ABSTRACT

Title of Document: SPECTROTEMPORAL MODULATION SENSITIVITY IN HEARING-IMPAIRED LISTENERS

Golbarg Mehraei, Master of Science, 2009

Directed By: Professor, Dr.Shihab Shamma, Department of Electrical Engineering

Speech is characterized by temporal and spectral modulations. Hearing-impaired (HI) listeners may have reduced spectrotemporal modulation (STM) sensitivity, which could affect their speech understanding. This study examined effects of hearing loss and absolute frequency on STM sensitivity and their relationship to speech

intelligibility, frequency selectivity and temporal fine-structure (TFS) sensitivity. Sensitivity to STM applied to four-octave or one-octave noise carriers were measured for normal-hearing and HI listeners as a function of spectral modulation, temporal modulation and absolute frequency. Across-frequency variation in STM sensitivity suggests that broadband measurements do not sufficiently characterize performance. Results were simulated with a cortical STM-sensitivity model. No correlation was found between the reduced frequency selectivity required in the model to explain the HI STM data and more direct notched-noise estimates. Correlations between low-frequency and broadband STM performance, speech intelligibility and frequency-modulation sensitivity suggest that speech and STM processing may depend on the ability to use TFS.

SPECTROTEMPORAL MODULATION SENSITIVITY IN HEARING-IMPAIRED LISTENERS

By

Golbarg Mehraei

Thesis submitted to the Faculty of the Graduate School of the

University of Maryland, College Park, in partial fulfillment

of the requirements for the degree of

Master of Science

2009

Advisory Committee:

Professor Dr. Shihab Shamma, Chair

Dr. Joshua Bernstein

Dr.Monita Chatterjee

Acknowledgements

This work was supported by a grant from the Oticon Foundation. Work was performed in the Psychoacoustic Laboratory of the Speech and Audiology department at Walter Reed Army Medical Center, Washington, DC, under the direction of Joshua Bernstein (Walter Reed) and Shihab Shamma (UMCP). I would like to thank Van Summers, Matt Makashay and Sandeep Phatak (Walter Reed) for providing the Notched-noise ERB, FM detection and speech intelligibility data. I would also like to thank Marjorie Leek, Sarah Melamed, Michelle Molis and Erick Gallun (National Center for Rehabilitative Auditory Research, Portland-VA, OR) for providing data for several of the listeners in all of the experiments and Ken Grant, Doug Brungart and Elena Grassi (Walter Reed) for general consultations.

Special thanks to Dr. Joshua Bernstein for being an exceptional mentor and introducing me to the field of Hearing & Speech and Dr. Shihab Shamma and Dr.Monita Chatterjee for their guidance. Additionally, I would like to thank my parents, Kobra Yaranivand and Parviz Mehraei and my brother Payam Mehraei for their encouragement and love. Finally, thanks to all my friends for supporting me throughout the good and bad days. Special thanks to Hoda Eydgahi, Ruxandra Luca, and Keesler Welch for telling me to hold on when times got rough. I am privileged to have all of you in my life.

The opinions and assertions presented are the private views of the authors and are not to be construed as official or as necessarily reflecting the views of the Department of the Army, or the Department of Defense.

Table of Contents

List of Tables

List of Figures

Figure 5: STM threshold difference between the broadband conditions and corresponding octave-band conditions for both NH and HI listeners. The top panel plots are results for an upward-moving ripple and bottom panel plots are results of a downward-moving ripple. Note that the HI data has been horizontally shifted on the plots for a clearer comparison between the two groups. Line through 0 depicts no difference between broadband performance and the octave band performance. Negative values indicate poorer sensitivity in the narrowband case.

Figure 6: Subject 250 sensitivity measurements of certain ripple conditions at the 500Hz octave region before and after low frequency flanking noise was added to the stimuli. The subject's performance significantly decreases once the extended masking noise is added. The biggest change is seen in the <32Hz,4cyc/oct> condition. The flanking noise was also extended at the octave region centered at 4000Hz; however, no significant change in sensitivity was observed.

Figure 7: Collapsed STM sensitivity data. (Left panels) Temporal modulation sensitivity. (Right panels) Spectral modulation sensitivity. (no scale 4)

Figure 8: Process of the early stage of the auditory model. This stage consists of the periphery filterbank, the transduction stage and a lateral inhibition process (Wang, Shamma 1992).

