ANOVA), 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-Hz 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 non-parametric tests.)

<table>
<thead>
<tr>
<th>Listener</th>
<th>Compression estimate for 4000 Hz signal</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.23</td>
</tr>
<tr>
<td>S2</td>
<td>0.16</td>
</tr>
<tr>
<td>S3</td>
<td>0.39</td>
</tr>
<tr>
<td>S4</td>
<td>0.44</td>
</tr>
<tr>
<td>S5</td>
<td>0.26</td>
</tr>
<tr>
<td>S6</td>
<td>0.30</td>
</tr>
<tr>
<td><strong>Mean data</strong></td>
<td><strong>0.30</strong></td>
</tr>
</tbody>
</table>

Table 3.3: 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.

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 4000 Hz. As can be seen by comparing the compression estimates from column 3 of Table 3.2 with those in Table 3.3, the results obtained using
the AFM technique indicate less compression than those obtained using forward GOM, even 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 approximately 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 approximately 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, while 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 6 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 manipulations 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. 3.5. 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 3.6 shows the data for low-level stimuli measured in quiet. Note that S4 was not available to run the low-level condition.
Fig. 3.5 Additivity results for a 4000-Hz signal using the same M1 and M2 levels as for Fig. 3.4, 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 text for further details.
Fig. 3.6 Thresholds for a 4000-Hz signal using lower signal levels in the absence of TEN. Symbol legend same as for previous figures.

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. 3.5, the abscissa
shows the nominal signal levels in dB SL, relative to thresholds in the presence of TEN, for ease of comparison with Fig. 3.4. The derived compression exponents are shown in Table 3.4.

<table>
<thead>
<tr>
<th>Listener</th>
<th>4000 Hz signal</th>
<th>4000 Hz signal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No TEN</td>
<td>Low SPL/no TEN</td>
</tr>
<tr>
<td>S1</td>
<td>0.34</td>
<td>0.26</td>
</tr>
<tr>
<td>S2</td>
<td>0.29</td>
<td>0.11</td>
</tr>
<tr>
<td>S3</td>
<td>0.33</td>
<td>0.26</td>
</tr>
<tr>
<td>S4</td>
<td>0.58</td>
<td>--</td>
</tr>
<tr>
<td>S5</td>
<td>1.09</td>
<td>1.96</td>
</tr>
<tr>
<td>S6</td>
<td>1.20</td>
<td>0.37</td>
</tr>
<tr>
<td>Average fit to mean data</td>
<td>0.44</td>
<td>0.26</td>
</tr>
<tr>
<td>Compression exponent averaged across listeners</td>
<td>0.64</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Table 3.4. 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.

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 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 3.3
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 in the absence of TEN. Note that S4 was excluded from this
analysis as she did not have results for both conditions. The results were 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, while the trend towards 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).

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 approximately 60 dB SPL. This suggests that the present data, with minimum signal levels of approximately 60 dB SPL due to the presence of the TEN, would not have been influenced by efferent effects. A recent study by Plack and Arifianto (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 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 Arifianto (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 towards 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 approximately 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 Arifianto (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 about 70 to 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 4 kHz 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 non-compressive results for HI listeners when measured at similarly high stimulus levels.
CHAPTER 4: MASKING RELEASE

I. INTRODUCTION

Normal-hearing (NH) listeners experience an improvement in speech understanding when they are listening in a background noise that is either temporally or spectrally modulated, compared to their performance when the noise is unmodulated. This improvement has been termed “masking release” (MR). MR is thought to reflect the ability of listeners to take advantage of the improved signal-to-noise ratio (SNR) in the dips of the fluctuating masker. Listeners with cochlear hearing loss generally show less, or even no, MR under similar listening conditions (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; Buss et al., 2009; Desloge et al., 2010). Several studies have evaluated various possible explanations for this lack of MR, including audibility, temporal resolution, spectral resolution, and cochlear compression.

One of the most obvious reasons for the reduced ability of hearing-impaired (HI) listeners to take advantage of spectral or temporal masker dips is the reduced audibility of the speech within the dips as a result of the hearing loss. Several studies have attempted to control for the reduced audibility of 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., 1999a). 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 was 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.

Festen and Plomp (1990) suggested that a lack of MR in their HI listeners may be due to abnormal temporal resolution, specifically the “persistence of forward masking.” George et al. (2006) concluded that temporal resolution (measured by estimating the temporal window width centered around 1000 Hz) was correlated with degree of MR for their HI listeners. However, this conclusion was not supported by Jin and Nelson (2006), who did not find a strong relationship between a 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 is likely to be affected by changes in cochlear compression as well as any underlying changes in temporal resolution (Oxenham and Moore, 1997).

Cochlear compression and sharp frequency selectivity are both believed to be mediated by the functioning of the outer hair cells in the cochlea (e.g., Ruggero and Rich, 1991; Zheng et al., 2000). Loss of outer hair cell function leads to less gain at low sound levels, broader cochlear tuning, and a less compressive input-output function for tones at or near the CF of the place of measurement along the basilar membrane (BM). Using behavioral measures, Moore (1999b) found a strong and significant relationship between frequency selectivity and estimated compression in a group of HI listeners. It is possible
that the loss of low-level amplification, due to damage to outer hair cell function, may impair the ability of HI listeners to hear speech in the dips of a temporally modulated masker. This is because relatively low-level speech components present in the dips would not be amplified to the same extent as for a NH listener. In addition, poorer spectral resolution due to broader auditory filter bandwidths may result in decreased ability to separate speech from noise.

Jin and Nelson (2010) found a small but significant relationship between spectral resolution and performance in gated speech. They also noted that performance for gated speech was highly correlated with degree of MR found for speech in a modulated background. Another study (George et al., 2006) evaluated measures of spectral resolution as they related to MR for speech in an amplitude-modulated noise. Aside from level effects (deteriorating spectral resolution seen with increasing presentation level for both NH and HI listeners), they did not find any significant correlation with MR.

There is some limited support in the literature demonstrating a potential relationship between cochlear compression and MR. Moore et al. (1999a) 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. Fast-acting compression would result in gain changes at the syllabic level and could be considered grossly analogous to cochlear compression. Rosengard (2004) attempted to determine if there was a relationship between cochlear compression and MR for 5 HI listeners. Based on this limited number of subjects, her results suggested that behavioral estimates of compression and MR may be correlated. However, her stimuli
were bandlimited from 1000-4000 Hz, which does not reflect the more broadband spectrum that everyday speech possesses. In addition, the results from Jin and Nelson (2010) suggest that low and mid-frequency thresholds are important potential predictors of MR-like performance on a gated-speech task. Specifically, they found that their listeners with poorer low- and mid-frequency audiometric thresholds showed poorer performance with gated speech stimuli than did those with only high-frequency hearing loss. These results mirrored the results they obtained when measuring MR in gated noise, leading them to suggest that reduced MR may be due to reduced ability to separate the masker from the speech fluctuations, because of missing or degraded low-frequency information in the temporal fine structure. However, studies by Oxenham and Simonson (2009) and Buss et al. (2009) found similar degrees of MR for both low- and high-pass filtered speech, suggesting that low frequency information is not necessarily dominant in determining MR.

Studies have shown that behavioral estimates of cochlear compression can vary amongst HI listeners, especially those with mild-to-moderate losses (e.g., Plack et al., 2004; Lopez-Poveda et al., 2005). It does not appear possible to accurately predict the degree of underlying cochlear compression based on the audiogram alone (Plack et al., 2004; Lopez-Poveda et al., 2005) and, therefore, residual cochlear compression may explain the differences in MR seen among HI listeners with similar audiograms. It is possible that the HI listeners for whom performance cannot be imitated using a NH listener with artificially elevated thresholds may be experiencing the effects of reduced cochlear compression. This reduced compression may also result in poorer spectral
resolution, which Jin and Nelson (2010) found as a predictor of MR-like performance in their measurements of gated speech understanding.

A recent alternative approach has suggested that HI listeners may not exhibit abnormal MR, once overall SNR in steady-state noise is taken into account (Bernstein and Grant, 2009; Bernstein and Brungart, 2011). These authors noted that the highest magnitude of MR for both HI and NH listeners is seen 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.