Figure 9: A) The relationship between the psychoacoustic NH STM sensitivity estimates and the corresponding cortical response magnitude of the Gammatone filterbank defined by Glasberg and Moore (1990). Filter ERBs were adjusted based on the notched-noise ERB measurements for the NH listeners. B) The one-to-one relationship between STM data and the predicted STM thresholds based on cortical magnitudes and exponential fit in panel A.

Figure 10: Transformation of auditory spectrogram into plot of STRF in the central stage of the model.

Figure 11: a) Auditory spectrogram of ripples 4Hz, 1cyc/oct, upward direction at CF=500Hz BW=1 octave. b) Scale-rate plot of the ripple at the cortical stage. Note that negative value of the rate in the scale rate plot refers to the upward direction of the ripple in the model.

Figure 12: Comparison of average raw data with model for the HI group. (Left panel): Comparison of the STM sensitivity data with predicted thresholds based on the NH model peripheral filters. (Right panel): Comparison of data and model predictions with the bandwidths of the peripheral filters adjusted (i.e. broadened) to fit the data.

Figure 13: Comparison of raw data with model for HI subject 15. (Left panel): Comparison of the STM sensitivity data with predicted thresholds based on the NH

model peripheral filters. (Right panel): Comparison of data and model predictions with the bandwidths of the peripheral filters adjusted (i.e. broadened) to fit the data.

Figure 14: Comparison of Speech Intelligibility scores and STM sensitivity across absolute frequency. Speech was presented in stationary noise with a SNR of 0dB. The p values listed in each panel are one-tailed p values. It was assumed a priori that the correlations can only go one way - listeners who are worse at one task will also be worse at the other. Last plot compares broadband STM sensitivity to Speech intelligibility scores.

Figure 15: Comparison of model predicted ERB estimate to notched-noise ERB estimated for each HI listeners at each frequency region.

Figure 16: Comparison of model predicted ERB estimate to notched-noise ERB estimated for average HI listener.

Figure 17: A comparison between STM sensitivity and FM detection. Each plot compares the STM data for that absolute frequency region with the FM data that uses the corresponding carrier frequency.

Figure 18: A comparison between broadband STM sensitivity and FM detection. Each plot corresponds to a different FM carrier frequency.

Chapter 1: Introduction

Speech identification is often characterized by its formant peaks, spectral edges, and amplitude modulations at onsets/offsets. These significant features contribute to the energy modulations seen in speech spectrograms, both in time for any given frequency channel, and along the spectral axis at any instant. It has been suggested that speech intelligibility is highly dependent on these low spectral modulation densities and temporal modulations rates (<30Hz) that reflect the phonetic and syllabic rate of speech (Houtgast and Steeneken, 1985; Drullman et al., 1994a,b; Henry et al 2005). Although sensitivity to temporal and spectral modulation has been investigated extensively, these two measurements are frequently studied separately. Measurements of purely temporal and spectral modulations in normal hearing (NH) and hearing impaired (HI) listeners generally exhibit a low pass response, reflecting the limits of temporal and spectral processing by humans (Viemeister, 1979; Green 1986).

The temporal fluctuations of speech waveforms are important for providing information about segmental speech properties such as consonant articulation and about prosodic aspects of speech. Smearing of the temporal envelope causes severe reduction in sentence intelligibility (Drullman et al., 1994a, b). Studies investigating the effect of hearing impairment on temporal resolution have generally found that performance of temporal modulation detection for a broadband noise carrier is not significantly affected in listeners with sensorineural hearing loss for signals presented at equal spectrum levels or at equal SL to NH listeners (Bacon and Viemeister, 1985;

Bacon Gleitman, 1992; Moore et al, 1992). In the cases that have shown weaker temporal sensitivity in HI listeners, this was largely a consequence of the fact that high frequencies were inaudible for these listeners as most subjects had greater high frequency hearing loss. When the modulated noise was low pass filtered, simulating the effects of threshold elevation at high frequencies, NH listeners also showed a reduced ability to detect high modulation rates (Bacon and Viemeister, 1985). Overall, similar temporal modulation transfer functions (TMTFs) seen between NH and HI listeners at equal spectrum levels suggests that temporal resolution is not significantly affected by hearing loss.

In contrast to their relatively normal temporal processing abilities, there is evidence that listeners with cochlear damage have spectral modulation deficits as a result of broader auditory filters compared to NH listeners (Glasberg and Moore 1986). As a result of these broader filters, smearing of spectral details in the internal representation of an acoustic signal may occur. This smearing causes an amplitude reduction between the peaks and valleys of a signal resulting in identification difficulties of the frequency locations of spectral peaks. The locations of spectral peaks are important cues for speech identification, and as such, the spectral flattening resulting from the broader filters may result in impaired speech perception ability. Listeners with normal hearing show peak spectral sensitivity between 2-4cycles/octave with a substantial increase in modulation detection threshold for higher modulation frequencies due to limited spectral resolution (Bernstein and Green, 1987a,b;1988; Summers and Leek, 1994; Amagai et al 1999; Chi et al., 1999, Eddins and Bero, 2006; Hillier, 1991). In comparison, spectral sensitivity in HI

listeners maintains the same low pass shape but performance is relatively worse (Summers and Leek 1994). Specifically, Summers and Leek (1994) reported that relative bandwidths measured for HI subjects fell outside the range of normal bandwidths for filters centered at 3000Hz and 1000Hz and that reduced performance of the individual hearing impaired listeners in the spectral modulation detection task was correlated to the extent to which their filters were broadened.

Reduced spectral resolution may be a significant factor that limits speech perception for HI listeners by disrupting perception of the spectral shape of speech sounds. Studies have shown that in NH listeners, spectral smearing reduces speech intelligibility (Baer & Moore, 1993,1994; Ter Keurs et al 1992,1993). Henry et al (2005) found that the degree of spectral peak resolution required for accurate vowel and consonant recognition in quiet is about 4 cyc/oct and that spectral peak resolution poorer than 1–2 cyc/oct may result in highly degraded speech recognition. In addition, most current models of speech intelligibility focus on frequency content (e.g. AI, SII) (ANSI S3.5-1997, American National Standards Institute, New York), and in some cases, temporal modulations (Speech Transmission Index, Steeneken and Houtgast, 1980, 1998). Since frequency selectivity is reduced in HI listeners, it may be necessary to include the spectral dimension in quantitative models of speech intelligibility for HI listeners. This approach has only been applied for NH listeners (Elhilali et al 2003).

While studies have established much about the effects of hearing impairment on spectral and temporal resolution separately, these one dimensional MTFs do not directly reflect the characteristics seen in natural sounds that often have combined

spectrotemporal modulations. For example, speech is rarely a flat modulated spectrum nor is it a stationary peaked spectrum, but rather it is a spectrum with dynamic peaks. Chi et al (1999) measured sensitivity to combined spectral and temporal modulations using spectrotemporal "ripple" stimuli in NH listeners. They showed that the combined spectrotemporal MTFs are separable (i.e. product of spectral and temporal MTFs) and that the measurements replicate the low pass characteristics of purely temporal and spectral MTFs seen in previous studies. In addition, they found that a model combining peripheral filtering with the cortical STM model, which models the representation of spectrotemporal modulation in the auditory cortex, was able to account for the observed roll off sensitivity with increased spectral modulation density. Based on these measurements, it has been shown that speech intelligibility by normal hearing listeners in noise and reverberation can indeed be predicted by a model of spectrotemporal modulation (STM) strength in the auditory periphery (Elhilali et al 2003). Hence, the clarity of joint spectrotemporal modulations is quite significant in speech perception.

Listeners with sensorineural hearing loss have extreme difficulty understanding speech in background noise. Although amplification via a hearing aid compensates for speech perception to some extent, for those HI listeners with hearing loss in the moderate range, audibility does not account for the entire deficit in speech perception; thus, suggesting abnormalities in the perceptual analysis of sound at suprathreshold levels (Henry et al 2005). Among these suprathreshold distortions is the possible impairment in processing complex STMs. To this date, no attempts have been made to characterize STM sensitivity in listeners with hearing loss.