In the current study, a group of age-matched NH listeners served as controls and had their thresholds artificially elevated to control for audibility across the two groups. 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. In addition, George et al. (2006) found a significant effect of age on their measures of temporal resolution. Our goal was to determine how cochlear compression and temporal resolution are related to measures of MR, both for pure tones and for speech stimuli. Various experiments measured MR for sentences in listeners with hearing loss and their age- and audibility- matched NH counterparts. The relationship between the degree of MR for speech and estimates of
compression using the temporal masking curve (TMC) method (Nelson et al., 2001) was explored. Other factors that were explored as possible predictors of MR included recovery from forward masking, as a measure of temporal resolution, and MR for pure tone stimuli, as a measure of dynamic audibility.

II. EXPERIMENT 1: MASKING RELEASE WITH SPEECH

A. Listeners

Twenty-four listeners participated in the study. Twelve had varying degrees of cochlear hearing loss (as confirmed by air and bone conduction audiometry, as well as tympanometry), and their ages ranged from 21 to 69 years (mean age of 50.2 years). The remaining twelve listeners had NH for at least one ear (which was used for testing). Their ages ranged from 23-69 years (mean age of 50.1 years). 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 4.1. Two hours of training was provided for each listener prior to data collection. All listeners were paid an hourly rate for their participation.
<table>
<thead>
<tr>
<th></th>
<th>Age</th>
<th>TEN</th>
<th>500 HL</th>
<th>1000 HL</th>
<th>2000 HL</th>
<th>4000 Hz</th>
<th>500 Q</th>
<th>1500 Q</th>
<th>4000 Q</th>
<th>AVG Q</th>
<th>500 TEN</th>
<th>1500 TEN</th>
<th>4000 TEN</th>
<th>TEN AVG</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI 1</td>
<td>50</td>
<td>60</td>
<td>40</td>
<td>55</td>
<td>60</td>
<td>65</td>
<td>56.4</td>
<td>65.4</td>
<td>64.4</td>
<td>62.1</td>
<td>74</td>
<td>77.3</td>
<td>73.7</td>
<td>75.0</td>
</tr>
<tr>
<td>NMNH1</td>
<td>52</td>
<td>55</td>
<td>5</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>33.4</td>
<td>31.8</td>
<td>25</td>
<td>30.1</td>
<td>77.2</td>
<td>77.7</td>
<td>72.9</td>
<td>75.9</td>
</tr>
<tr>
<td>HI 2</td>
<td>45</td>
<td>45</td>
<td>15</td>
<td>20</td>
<td>25</td>
<td>40</td>
<td>34.4</td>
<td>52.9</td>
<td>49.8</td>
<td>45.7</td>
<td>63.7</td>
<td>65.3</td>
<td>58.8</td>
<td>62.6</td>
</tr>
<tr>
<td>NMNH2</td>
<td>50</td>
<td>45</td>
<td>5</td>
<td>15</td>
<td>20</td>
<td>15</td>
<td>40.1</td>
<td>40.9</td>
<td>33.8</td>
<td>38.3</td>
<td>60.1</td>
<td>62.2</td>
<td>60.8</td>
<td>61.0</td>
</tr>
<tr>
<td>HI 3</td>
<td>53</td>
<td>50</td>
<td>10</td>
<td>5</td>
<td>30</td>
<td>50</td>
<td>23.2</td>
<td>28</td>
<td>56.7</td>
<td>36.0</td>
<td>75.1</td>
<td>71.2</td>
<td>67.8</td>
<td>71.4</td>
</tr>
<tr>
<td>NMNH3</td>
<td>56</td>
<td>50</td>
<td>15</td>
<td>10</td>
<td>0</td>
<td>20</td>
<td>33.1</td>
<td>26.7</td>
<td>27.9</td>
<td>29.2</td>
<td>66.4</td>
<td>77.9</td>
<td>66</td>
<td>70.1</td>
</tr>
<tr>
<td>HI 4</td>
<td>59</td>
<td>55</td>
<td>20</td>
<td>30</td>
<td>40</td>
<td>50</td>
<td>39.4</td>
<td>44.3</td>
<td>53</td>
<td>45.6</td>
<td>73.3</td>
<td>72.1</td>
<td>70</td>
<td>71.8</td>
</tr>
<tr>
<td>NMNH4</td>
<td>51</td>
<td>50</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>25.9</td>
<td>29.2</td>
<td>24.1</td>
<td>26.4</td>
<td>70.1</td>
<td>70.9</td>
<td>73.1</td>
<td>71.4</td>
</tr>
<tr>
<td>HI 5</td>
<td>52</td>
<td>45</td>
<td>15</td>
<td>25</td>
<td>35</td>
<td>35</td>
<td>34.2</td>
<td>40.8</td>
<td>45.8</td>
<td>40.3</td>
<td>60.6</td>
<td>63.4</td>
<td>59.4</td>
<td>61.1</td>
</tr>
<tr>
<td>NMNH5</td>
<td>61</td>
<td>45</td>
<td>10</td>
<td>15</td>
<td>10</td>
<td>20</td>
<td>24.3</td>
<td>25.6</td>
<td>27.3</td>
<td>25.7</td>
<td>63.3</td>
<td>65.2</td>
<td>61.7</td>
<td>63.4</td>
</tr>
<tr>
<td>HI 6</td>
<td>21</td>
<td>60</td>
<td>55</td>
<td>45</td>
<td>65</td>
<td>60</td>
<td>69.3</td>
<td>62.9</td>
<td>72.9</td>
<td>68.4</td>
<td>82.9</td>
<td>80.1</td>
<td>77</td>
<td>80.0</td>
</tr>
<tr>
<td>NMNH6</td>
<td>25</td>
<td>60</td>
<td>10</td>
<td>10</td>
<td>15</td>
<td>15</td>
<td>31.8</td>
<td>24.9</td>
<td>28.3</td>
<td>28.3</td>
<td>79.3</td>
<td>84</td>
<td>76.3</td>
<td>79.9</td>
</tr>
<tr>
<td>HI 7</td>
<td>23</td>
<td>60</td>
<td>55</td>
<td>60</td>
<td>50</td>
<td>10</td>
<td>78</td>
<td>72</td>
<td>26.8</td>
<td>58.9</td>
<td>78</td>
<td>78.7</td>
<td>76.3</td>
<td>77.7</td>
</tr>
<tr>
<td>NMNH7</td>
<td>23</td>
<td>55</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-10</td>
<td>23.1</td>
<td>28.9</td>
<td>23.3</td>
<td>25.1</td>
<td>74.8</td>
<td>81.7</td>
<td>74.6</td>
<td>77.0</td>
</tr>
<tr>
<td>HI 8</td>
<td>57</td>
<td>60</td>
<td>25</td>
<td>40</td>
<td>40</td>
<td>45</td>
<td>37</td>
<td>54.4</td>
<td>65</td>
<td>52.1</td>
<td>77.1</td>
<td>76.7</td>
<td>72.2</td>
<td>75.3</td>
</tr>
<tr>
<td>NMNH8</td>
<td>56</td>
<td>55</td>
<td>10</td>
<td>5</td>
<td>0</td>
<td>10</td>
<td>35.8</td>
<td>28.2</td>
<td>30.1</td>
<td>31.4</td>
<td>76.1</td>
<td>83.9</td>
<td>75.2</td>
<td>78.4</td>
</tr>
<tr>
<td>HI 9</td>
<td>54</td>
<td>45</td>
<td>35</td>
<td>30</td>
<td>25</td>
<td>15</td>
<td>39.1</td>
<td>40.2</td>
<td>30.8</td>
<td>36.7</td>
<td>61.1</td>
<td>65.1</td>
<td>63.5</td>
<td>63.2</td>
</tr>
<tr>
<td>NMNH9</td>
<td>59</td>
<td>45</td>
<td>5</td>
<td>5</td>
<td>10</td>
<td>5</td>
<td>23.1</td>
<td>26</td>
<td>23.4</td>
<td>24.2</td>
<td>65.1</td>
<td>64.7</td>
<td>62.4</td>
<td>64.1</td>
</tr>
<tr>
<td>HI 10</td>
<td>64</td>
<td>55</td>
<td>45</td>
<td>60</td>
<td>50</td>
<td>55</td>
<td>56</td>
<td>67.1</td>
<td>59.4</td>
<td>60.8</td>
<td>69.9</td>
<td>73.8</td>
<td>68.8</td>
<td>70.8</td>
</tr>
<tr>
<td>NMNH10</td>
<td>61</td>
<td>50</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>26.2</td>
<td>39.3</td>
<td>26.3</td>
<td>30.6</td>
<td>77.8</td>
<td>76.7</td>
<td>69.2</td>
<td>74.6</td>
</tr>
<tr>
<td>HI 11</td>
<td>66</td>
<td>40</td>
<td>0</td>
<td>10</td>
<td>15</td>
<td>45</td>
<td>25.3</td>
<td>30.6</td>
<td>46.3</td>
<td>34.1</td>
<td>60.2</td>
<td>62.1</td>
<td>54.8</td>
<td>59.0</td>
</tr>
<tr>
<td>NMNH11</td>
<td>69</td>
<td>40</td>
<td>10</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td>30.7</td>
<td>31.9</td>
<td>31.9</td>
<td>31.5</td>
<td>59.2</td>
<td>59.3</td>
<td>55.2</td>
<td>57.9</td>
</tr>
<tr>
<td>HI 12</td>
<td>58</td>
<td>65</td>
<td>45</td>
<td>45</td>
<td>55</td>
<td>70</td>
<td>50.6</td>
<td>58.3</td>
<td>75.9</td>
<td>61.6</td>
<td>80.2</td>
<td>82.6</td>
<td>85.2</td>
<td>82.7</td>
</tr>
<tr>
<td>NMNH12</td>
<td>62</td>
<td>65</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>15</td>
<td>30.6</td>
<td>27.7</td>
<td>28.6</td>
<td>29.0</td>
<td>86.9</td>
<td>83</td>
<td>84.6</td>
<td>84.8</td>
</tr>
</tbody>
</table>

**TABLE 4.1:** Demographic and threshold information for each HI/NMH1 listening pair. See text for details.
B. Threshold elevation

To control for differences in audibility between the NH and HI listener groups, all listeners had their thresholds elevated using TEN (threshold equalizing noise, Moore et al., 2000). It was decided to have both groups listen in noise to remove any possible confounds due to the distraction that may have been caused by having one group listen in background noise and the other in quiet. The assumption is that the NMNH listeners will have normally compressive responses and that the background noise will only serve to decrease their audibility (e.g., Gregan et al., 2010). 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). 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; to avoid using an uncomfortably loud TEN level, 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 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 4.1. 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.

C. Stimuli and procedure

Intelligibility of IEEE sentences (IEEE, 1969) was measured in a background noise that was spectrally shaped to match the long-term power spectrum of the sentences and was either unmodulated or modulated with a 10-Hz square wave. The steady and gated noises were scaled to have equal rms values overall. The choice of a 10-Hz modulation rate was based on results from previous studies that show 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 from 500-4000 Hz. The speech level was fixed across conditions and the masker level was varied to evaluate the percentage of correct keywords in the sentences for a variety of SNRs. 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. 4.1 and 4.2. The level of the speech was determined as part of the pilot testing for each HI listener and was set to be yield 80% or better performance in quiet while not exceeding tolerable loudness levels. This was possible for all HI listeners except HI12, for whom a quiet score 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 background 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 set-up. 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 to 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. However, the audibility should have been the same for both HI and NMNH listeners in each pair, due to the use of the TEN. This suggests that while 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 across an age- and audibility-matched HI and NMNH pair. Therefore, differences in MR seen within a given HI and NMNH listener pair would not 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 earphones. 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. 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. 4.1 and 4.2, 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. 4.1 is the HI counterpart for NMNH1 in the upper left hand corner of Fig. 4.2). The raw data points are indicated by circles (filled for results in the square-wave noise and open for the results in the steady-state noise). The data are plotted as percent correct versus SNR. Note that not all SNRs were tested for each listener (due to time and individual performance constraints). This was particularly true for HI12 who 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 percent correct in quiet (no steady-state or gated noise, only the always-present TEN). It can be seen that when MR was seen for a particular SNR, the SNR was typically less than 0 dB. In other words, the more difficult listening 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).
Fig. 4.1 HI listener results for IEEE sentence percent correct as a function of signal-to-noise ratio (SNR) in a steady-state noise background (open circles) or a square-wave gated noise background (filled circles). Lines represent best fits using a 3 parameter sigmoidal function. Note that the NMNH counterpart for a given HI listener will be found in the same panel location in Fig. 4.2.
Fig. 4.2 As Fig. 4.1, but for NMNH listeners.
The data were fitted with a 3-parameter sigmoidal function, to allow for prediction of SNR required for 50% correct performance in each noise background. The fits are show as solid lines in the figures. The equation used was:

\[ y = \frac{a}{1 + \exp\left(-\frac{x-x_0}{b}\right)} \]  

(Eq. 4.1)

where \( y \) is the percent of words correctly reported, \( x \) is the SNR in dB, \( a \) is maximum (asymptotic) \( y \) value, \( b \) controls the slope of the transition region, and \( x_0 \) is the SNR for 50% correct.

<table>
<thead>
<tr>
<th></th>
<th>( R^2 ) SQ fit</th>
<th>( R^2 ) SSN fit</th>
<th>MR: -5 dB SNR</th>
<th>MR: SRT diff (SQ-SSN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI 1</td>
<td>0.94</td>
<td>0.99</td>
<td>10</td>
<td>-0.28</td>
</tr>
<tr>
<td>HI 2</td>
<td>0.98</td>
<td>0.92</td>
<td>16</td>
<td>-1.46</td>
</tr>
<tr>
<td>HI 3</td>
<td>0.96</td>
<td>0.99</td>
<td>12</td>
<td>-2.19</td>
</tr>
<tr>
<td>HI 4</td>
<td>0.92</td>
<td>0.89</td>
<td>30</td>
<td>-4.14</td>
</tr>
<tr>
<td>HI 5</td>
<td>0.94</td>
<td>0.99</td>
<td>54</td>
<td>-5.16</td>
</tr>
<tr>
<td>HI 6</td>
<td>0.99</td>
<td>0.96</td>
<td>-9</td>
<td>2.70</td>
</tr>
<tr>
<td>HI 7</td>
<td>0.99</td>
<td>0.99</td>
<td>15</td>
<td>0.68</td>
</tr>
<tr>
<td>HI 8</td>
<td>0.88</td>
<td>0.97</td>
<td>7</td>
<td>6.42</td>
</tr>
<tr>
<td>HI 9</td>
<td>0.77</td>
<td>0.99</td>
<td>18</td>
<td>-5.44</td>
</tr>
<tr>
<td>HI 10</td>
<td>0.99</td>
<td>0.98</td>
<td>2</td>
<td>2.54</td>
</tr>
<tr>
<td>HI 11</td>
<td>0.92</td>
<td>0.99</td>
<td>68</td>
<td>-22.79</td>
</tr>
<tr>
<td>HI 12</td>
<td>0.86</td>
<td>0.97</td>
<td>9*</td>
<td>4.40</td>
</tr>
<tr>
<td>NMNH 1</td>
<td>0.99</td>
<td>0.99</td>
<td>42</td>
<td>-5.29</td>
</tr>
<tr>
<td>NMNH 2</td>
<td>0.96</td>
<td>0.99</td>
<td>51</td>
<td>-11.40</td>
</tr>
<tr>
<td>NMNH 3</td>
<td>0.99</td>
<td>0.99</td>
<td>49</td>
<td>-10.79</td>
</tr>
<tr>
<td>NMNH 4</td>
<td>0.93</td>
<td>0.98</td>
<td>68</td>
<td>-8.88</td>
</tr>
<tr>
<td>NMNH 5</td>
<td>0.99</td>
<td>1.00</td>
<td>24</td>
<td>-2.46</td>
</tr>
<tr>
<td>NMNH 6</td>
<td>0.99</td>
<td>0.99</td>
<td>35</td>
<td>-6.34</td>
</tr>
<tr>
<td>NMNH 7</td>
<td>0.99</td>
<td>0.99</td>
<td>48</td>
<td>-5.83</td>
</tr>
<tr>
<td>NMNH 8</td>
<td>0.99</td>
<td>0.99</td>
<td>41</td>
<td>-3.24</td>
</tr>
<tr>
<td>NMNH 9</td>
<td>0.99</td>
<td>0.94</td>
<td>50</td>
<td>-5.96</td>
</tr>
<tr>
<td>NMNH 10</td>
<td>0.77</td>
<td>0.99</td>
<td>69</td>
<td>-16.81</td>
</tr>
<tr>
<td>NMNH 11</td>
<td>0.93</td>
<td>0.96</td>
<td>30</td>
<td>-2.62</td>
</tr>
<tr>
<td>NMNH 12</td>
<td>0.99</td>
<td>1.00</td>
<td>9</td>
<td>-1.99</td>
</tr>
</tbody>
</table>
The $R^2$ values indicating the goodness of fit are shown in Table 4.2. 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 used. The first summary measure was based on the percentage-point improvement in 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 percent-correct scores at -5 dB SNR for square-wave gated and steady-state noise conditions is shown in third column of Table II. By comparing this value of MR across each HI and NMNH pair, it can be seen that all NMNH listeners showed more MR than their HI counterparts, with the exception of HI11 and NMNH11. A paired-samples t-test confirmed that the MR at -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 was derived using the $x_0$ values from the sigmoidal fits to the data. These values are shown in the last column of Table 4.2 (labeled “SRT diff”). 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 poorest SNRs tested. For examples of this problem, see the data from NMNH2, NMNH10 and HI11. Also note the results from HI12, who never reached 50% even in
quiet testing. Negative values indicate MR, in that the listener was able to achieve 50% correct in a less favorable SNR in square-wave gated noise than in steady-state noise. Positive values indicate the listener performed better in the steady-state noise condition ("negative" MR). Note that none of the NMNH listeners had positive $x_0$ values, while 5 of the HI listeners did. This indicates that for these 5 listeners, the modulated masker actually impaired performance compared to the steady-state masker instead of providing MR. This negative MR has been observed in some previous studies (Kwon and Turner, 2001; Nelson et al., 2003). For each HI/NMNH listener pair, the NMNH listener always showed more MR according to this measure (referred to as the "SRT difference") with the exception of HI11/NMNH11. In this case, while both listeners showed MR, HI11 showed more MR using this summary value.