Furthermore, previous studies of spectrotemporal modulation and spectral modulation detection have only used broadband carriers as their stimuli to test NH listeners (Chi et al 1999; Summers and Leek 1994; Bernstein and Green 1987a, b;1988). It is important to look across frequency regions in both NH and HI listeners: there is no indication from perception of the broadband stimuli which frequency region might be supporting STM detection. Sensitivity to STM as a function of absolute frequency can be particularly important in parametrizing the ability to process spectrotemporal modulations due to processing differences across the cochlea partition. Eddins and Bero (2006) reported that spectral modulation detection was not strongly dependent on carrier frequency region with the exception of carrier bands restricted to very low audio frequencies. However, this dependence has not yet been determined for STM. Moreover, differences in hearing loss across frequency in HI listeners may differentially affect STM sensitivity.

The present study aimed to determine the extent which STM sensitivity is compromised in listeners with sensorineural hearing loss and if there is variation across tonotopic frequency in STM sensitivity for NH and HI listeners. The STM detection threshold was determined by estimating the modulation depth required to discriminate a spectrally flat standard noise from a signal that was similar to the standard noise except for added spectral and temporal modulations (Chi et al 1999). This study measured NH and HI sensitivity to the STM modulations over perceptually important spectral and temporal ranges with broadband and octave band carriers. We hypothesized that the spectral and temporal dimensions are separable for

HI listeners as was shown for NH listeners by Chi et al (1999) and that HI listeners will have deficits in the spectral but not the temporal dimensions.

Additionally, the study attempted to predict HI listeners' STM sensitivity based on performance in a standard measure of frequency selectivity using the notched-noise technique (Rosen and Baker, 1994). The two measures were related using the auditory model approach of Chi et al (1999). The purpose was to determine the extent to which differences in STM sensitivity between NH and HI listeners can be explained in terms of peripheral frequency selectivity.

Chapter 2: Methods

Psychoacoustic spectrotemporal modulation transfer functions (STMTFs) were measured for NH and HI listeners for octave-band and broadband (four-octaves) stimuli. A two alternative forced choice adaptive task, where one interval contained unmodulated noise and the other contained the STM stimulus, was used to estimate STM detection thresholds. STM sensitivity was characterized in terms of the modulation depth required for modulation detection.

Spectrotemporal ripple Stimuli

Broadband Ripples

The broadband ripple stimuli consisted of equal amplitude tones that were equally spaced along the logarithmic frequency axis spanning four octaves (0.3535-5.656kHz). Sinusoidal amplitude modulation was applied to each carrier tone. Spectral modulation was induced by adjusting the relative phase of the temporal modulation for each successive carrier tone yielding a sinusoidal envelope at each point in time along the log frequency axis. The STM is fully characterized by equation (1) where *S* represents the amplitude of each carrier tone as a function of time and frequency, ω is the ripple velocity defined as the number of ripple cycles-per-second, and Ω represents the spectral density (cycles/octave). The position, *x*, in octaves is defined as $x = \log_2(\frac{f}{f0})$ with f_0 being the lower edge of the spectrum and *f*

the frequency (octaves). The phase, φ, in this spectrum is selected randomly on each stimulus representation. The amplitude (A) of each carrier tone at each point in time is determined by the modulation depth (0=no modulation and 1=100% modulation).

$$S(x,t) = A \cdot sin(2\pi(\omega \cdot t + \Omega \cdot x) + \varphi). \quad (1)$$

The direction of the ripple was determined by ω; a negative ω corresponds to a ripple envelope drifting upward and vice versa. Example auditory spectrograms for various STM stimuli are shown in Fig. 1. The auditory spectrograms are the time-frequency representations of the stimuli passed through an auditory model (Chi et al 1999) representing peripheral processing in four stages (filtering, half-wave rectification, lowpass filtering, lateral inhibition discussed further in Chapter 4). The patterns seen in the frequency (vertical) dimension of the auditory spectrograms depict the spectral modulation of the signal while the patterns in the time (horizontal) dimenstion represent the temporal modulation. For example, in Fig.1A, there are four spectral peaks across four octaves in the vertical dimenstion (1 cyc/oct) and two cycles across 500ms in the horizontal dimension (4Hz). The sweeping direction of the spectrotemporal modulated signal is also seen in the auditory spectrograms where the upward direction (Fig. 1A) depicts a negative ω and the downward direction represents a positive ω (Fig.1B).