In contrast to the results from the MR at -5 dB SNR, a paired t-test comparison of the SRT difference values for HI versus NMNH did not indicate a significant difference between groups [$t(11) = 1.723; p=0.113$]. The lack of a significant effect may reflect the fact that 50% correct was achieved on average at an SNR of 0 dB, whereas the largest MR effects are typically observed at negative SNRs. In addition, the accuracy of the $x_0$ measure was compromised by the fact that extrapolation was necessary in several cases. Nevertheless, the two summary values (MR at -5 dB SNR and the SRT difference) were highly correlated with each other, as indicated by linear regression [$R^2 = 0.728; p<0.001$], and both measures will be used for comparison throughout the remainder of this paper.

As mentioned in the introduction, recent work by Bernstein and Grant (2009) and Bernstein and Brungart (2011) has suggested that the reason HI listeners seemingly show
less MR is due to the fact that they require higher SNRs to achieve a given percent
correct in steady-state noise and, therefore, are not typically tested at negative SNRs,
where the most MR is seen. This hypothesis was evaluated with our data by plotting the
MR at -5 dB SNR and SRT difference as a function of the SNR required for 50% correct
in steady-state noise. According to the hypothesis, the measures should be correlated.
The data are shown in Figs. 4.3A (MR at -5 dB SNR) and 4.3B (SRT difference). A
linear regression analysis failed to find a significant relationship between either of these
comparisons 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 is not due simply to differences in the SNR in
steady-state noise between the two groups.

Fig. 4.3A Percent MR at -5 dB SNR as a function of SNR required for 50% correct in
steady-state noise.
The use of TEN rules out audibility differences as an explanation of differences in MR between the HI listeners and their matched NMNH listeners. However, the TEN level and overall speech level differed between each pair 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 both measures of MR in the HI group, NMNH group, and in the combined HI/NMNH group. Results for the HI comparison are shown in Figs. 4.4A and 4.4B (for MR at -5 dB and SRT difference, respectively). Results for the NMNH comparison are shown in Figs. 4.4C and 4.4D. Results for the combined HI/NMNH data are show in Figs. 4.5A and 4.5B. Both HI-only comparisons (MR at -5 dB and SRT difference) and the one of the combined analyses (SRT difference) indicated that MR decreased significantly as the speech presentation
level increased. The combined group comparison for MR at -5 SNR did not reach significance, but showed a trend in the same direction as the HI only data. Neither of the NMNH only comparisons reached significance, indicating that the trend shown in the combined figures is due to the HI results. This significant effect of level is in the opposite direction to what was predicted (but not found) in the study by Summers and Molis (2004). These authors predicted that HI listeners would show more MR for speech as the presentation level was increased and, therefore, as the speech stimuli became more audible. They did note some rollover effect (worsening of performance at higher intensity levels) for their NH listeners and two of their six HI listeners. It is possible that rollover due to high intensity levels could account for some of the reduced MR seen at high speech levels for the current subjects as well. However, this would not explain the difference in MR seen between HI and NMNH listeners due to the steps taken to equate audibility and level. In addition, it is not likely that rollover would have occurred for the speech levels used in this study, which approximated 75 dB HL and below.
**Fig. 4.4A** Percent improvement in square-wave gated noise versus steady-state noise at a -5 dB SNR as a function of the RMS speech levels for HI listeners.

**Fig. 4.4B** As for Fig. 4.4A, but the MR summary measure used is SRT difference.
Fig. 4.4C As for Fig. 4.4A, but for the NMNH listeners only

Fig. 4.4D As for Fig. 4.4C, but with SRT difference as the summary measure
**Fig. 4.5A** As for Fig 4.4A, but for combined HI and NMNH results.

**Fig. 4.5B** As for Fig. 4.4B, but for combined HI and NMNH results.
The MR summary measures for HI listeners were also plotted as a function of average quiet threshold for the three test signals, which were selected to span the range of the speech bandwidth (see Figs. 4.6A and 4.6B). These plots show a significant correlation between average absolute threshold and degree of MR [MR at -5 dB SNR: $R^2 = 0.5004, p = 0.01$; SRT difference: $R^2 = 0.4837, p = 0.012$]. In other words, listeners with less hearing loss tended to show more MR. However, in both cases absolute thresholds account for only around half the variance of the MR measures, suggesting other factors may play a role.

![Graph showing the relationship between percent MR at -5 dB SNR and average quiet threshold for HI listeners. The graph includes a line of best fit with equation $R^2 = 0.5004$ and $p = 0.01$.]

**Fig. 4.6A** Percent MR at -5 dB SNR as a function of average quiet threshold for HI listeners.
Our finding of more MR for listeners with normal hearing than those with HI, even when equated for audibility, 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 factors, other than audibility, 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. 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 tonal stimuli located in a fluctuating masker.
valley compared to a masker 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.

A. Listeners

The same 24 listeners participated in this experiment as were used in experiment 1.

B. Stimuli

Signal tones were fixed in level at 8 dB SL (1500 and 4000 Hz) and 10 dB SL (500 Hz). The 500 Hz tone was presented at the slightly higher level due to the inability of several listeners to consistently detect the lower-level 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” (e.g., Shailer and Moore, 1987) and possible physical overlap between the BM response to the masker and signal at this low frequency. The masker was either a steady-state noise with the same long-term power
spectrum as the IEEE sentences, or the same noise modulated with 10-Hz square-wave, as described for experiment 1, with the exception that the IEEE noise was not bandpass filtered in the present experiment. The masker duration was 500 ms (i.e., 5 periods of the gated masker). The same 10-Hz square-wave gated masker was used as in experiment 1, with 50-ms on-periods and 50-ms off-periods. For each signal frequency, 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 valley (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. 4.7. For the steady-state 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 a TEN, selected for each subject pair individually, to equate audibility in the absence of the masker. These levels are listed in Table 4.1.

C. Procedure

The pure tone masking release thresholds 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 (either broadband or gated) was presented in all three intervals, whereas 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 adaptively varying the masker level was 8 dB, which was reduced to 4 dB after the first reversal point, and was held constant at 2 dB 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 from three runs. The presentation of temporal position and masker type was randomized within and across subjects and repetitions. Listening sessions were 2 h in length, including frequent breaks. This particular experiment took approximately 2-4 h 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 that exceeded this masker level 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. The MRPT was calculated as the difference between the masker level necessary to mask the tonal signal at a particular temporal location in the gated masker and the masker level necessary to mask the tonal signal in the presence of steady state noise. A larger difference indicates a higher masker level was required in the gated masker than the steady masker, much like MR for speech.
Fig. 4.7 Schematic showing temporal location of signals in relation to the gated masker for the pure tone masking release experiment.

D. 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 masker to provide the summary measure of MRPT for each subject. These values are listed in Table 4.3 for each of the test signals, along with an average MRPT value across test frequencies.

<table>
<thead>
<tr>
<th></th>
<th>MRPT 500</th>
<th>MRpt 1.5</th>
<th>MRPT 4k</th>
<th>AVG MRPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI1</td>
<td>30.6</td>
<td>22.2</td>
<td>22.1</td>
<td>25.0</td>
</tr>
<tr>
<td>HI2</td>
<td>16.8</td>
<td>43.5</td>
<td>8.6</td>
<td>23.0</td>
</tr>
<tr>
<td>HI3</td>
<td>0.08</td>
<td>34.6</td>
<td>6.7</td>
<td>13.8</td>
</tr>
<tr>
<td>HI4</td>
<td>4.8</td>
<td>8.5</td>
<td>15.6</td>
<td>9.6</td>
</tr>
<tr>
<td>HI5</td>
<td>5.3</td>
<td>26.8</td>
<td>9.9</td>
<td>14.0</td>
</tr>
<tr>
<td>HI6</td>
<td>0.4</td>
<td>4.9</td>
<td>8.5</td>
<td>4.6</td>
</tr>
<tr>
<td>HI7</td>
<td>11.8</td>
<td>16.1</td>
<td>14.1</td>
<td>14.0</td>
</tr>
<tr>
<td>HI8</td>
<td>20.4</td>
<td>26.7</td>
<td>11.1</td>
<td>19.4</td>
</tr>
<tr>
<td>HI9</td>
<td>30.1</td>
<td>21</td>
<td>17.2</td>
<td>22.8</td>
</tr>
<tr>
<td>HI10</td>
<td>3.8</td>
<td>-0.6</td>
<td>3.2</td>
<td>2.1</td>
</tr>
<tr>
<td>HI11</td>
<td>49.6</td>
<td>33.2</td>
<td>10.1</td>
<td>31.0</td>
</tr>
<tr>
<td>HI12</td>
<td>0.6</td>
<td>9.1</td>
<td>6.2</td>
<td>5.3</td>
</tr>
<tr>
<td>NMNH1</td>
<td>12.4</td>
<td>8.5</td>
<td>5.3</td>
<td>8.7</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>NMNH2</td>
<td>34.9</td>
<td>30.4</td>
<td>26.3</td>
<td>30.5</td>
</tr>
<tr>
<td>NMNH3</td>
<td>35.6</td>
<td>40.9</td>
<td>26.7</td>
<td>34.4</td>
</tr>
<tr>
<td>NMNH4</td>
<td>15.2</td>
<td>29.8</td>
<td>17.6</td>
<td>20.9</td>
</tr>
<tr>
<td>NMNH5</td>
<td>8.6</td>
<td>18.8</td>
<td>23</td>
<td>16.8</td>
</tr>
<tr>
<td>NMNH6</td>
<td>24.2</td>
<td>11.6</td>
<td>11.1</td>
<td>15.6</td>
</tr>
<tr>
<td>NMNH7</td>
<td>16.8</td>
<td>14.29</td>
<td>14.92</td>
<td>15.3</td>
</tr>
<tr>
<td>NMNH8</td>
<td>10.7</td>
<td>7</td>
<td>13.5</td>
<td>10.4</td>
</tr>
<tr>
<td>NMNH9</td>
<td>3</td>
<td>27.4</td>
<td>16.9</td>
<td>15.8</td>
</tr>
<tr>
<td>NMNH10</td>
<td>25.2</td>
<td>20.2</td>
<td>19.5</td>
<td>21.6</td>
</tr>
<tr>
<td>NMNH11</td>
<td>28.3</td>
<td>36.3</td>
<td>31.5</td>
<td>32.0</td>
</tr>
<tr>
<td>NMNH12</td>
<td>11.9</td>
<td>11.4</td>
<td>7.9</td>
<td>10.4</td>
</tr>
</tbody>
</table>

**Table 4.3:** Masking release for pure tone signals. See text for details regarding calculation.

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 4.3, there is a great deal of variability in the amount of MRPT within each group of listeners.

![Graph showing the relationship between MRPT and percent MR at -5 dB SNR.]

**Fig. 4.8A** Average MRPT across all 3 signal frequencies as a function of the MR for speech at -5 dB SNR.
The possible relationship between MRPT and MR for speech was also explored. Figure 4.8A shows the MRPT for each subject, averaged across the three signal frequencies, plotted against the MR for speech at -5 dB SNR. Regression analysis confirmed that this relationship was significant [$R^2 = 0.2191$, $p = 0.021$], meaning that individuals with larger MRPT also tended to show more MR for speech at -5 dB SNR. Similarly, Fig. 4.8B plots the average MRPT versus MR in terms of SRT difference and this result is also significant [$R^2 = 0.3276$, $p = 0.004$]. This significant negative correlation shows that as the SRT difference between the square-wave and steady-state conditions becomes smaller (i.e. less MR for speech), the MRPT also becomes smaller.

**Fig. 4.8B** Average MRPT across the 3 signal frequencies as a function of SRT difference for 50% correct for speech.

![Graph showing the relationship between MRPT and SRT difference](image)
IV. EXPERIMENT 3: ESTIMATING COCHLEAR COMPRESSION

The role of cochlear compression in the MR results seen for speech and pure tones in the previous 2 experiments was explored in the present experiment by estimating cochlear compression for frequencies across the speech bandwidth (500, 1500, and 4000 Hz) using the TMC technique (Nelson et al., 2001). 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. Growth of masking (GOM) 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 issue 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). This approach is also used in the present study to estimate compression at CFs between 500 and 4000 Hz.

A. Listeners

The same 24 listeners from experiments 1 and 2 took part in this experiment.

B. Stimuli

The stimuli were generated digitally as previously described at a sampling rate of 48 kHz. For all listeners, the signal levels were fixed at a low sensation level (8 dB SL for 1500 and 4000 Hz and 10 dB SL for 500 Hz). TEN was present to elevate thresholds, as described in Section II.B. As in experiment 2, the slightly higher SL for 500 Hz was used after pilot data indicated that many listeners could not consistently detect the 500 Hz signal at the lower SL. 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 (0.45x4000). 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. Similarly, 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.

C. 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 pure-tone masker was presented in all three intervals, whereas 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 adaptively varying the 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 from three runs. The presentation order of delay was randomized within and across subjects and repetitions. Listening sessions were 2 h in length, including frequent breaks. The TMC experiment took approximately 8-10 h per subject to complete.
D. Results

The raw TMC data for HI and NMNH listeners are shown in Figs. 4.9A and 4.9B, respectively. Error bars indicate ±1 standard deviation across runs. As described in Nelson et al., (2001) these raw 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 both 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 fitted function of off-frequency masker level as a function of masker-signal delay was then used as the “linear” reference to derive estimates of BM input-output function for each 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 to do so in this experiment 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 more shallow 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).
Fig. 4.9A TMC data for the 3 on-frequency conditions and the one off-frequency condition for HI listeners. Error bars indicate standard deviations.
Fig. 4.9B As for Fig. 4.9A, but for the NMNH listeners.
Fig. 4.10A Derived input-output functions from the TMC data in Fig. 4.9A. 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).
**Fig. 4.10B** As for Fig. 4.10A, but for the NMNH listeners.
### TABLE 4.4: Fits to TMC derived input-output data for the 3 “on-frequency” conditions tested and the “off-frequency” condition. Numbers in parentheses are $R^2$ values.