Narrowband Ripples

Narrowband ripples were constructed in the same way as the broadband stimuli as described in equation (1) except that the modulated carrier tone frequencies were limited to one octave centered at 500, 1000, 2000 or 4000Hz. In the remaining regions of the four-octave band associated with the broadband ripples, standard noise (i.e. 1000 logarithmically spaced random-phase tones per octave) was presented, with a level per component lower than the tones in the modulated region. This was done so that performance in the narrowband conditions could be compared to performance in the broadband case while limiting spectral cues at the edges of each octave band that would not have been available in the wideband case. These possible spectral cues could arise due to modulation components extending the bandwidth of the carrier region. The unmodulated noise, extending the remainder of the four octaves, was 15dB lower than the modulated octave band to draw listener's attention to the modulation. Figures 1C and D show auditory spectrograms for two narrowband STM stimuli (1C: 4Hz, 1 cyc/oct centered at 500Hz,1D: 4Hz, 2 cyc/oct centered at 4000Hz).

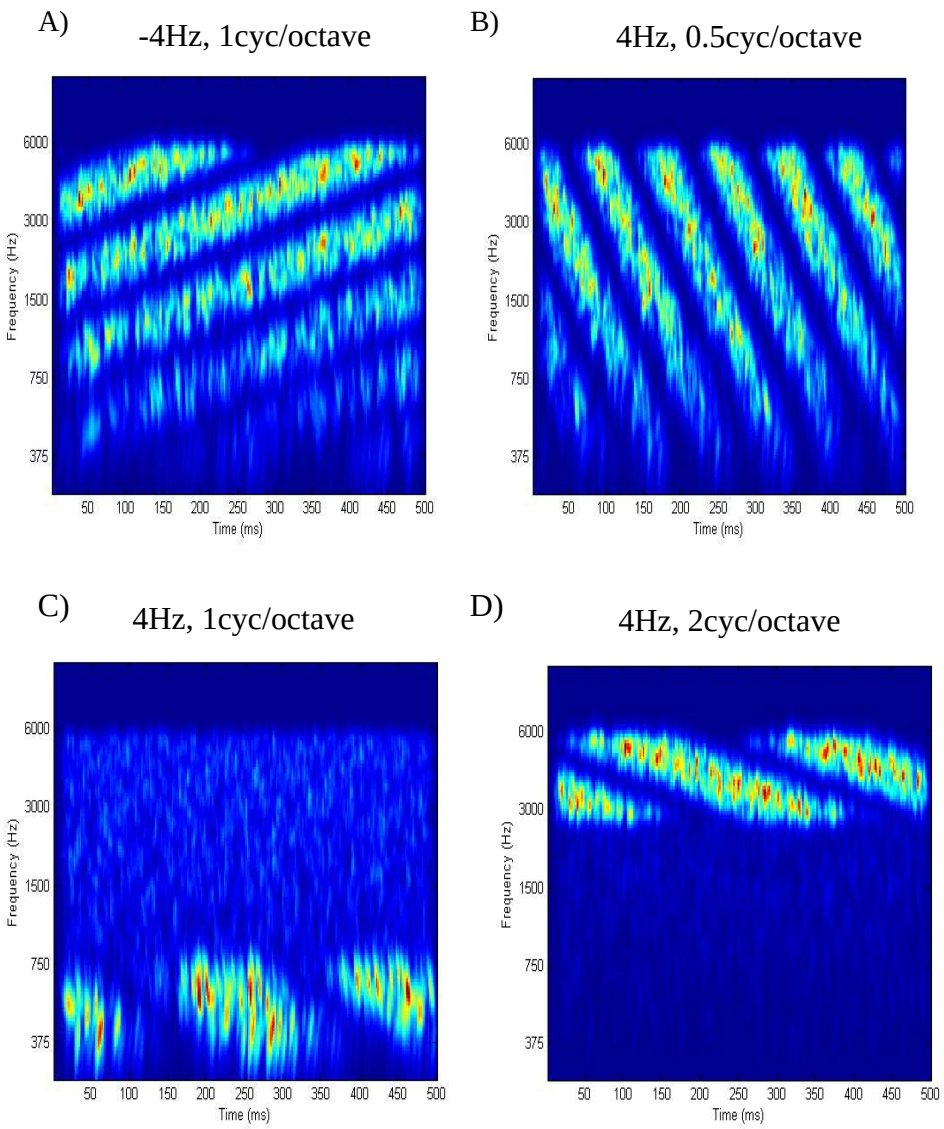

Figure 1: a) Auditory Spectrogram of broadband STM with rate=-4Hz, scale=1cycle/oct, upward direction. b) Broad band stimulus rate=12Hz, scale= 0.5 cycle/octave , downward direction. c) Spectrogram of octave band STM centered at 500Hz with rate=4 Hz, scale=1 cyc/oct, downward direction c) octave band centered at 4000Hz with rate=4Hz, scale= 2cyc/oct, downward direction.

Testing Procedures

STM detection thresholds were measured using a two-alternative forced choice adaptive procedure. Subjects were asked to discriminate between a spectrally flat stationary standard noise and a STM noise randomly presented to either interval

(p=0.5). The modulation depth was varied in a three down one up adaptive procedure tracking the 79.4% correct point (Levitt 1971). The modulation depth of the STM signal was tracked during each run and was reported in dB as described in equation (2) where *m* is the modulation depth.

$$dB = 20log_{10}(m) \quad (2)$$

The starting modulation depth for each run was 1 (full modulation). The modulation depth was adjusted by 6dB until the first reversal, 4 dB for the next two reversals, and 2 dB for the last six reversals, for a total of nine reversals per run. The threshold was determined by taking the mean of the modulation depth (in dB) of the last six reversal points. If the subject was unable to detect the signal at the maximum modulation depth more than five times in any run, the run was terminated and a threshold was not collected. The signal and the standard noise were presented at a nominal level of 80dB SPL/octave to the test ear. This level was chosen such that both groups can hear the stimuli clearly without the signal being too loud. As shown in the audiograms in Fig.2, a level of 80dB SPL/octave is above threshold for both HI and NH listeners. Additionally, an 80dB SPL/octave level was used for both groups to reduce the influence of level on frequency selectivity. The overall presentation level was roved randomly across trials over a ±2.5dB range to reduce the effectiveness of possible loudness cues.

Two runs were presented for each combination of density(0.5 , 1 , 2 , 4 cyc/oct), rate(4,12,32Hz), frequency (broadband or 0.5 , 1 , 2 , 4kHz narrowband),

and direction (Ω, ω). If the two threshold estimates for any of combination differed by 3dB or more, an additional threshold was collected for that condition. Additionally, a third run was conducted if one of the two runs was terminated due to frequent incorrect responses at full modulation. A fourth threshold estimate was performed if the two of the three threshold estimates collected for a specific condition differed by more than 6 dB. A short visual feedback was displayed after each trial in that particular run.

Subjects

Eight NH listeners (four female, mean age: 44.5, age range: 24-60) and twelve HI listeners (one female, mean age: 75.7, age range: 70-87) took part in this study. Of the twenty listeners, fifteen were tested at Walter Reed Army Medical Center, Washington DC, and five at the National Center for Rehabilitative Auditory Research, Portland, OR. The mean audiogram (±1 standard error or deviation) for each listener group is shown in Fig.2. NH listeners had pure-tone threshold better than or equal to 20 dB HL at octave frequencies between 250-8000Hz plus 3000 and 6000 Hz. On average, HI listeners had high frequency hearing loss, and near normal thresholds below 1000Hz. The ear tested for each HI listener was determined by his or her audiogram: in general, the better ear was tested. In some cases where a HI listener had nearly equal audiograms for both ears, the decision was determined by the ear that yielded a lower detection threshold for a 1000Hz tone. NH listeners were tested in the ear of their choice.