<table>
<thead>
<tr>
<th></th>
<th>500 Hz (On)</th>
<th>1500 Hz (On)</th>
<th>4000 Hz (On)</th>
<th>4000 Hz (Off)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HI 1</td>
<td>0.675 (0.96)</td>
<td>0.690 (0.94)</td>
<td>0.854 (0.99)</td>
<td>0.33 (0.97)</td>
</tr>
<tr>
<td>HI 2</td>
<td>1.91 (0.78)</td>
<td>0.308 (0.89)</td>
<td>0.673 (0.85)</td>
<td>0.26 (0.51)</td>
</tr>
<tr>
<td>HI 3</td>
<td>0.671 (0.83)</td>
<td>0.972 (0.96)</td>
<td>1.57 (0.67)</td>
<td>0.39 (0.92)</td>
</tr>
<tr>
<td>HI 4</td>
<td>0.234 (0.86)</td>
<td>0.281 (0.98)</td>
<td>0.358 (0.95)</td>
<td>0.21 (0.96)</td>
</tr>
<tr>
<td>HI 5</td>
<td>0.220 (0.89)</td>
<td>0.222 (0.98)</td>
<td>0.358 (0.98)</td>
<td>0.24 (0.93)</td>
</tr>
<tr>
<td>HI 6</td>
<td>0.580 (0.96)</td>
<td>0.442 (0.99)</td>
<td>0.739 (0.89)</td>
<td>0.09 (0.73)</td>
</tr>
<tr>
<td>HI 7</td>
<td>1.146 (0.89)</td>
<td>0.932 (0.94)</td>
<td>0.257 (0.85)</td>
<td>0.45 (0.89)</td>
</tr>
<tr>
<td>HI 8</td>
<td>0.164 (0.99)</td>
<td>0.320 (0.95)</td>
<td>0.673 (0.96)</td>
<td>0.23 (0.51)</td>
</tr>
<tr>
<td>HI 9</td>
<td>0.418 (0.98)</td>
<td>0.426 (0.94)</td>
<td>0.499 (0.86)</td>
<td>0.29 (0.81)</td>
</tr>
<tr>
<td>HI 10</td>
<td>0.717 (0.62)</td>
<td>0.504 (0.97)</td>
<td>0.692 (0.89)</td>
<td>0.18 (0.76)</td>
</tr>
<tr>
<td>HI 11</td>
<td>0.173 (0.76)</td>
<td>0.272 (0.85)</td>
<td>0.569 (0.87)</td>
<td>0.28 (0.92)</td>
</tr>
<tr>
<td>HI 12</td>
<td>0.406 (0.97)</td>
<td>0.332 (0.97)</td>
<td>0.737 (0.90)</td>
<td>0.14 (0.81)</td>
</tr>
<tr>
<td>NMNH 1</td>
<td>0.438 (0.78)</td>
<td>0.587 (0.98)</td>
<td>0.577 (0.94)</td>
<td>0.49 (0.73)</td>
</tr>
<tr>
<td>NMNH 2</td>
<td>0.402 (0.94)</td>
<td>0.703 (0.94)</td>
<td>0.615 (0.97)</td>
<td>0.47 (0.96)</td>
</tr>
<tr>
<td>NMNH 3</td>
<td>0.613 (0.56)</td>
<td>0.276 (0.76)</td>
<td>0.557 (0.94)</td>
<td>0.35 (0.94)</td>
</tr>
<tr>
<td>NMNH 4</td>
<td>0.122 (0.92)</td>
<td>0.085 (0.94)</td>
<td>0.059 (0.93)</td>
<td>0.086 (0.10)</td>
</tr>
<tr>
<td>NMNH 5</td>
<td>0.949 (0.85)</td>
<td>0.717 (0.94)</td>
<td>0.555 (0.93)</td>
<td>0.47 (0.96)</td>
</tr>
<tr>
<td>NMNH 6</td>
<td>0.356 (0.91)</td>
<td>0.327 (0.94)</td>
<td>0.206 (0.96)</td>
<td>0.26 (0.86)</td>
</tr>
<tr>
<td>NMNH 7</td>
<td>0.555 (0.91)</td>
<td>0.382 (0.75)</td>
<td>0.299 (0.86)</td>
<td>0.42 (0.89)</td>
</tr>
<tr>
<td>NMNH 8</td>
<td>0.487 (0.95)</td>
<td>0.533 (0.99)</td>
<td>0.421 (0.98)</td>
<td>0.42 (0.90)</td>
</tr>
<tr>
<td>NMNH 9</td>
<td>0.929 (0.95)</td>
<td>0.655 (0.94)</td>
<td>1.23 (0.96)</td>
<td>0.92 (0.95)</td>
</tr>
<tr>
<td>NMNH 10</td>
<td>0.628 (0.96)</td>
<td>0.710 (0.95)</td>
<td>1.38 (0.97)</td>
<td>0.68 (0.93)</td>
</tr>
<tr>
<td>NMNH 11</td>
<td>0.286 (0.94)</td>
<td>0.212 (0.95)</td>
<td>0.429 (0.81)</td>
<td>0.40 (0.98)</td>
</tr>
<tr>
<td>NMNH 12</td>
<td>0.428 (0.83)</td>
<td>0.356 (0.96)</td>
<td>0.344 (0.90)</td>
<td>0.20 (0.70)</td>
</tr>
</tbody>
</table>

The raw TMC data from Figs. 4.9A and 4.9B were used to obtain the derived input-output functions for the HI and NMNH listeners, as shown in Fig. 4.10A and 4.10B, respectively. To estimate compression exponents, the function for each signal frequency was fitted with a straight line and the slope was taken as the compression estimate. These compression estimates (along with the $R^2$ values for the linear regressions) are shown in Table 4.4. The dashed line depicts 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 a more complex fitting routine (i.e. third order polynomial) as has 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, 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 $R^2$ values, this simple fitting process generally describes the underlying functions well ($R^2$ values were generally higher than 0.75), with a few exceptions (i.e. NMNH3 for 500 Hz; HI10 for 500 Hz; and HI3 at 4000 Hz).

E. Discussion

At first glance, it is interesting to note that HI listeners with higher degrees of hearing loss (like 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 signals. Previous studies have noted that estimates of compression do not appear to be strongly correlated with the underlying audiometric thresholds (Plack et al., 2004; Lopez-Poveda et al., 2005), and the present results appear to 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 the signal stimulus (in this case, 1000 Hz). The resulting compression estimates for these individuals with audiometrically normal hearing at the test frequency ranged from 0.083 up to 1.749. The fact that Poling et al. (2011) showed similar variability in compression estimates using the TMC method in quiet suggests that the variability in compression estimates 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 off-frequency TMC curve used as a linear reference is assumed to reflect only the recovery from forward masking and, therefore, it can be used as an estimate of temporal resolution. As stated in the introduction, there have been conflicting results in the literature to date as per 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 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 4.4). As can be seen from Table 4.4, there was quite a bit of variability in the slopes, even within the NMNH group. However, paired t-test showed a significant difference between the slopes for the NMNH and the HI
group \( t(11) = -2.78; p=0.018 \). Therefore, the slopes of the off-frequency TMC curves were significantly shallower for the HI versus the NMNH group. This is consistent with the finding that the NMNH listeners had more MR for speech than did the HI group.

**Fig. 4.11** Compression exponents derived from TMC functions plotted as a function of quiet threshold (for the same frequency). HI data.

**Fig. 4.12** As for Fig. 4.11, but for NMNH listeners.
V. COMPARISONS OF COMPRESSION AND TEMPORAL RESOLUTION WITH MASKING RELEASE IN SPEECH AND TONES

A. Compression estimates and absolute threshold

Our estimates of cochlear compression were compared with absolute thresholds for the test signals used for the HI listeners (Fig. 4.11). 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. As can be seen from the scatterplot, there is a slight trend towards less compression with increasing absolute threshold, but there is 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., Yates et al., 1990; Ruggero, 1992; Rosengard et al., 2005). 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. Fig. 4.12 shows the results from the NMNH listeners. As expected, no significant relationships were found between compression exponents and absolute thresholds in the NMNH group.

B. Compression estimates and MR for speech

To determine if our hypothesis that underlying compression estimates are correlated with MR for speech, the two summary measures of MR (percent correct
improvement seen at -5 SNR and the difference between SNR required for 50% correct in gated versus steady noise) were plotted as a function of several compression summary values. Although compression was estimated for three signal frequencies that spanned the bandwidth of the speech stimulus, it was necessary to use a summary compression estimate measure to explore the relationship between MR and compression. Three summary measures of compression were selected: the average compression exponent across the three signal frequencies, the least compressive exponent of the three, and the most compressive exponent of the three. The data from the HI listeners were plotted separately from NMNH data to improve readability of the figures. As can be seen from Figs. 4.13A and 4.13B (HI) and Figs. 4.14A and 4.14B (NMNH), and confirmed by regression analysis, there were no significant correlations between any of the compression estimates and either of the MR summary measures (p > 0.05 in all cases, even with no post-hoc corrections for multiple comparisons). This is inconsistent with results from Oxenham and Dau (2004) who found a weak but significant correlation between the degree of masking difference (between flat and modulated Schroeder-phase maskers) and auditory filter bandwidth. In addition, auditory filter bandwidth has been shown to be correlated with estimates of BM compression (using forward growth of masking results) (Moore et al., 1999b).
Fig. 4.13A  MR at -5 SNR as a function of various compression exponents (see legend). HI results.

Fig. 4.13B  SRT difference for speech MR as a function of various compression exponent summary measures (see legend). HI results.
Fig. 4.14A As for Fig. 4.13A, but for NMNH results.