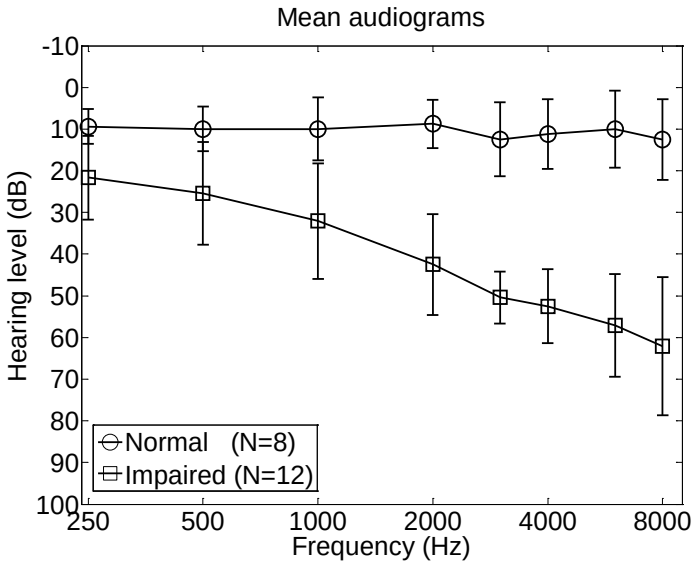

Figure 2: Mean audiogram for twelve HI and eight NH listeners.

Training

Each subject completed a minimum of an hour of training. Training runs were similar to the experiment runs with the exception of an additional interval. The listener was asked to identify the modulated stimulus randomly presented in interval two or three. The first interval always contained the standard noise reference. The purpose of this reference was to help the listener to better identify the stimulus among the three intervals and to become familiar with the differences between the standard noise and the STM signals. Training was done on a pseudorandom sampling of the spectrotemporal conditions presented in the experiment, with emphasis placed on higher scales and lower frequency regions where listeners experienced the most difficulty. The training period continued for each listener until performance had stabilized.

Sounds were generated digitally with a 32-bit amplitude resolution and 48848Hz sampling rate. The 500ms long digitized samples were ramped on and off (20-ms raised cosine) and normalized in level so that all stimuli had the same average root-mean-squared amplitude. The ramping of the signals helped prevent the production of sudden audible “clicks” during the presentation. The digital audio signal was sent to an enhanced real-time processor (TDT RP2.1) where it was stored in a buffer. The audio signal was then converted to analog by the TDT RP2.1 and was passed through a headphone buffer (TDT HB7) before being presented to the listener through one earpiece of Sennheiser HD580 headset. To prevent detection of the target speech signal in the contralateral ear, standard uncorrelated noise with a level 20dB below that of the target signal was presented to the non-test ear. The listener was seated inside a double-walled sound attenuating chamber.

Chapter 3: Results

Mean STM detection thresholds across eight NH (grey symbols) and twelve HI (open symbols) listeners are shown in Fig. 3 as a function of spectral modulation scale (Ω, horizontal axis) and temporal modulation rate (ω, shapes) for upward(upper plots) and downward (lower plots) moving ripples. More negative values in Fig. 3 indicate better performance, with STM detectable for smaller modulation depths.

Overall, STM sensitivity in the spectral and temporal dimensions demonstrated the lowpass characteristics shown previously (Chi et al 1999). As shown in Fig.3, sensitivity generally decreased as a function of increasing scale (horizontal axis), increasing rate (squares to circles to triangles), decreasing absolute frequency (first through fourth panel in each row), and hearing loss. To confirm these trends statistically, an analysis of variance (ANOVA) was implemented on the narrowband STM measurements and will be discussed in conjunction with the results. The analysis included four within-subject factors (rate, scale, direction, frequency) and one between-subjects factor (hearing loss). However, the ANOVA analysis was complicated by floor performance for several combinations of conditions and an individual subject who unexpectedly had high senstivity in some high temporal modulation rates.

Although individual listeners generally showed the lowpass characteristic in the temporal and spectral domain, one listener demonstrated uncharacteristically high senstivity to 32Hz and 12Hz ripples at 500Hz. This subject informally reported that those stimuli did not sound modulated but instead were discriminable based on pitch

differences. Modulation is imposed on each tone carrier by creating sidebands above and below the carrier frequency. In most cases, the presence of noise in the non-modulated regions likely masked the ability to detect these spectral changes. However, for the 500Hz, 4000Hz narrowband and the broadband conditions, no additional noise was present above (broadband) or below (broadband) the modulated regions. In the 500-Hz and broadband cases, the 32Hz modulation would have extended the lower frequency edge of the stimulus (353Hz) downward by about 10%, yielding a potentially salient spectral-edge cue. The possible use of a spectral-edge cue in the 500Hz condition was estimated for this NH listener in Fig. 6. STM sensitivity is shown with and without the addition of an octave-wide flanking noise with a level 15dB below that of the modulated band, just below the 500Hz region. The addition of the flanking noise yielded a significant reduction in sensitivity for the <32Hz,4cyc/oct> condition (black squares) supporting the idea that this listener relied on spectral-edge cues for this condition. No other listener demonstrated a trend of better performance for 32Hz than for lower rates for any combination of spectral scale and frequency region. This listener's data was not included in the plots shown in Fig.3 nor in the statistical or modeling analysis.