Fig. 4.14B As for Fig. 4.13B, but for NMNH results.
C. Compression estimates and MRPT

To determine if the degree of MRPT varied systematically with the estimate of compression the comparable frequency, we plotted MRPT as a function of the estimated compression for each of the signal frequencies, as well as an average value, as shown in Figs. 4.15A (HI) and 4.15B (NMNH). As can be seen from the considerable scatter in the figures, and as was confirmed with regression estimates, there was no significant relationship for any of these comparisons. 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.

Fig. 4.15A MRPT across frequency as a function of compression estimate across frequency, for HI listeners.
**Fig. 4.15B** As for Fig. 4.15A, but for NMNH listeners.

**D. Relationship of temporal resolution to absolute thresholds and MR**

The slopes of the off-frequency TMC curves are plotted as a function of the amount of MR seen at -5 dB SNR and the SRT difference values, (see Fig. 4.16A and 4.16B, respectively). While not statistically significant for either summary measure (p=0.049 for 4.16A and p=0.1194 for 4.16B), there is a trend in the expected direction for both data sets. In other words, the magnitude of the MR (for either summary measure) increases as the estimated rate of recovery from forward masking increases.
**Fig. 4.16A** Slope of the off-frequency TMC function as a function of the MR seen at -5 dB SNR.

**Fig. 4.16B** Slope of the off-frequency TMC function (estimate of recovery from forward masking) as a function of MR (SRT difference for 50% correct).
Because some previous studies have observed a decrease in recovery from forward masking with increasing audiometric threshold (e.g., Derleth et al., 2001; Rhebergen et al., 2006), the off-frequency slope was plotted as a function of the thresholds in quiet for the brief tonal stimuli. As seen in Fig. 4.17, there is a trend for the slope of the off-frequency function to decrease as hearing loss increases and this relationship is significant [$R^2 = 0.2141; p=0.0228$]. However, there is also a great deal of scatter amongst the NH data (with slope values ranging from 0.086 to 0.92) that makes further interpretation of these results difficult. However, even with the variability in the NMNH data, the NMNH listeners did have significantly steeper recover functions than did the HI listeners, as per a paired t-test comparison described earlier.

![Graph](https://via.placeholder.com/150)

**Fig. 4.17** Slope of the off-frequency TMC function as a function of the average threshold for the tonal stimuli in quiet.
VI. SUMMARY

The results presented have confirmed those of previous studies that have shown less MR for speech for individuals with HI compared to those with NH but elevated thresholds (e.g., Eisenberg et al., 1995; Bacon et al., 1998). 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 that equated audibility for each pair of listeners. The remainder of the paper was concerned with determining if other factors, aside from audibility, could be used to predict MR in speech. The factors explored included MR in the same masker backgrounds for simpler, tonal stimuli (MRPT), peripheral compression (as estimated by the TMC method), and recovery from forward masking (as estimated from the slope of the TMC off-frequency linear reference curve). The findings can be summarized as follows:

1) All but two of the HI/NMNH pairs showed that NMNH listeners have more MR at -5 dB SNR and more MR as determined by the SRT difference for the NH counterpart, even though they were equated for audibility.

2) There was a significant relationship between thresholds in TEN and MR (i.e. higher masked thresholds were associated with less MR). This may have resulted in decreased MR for NMNH/HI pairs that listened in the higher TEN levels. However, it would not differentially affect the MR across a given listening pair, as both listeners had equivalent masked thresholds.

3) The present results do not support the view that degree of MR is determined by SNR, as described in Bernstein and Grant (2009).
4) MRPT was significantly correlated with both summary measures for speech MR. This is consistent with a common underlying mechanism for both speech and pure tone MR.

5) The compression estimates obtained using the TMC method were not significantly correlated with either of the MR for speech summary measures nor were they significantly correlated with the MRPT measures. They were also not strongly correlated with the quiet thresholds.

6) TMC results were highly variable even for the NMNH listeners despite their normal audiometric thresholds. It is unlikely that this is due to our use of TEN given the recent results from Poling et al. (2011) that showed a similar range of TMC-derived compression estimates for normal hearing listeners tested in quiet. Those researchers also ruled out potential subject behavior-related factors (such as test-retest variability) as contributing to the result variability.

7) Overall, NMNH listeners had significantly steeper recovery slopes than did the HI listeners. Recovery from forward masking (as estimated from the off-frequency slope of the TMC functions) showed a trend in the expected direction (slower recovery from forward masking being associated with decreased MR) for both of the MR for speech summary measures, but neither relationship reached significance. In addition, the off-frequency slope was significantly correlated with the average threshold in quiet for the test stimuli.

Overall, the study failed to find a strong relationship between MR in speech and cochlear compression. However, such a relationship cannot be completely ruled out,
given the wide variability in the TMC measures (even for the NMNH group). Further work in this area is needed to determine the best means of behaviorally estimating compression over a wide-frequency range before the question of the relationship between BM compression and MR can be fully addressed. In addition, the fact that two recent studies (the present study and that of Poling et al. (2011) have shown a large degree of variability in TMC results for listeners with audiometrically normal hearing suggests this method may need to be used with caution when attempting to determine estimates of BM compression (for both NH and HI populations).
CHAPTER 5: SUMMARY OF RESULTS

The primary goal of this thesis was to examine the relationship between masking release (MR) for speech and behavioral estimates of basilar membrane (BM) compression. The hypothesis was that hearing-impaired (HI) listeners who show reduced MR compared to normal hearing (NH) individuals, with thresholds elevated to equate audibility, have reduced amplification of low- and mid-level sounds due to loss or reduction of the cochlear active process, which in turn should result in reduced MR in a temporally-varying masker. Therefore, differences in underlying cochlear compression seen in HI listeners with sensorineural hearing loss should be significantly correlated with behavioral estimates of compression.

A. Forward growth of masking

Control for differences in audibility across NH and HI groups was accomplished by presenting a threshold equalizing noise (TEN, Moore et al., 2000) to both listener groups. To ensure that the addition of this background noise would not alter underlying BM response characteristics (and, therefore, estimates of compression), forward growth of masking (GOM) functions in the presence of TEN were measured for a group of NH listeners (Chapter 2). The results from Chapter 2 demonstrated no significant change in the slope of the off-frequency forward GOM function with increasing TEN level, nor was there a significant difference in the slope of the off-frequency function measured in quiet versus those obtain in TEN. These findings suggest that it is valid to assume that the addition of background noise to elevate thresholds in NH listeners does not alter behavioral estimates of compression, at least for the forward GOM method.
Unfortunately for our purposes, the forward GOM method is limited in its ability to accurately estimate compression at lower frequencies. The forward GOM method requires the use of an off-frequency masker which is assumed to be processed linearly at the place along the BM with a CF corresponding to the (higher) signal frequency.

Although this assumption is supported by results from physiological studies in the base of the cochlea (e.g. Robles et al., 1986; Ruggero et al., 1997), it seems unlikely to hold at frequencies corresponding to more apical locations in the cochlea (e.g. Cooper and Yates, 1994; Oxenham and Plack, 1997; Hicks and Bacon, 1999), so that the method may only be valid for signal frequencies of 4 kHz and above. Because hearing loss frequently varies with frequency, it is important to be able to obtain estimates of compression across a range of frequencies. Therefore, methods that rely on the assumption of linear processing of frequencies below CF (such as the GOM and temporal masking curve methods) cannot be used in comprehensive measures of compression across a wide range of frequencies.

B. Additivity of forward masking

An alternative method for estimating compression that does not require a linear off-frequency reference is the additivity of forward masking (AFM) method (e.g. Penner and Shiffrin, 1980; Widin and Viemeister, 1980; Plack and O'Hanlon, 2003; Plack et al., 2008). With the AFM method, two forward maskers are presented contiguously after having been set in level so that each is equally effective at independently masking a given signal level. When both of the maskers are presented, the assumption is that they are compressed independently prior to their effects being summed and this predicts an
additional 3 dB of masking in a linear system, according to the energy summation model (Green and Swets, 1966). In a non-linear, compressive system, the combination of these maskers will result in more than 3 dB of masking. This is because the internal effect of the signal will be compressed, requiring an increase in external signal level of more than 3 dB to achieve a 3 dB internal increase after compression. The magnitude of this “excess masking” is thought to be proportional to the degree of underlying compression, in that the more the signal is compressed internally, the greater the increase in external signal level that will be required to compensate.

The AFM technique had not been directly compared with the GOM technique, nor have the effects of AFM in threshold elevating background noise been evaluated previously. Therefore, Chapter 3 measured AFM for 4000 Hz signal in 40 dB TEN to allow for comparison of compression estimates with the results from Chapter 2 using the forward GOM technique in the same NH listeners. In addition, AFM was used to estimate compression for 500 and 1500 Hz signals, also in 40 dB TEN. Results of these measurements showed compressive responses for all listeners for the 500-Hz signal and for 5 out of 6 listeners for the 1500-Hz signal. However, only 2 listeners showed a compressive response for the 4000-Hz signal, in contrast with the forward GOM results.

Although the results from Chapter 2 strongly suggest that the use of TEN did not alter the compression characteristics in these NH listeners, it was necessary to test the possibility that the effects of the TEN may have resulted in decreased AFM-based compression estimates at 4000 Hz. Two additional conditions were measured: repetition of the original 4000-Hz condition without TEN and measurement of AFM using lower
level signals without background TEN. There was no statistical difference in the estimated compression slopes between these new measurements and the original 4000-Hz measurements in TEN, suggesting that it was neither high-stimulus levels nor the use of TEN that resulted in the finding of limited compression for the 4000 Hz signal.

It was deemed unlikely that the finding of reduced compression for the 4000 Hz signal was due to gain reduction effects of the medial olivocochlear efferent system (MOCR) based on findings from previous studies. Another possibility was that the acoustic reflex (AR) resulted in the differential findings of compression in the present study. The AR effect is most pronounced for frequencies under 2000 Hz (Humes, 1978); therefore, it is possible that the AR activation resulted in lower effective stimulus levels for the 500 and the 1500 Hz signals which kept the response on the more compressive portion of the BM response function. Even though there was not a statistical difference in compression estimates between the results for the 4000 Hz high-level and low-level stimuli in the present study, there was a non-significant trend for compression estimates to become more compressive as the stimulus level was lowered. Further study regarding this issue is required to better determine the underlying cause(s) for the pattern of results obtained in the present study.

As neither forward GOM nor AFM appeared to be a good choice to estimate compression across a range of frequencies, it was decided to use a modified temporal masking curve (TMC) approach to estimate compression for the noise-masked normal hearing (NMNH) and HI listeners who participated in the MR experiment. The TMC approach utilizes a fixed, low sensation level (SL) signal and determines the forward
masker level required to just mask this signal for a variety of masker-signal delays. The masker is either the same as the signal frequency (on-frequency) or it is much lower than the signal frequency (off-frequency). As a means of circumventing the uncertainty surrounding compression characteristics in the apical portion of the cochlea, previous studies have opted to use one off-frequency reference (obtained for a high-frequency signal) as the linear reference for lower signal frequencies. This approach was also adopted for the present study. On-frequency TMCs were measured for 500-, 1500-, and 4000-Hz signals. One off-frequency TMC was measured with a 4000-Hz signal and an 1800-Hz masker and this linear reference was used for all three signal frequencies.

C. Masking release

In the final set of experiments, described in Chapter 4, a group of 12 HI listeners were paired with age-matched NMNH counterparts. MR for IEEE sentences was evaluated by measuring percent correct (keyword identification) at a variety of signal-to-noise ratios (SNRs) in a background of speech-shaped steady-state noise (SSN) or square-wave noise gated at a rate of 10 Hz (SQ). For all but one NMNH/HI listener pair, the NMNH listeners showed more MR (defined as the difference in percent correct scores in the SQ and SSN conditions), confirming that HI generally show reduced MR, and that it cannot be attributed solely to differences in audibility of the speech in quiet.

The second experiment measured MR release for pure tones (MRPT). This experiment used the same signals and fixed signal levels as were used for the TMC measurements in experiment 3 and presented the signals either in SSN or SQ (the same background stimuli as were used to determine MR for speech). Masker level was varied
adaptively to just mask the signal and the signal was located (in the SQ noise) either at the masker peak or at one of three locations in the masker valley. The MRPT values were significantly correlated with the MR for speech, in that the listeners who showed more MR for speech tended to show more MRPT.

In an attempt to explain the speech and pure-tone MR findings from experiments 1 and 2, experiment 3 measured compression estimates using the TMC method. Interestingly, there was no significant correlation between the compression estimates and either the speech or pure-tone MR measures. This results is inconsistent with previous studies that have shown a relationship between MRPT and cochlear compression (e.g., Oxenham and Dau, 2004). In addition, the TMC-derived compression measures were not correlated with the listeners’ quiet thresholds. This result has been seen in previous studies (Plack et al., 2004). Indeed, even the NH listeners showed a wide range of compression estimate values ranging from highly compressive (slope of 0.085) to essentially linear (slope of 1.38). Although we did not measure TMC in the absence of TEN, it appears unlikely that the use of TEN can explain the variability in these results since another recent study (which tested listeners in quiet) showed similar range of variability in TMC-estimated compression for NH listeners (Poling et al., 2011).

Therefore, while the present results suggest no relationship between estimates of compression and MR for speech or pure tones, the finding of highly variable responses using the TMC technique calls into question its validity as a measure of cochlear compression. In other words, it is possible that there is actually a relationship between compression and MR, but that the TMC method is not a suitable means of estimating
compression. It is also possible that the TMC results are tapping into some underlying cochlear dysfunction that is not apparent from simple threshold estimation. However, it seems unlikely that this is the case for listeners with audiometric thresholds in the 5-20 dB HL range. Further work is required to clarify if the wide range of TMC responses is due to spurious measurement variability or is actually being influenced by some other, as yet unknown, underlying variable.

Other studies have suggested that reduction in MR for HI listeners may be related to impaired temporal resolution (e.g., Festen and Plomp, 1990; Bacon et al., 1998; George et al., 2006; Jin and Nelson, 2006). Some studies have evaluated forward masking recovery and attempted to relate it to MR with varying results (e.g., Dubno et al., 2002; 2003; Jin and Nelson, 2006). One problem with using recovery from forward masking as an estimate of temporal resolution is that it also includes the effects of compression. The off-frequency TMC function is assumed to reflect the recovery rate from forward masking independent of compression. Therefore, the slopes of the off-frequency TMC functions were compared to the MR results to see if there was a correlation. While there was a trend for more MR with faster recovery (i.e. steeper slopes), this trend did not reach significance.

In summary, the present study results were not able to explain degree of MR (for speech or pure tone stimuli) based on degree of cochlear compression as estimated by the TMC technique.
D. Outlook and future research

There is a trend for faster recovery from forward masking (as estimated by the off-frequency TMC function) to be associated with higher magnitude of MR for speech, but this trend did not reach significance. Data from studies such as Jin and Nelson (2010) that have shown comparable results for speech in gated noise as well as gated speech in quiet suggest that perhaps forward masking is not a sufficient predictor of performance in gated noise. One factor that was not evaluated in the present study, but that had been shown to be related to MR is spectral resolution (Jin and Nelson, 2010). Spectral resolution is also strongly related to behavioral compression estimates made using the forward GOM method (Moore et al., 1999b). It may be that measures of frequency selectivity provide a more robust, albeit more indirect, measure of cochlear compression. Further study is required in this area to determine if degree of spectral resolution may predict degree of MR in a modulated background.

The present study supports the assumption that the addition of background noise to elevate thresholds of NH listeners does not significantly alter the underlying BM compression characteristics. However, there may be other repercussions to using background noise that were not specifically evaluated here, such as the effects of introducing a higher neural firing rate with the presence of the noise, or activating the MOC efferent system, that is not found or is reduced in listeners with HL.

The use of AFM requires further study, especially at high levels as would be required for HI listeners, before it is used as a means of comparing compression between NH and HI listeners. The finding that NH listeners did not show results consistent with
compression for a 4000 Hz signal makes interpretation of similar results in HI individuals problematic. In other words, if a HI listener did not show results consistent with compression at 4000 Hz using the AFM method, it cannot be said whether this is due to the technique being used versus detection of an actual loss of compression at that frequency.

Overall, these results support the findings in the existing literature that NH listeners show more MR for speech and tonal signals, even when audibility is accounted for. The results do not point strongly to either behavioral estimates of compression or temporal resolution as a means of explaining this discrepancy in performance. However, the estimates of compression were highly variable, suggesting more work is necessary in this area to better understand how to appropriately use this technique to define cochlear compression for both NH and HI listeners and making it difficult to completely rule out underlying basilar membrane compression as a factor in MR. It may be that another measure strongly related to basilar membrane compression, such as spectral resolution, may provide a more reliable test.
References


Rosengard, P. S. (2004). "Relationship between measures related to the cochlear active mechanism and speech reception thresholds in backgrounds with and without spectral and/or temporal fluctuations.," (Massachusetts Institute of Technology, Cambridge).