Effects of Scale and Rate

As shown in Fig. 3, NH and HI listeners exhibited a decrease in sensitivity as the spectral modulation Ω increased. Generally, both groups maintained high sensitivity across frequency regions to low scales (0.5-1 cyc/oct) and diminished sensitivity at 4 cyc/oct. In the temporal domain, sensitivity was generally maximum

at a low temporal rate of 4 Hz (squares) and worsened at 32Hz (triangles) in both directions. However, sometimes performance was better at 12Hz than 4Hz suggesting that maybe the signal duration was not long enough to detect the 4Hz modulation. This is in agreement with previous studies (Viemesiter 1979) where a bandpass characteristic with a reduction in performance for very low temporal rates was found. The effects of temporal and spectral modulation on STM sensitivity were evident in the ANOVA STM (Table 1) where both factors were shown to be significant. The temporal functions generally maintained their shape across all values of Ω as shown in Fig. 4. As the Ω increased, the temporal transfer functions were shifted upwards relative to each other reflecting the decrease in senstivity to high spectral modulations in both ripple directions and across all frequencies as seen in Fig. 3. However, this was not always the case, as STM sensitivity was not strictly driven by spectral modulation or temporal modulation independently but by the combination of the two, evidenced by a significant interaction between scale and rate (Table 1).

Effects of Absolute Frequency

The data in Fig. 3 shows a clear absolute frequency effect for both NH and HI groups where STM sensitivity improved as the absolute frequency increased. This effect was verified by a significant main effect of frequency in the ANOVA. Although, the many significant interactions between frequency and other factors (frequency and rate, frequency & scale & rate) suggest that the frequency effect was larger for certain combinations of rate and scale (Table 1) . This could be due, at least in part, to floor effects at 500Hz and 1000Hz that occurred for higher rates and scales.

Some individual subjects were unable to successfully detect certain combinations of STM ripples and so a threshold that was not collected for these trials was assigned to be 0dB (100% modulation depth). Fig. 3 denotes the STM ripples exhibiting floor effects by black shading. Floor effects were generally seen in the higher rate and scale combinations in both directions, specifically in <32Hz,2cyc/oct>, <4Hz, 4cyc/oct>, in the 500Hz and 1000Hz octave bands. Of the eight NH listeners, a threshold could not be estimated for two listeners for the <32Hz, 4cyc/oct> at 500Hz, three listeners for the <-32Hz,4cyc/oct> at 1000Hz, and two listeners for the <-32Hz, 4cyc/oct> at 1000Hz conditions. Similiarly, of the twelve HI listeners, a threshold could not be estimated for two and three listeners for the <4Hz, 4 cyc/oct> and <32Hz, 4 cyc/oct> 1000Hz conditions, respectively. In the 500Hz region, three HI listeners were unable to detect condition <-4Hz, 4 cyc/oct> and two listeners were unable to detect <-32Hz, 2cyc/oct>. Because these floor effects were mostly seen in combinations with 4cyc/oct, the ANOVA analysis was performed without this high scale. However, the exclusion of this scale did not eliminate floor effects for the 2 cyc/oct conditions for the ANOVA. Furthermore, because the maximum modulation depth was not allowed to exceed 0 dB (full modulation), sensitivity estimates may be artifically low even in some cases where a run was not terminated before a threshold could be collected.

A comparison between the broadband (right panels of Fig.3) and narrowband data reveals that the broadband performance was similar to the STM performance at 4000Hz for both groups as shown in Fig. 3. Fig. 5 plots the difference between the STM detection thresholds for the broadband conditions and the corresponding thresholds for each octave-band condition. The largest differences are seen for the

500Hz conditions, while the differences between the broadband and 2000 and 4000Hz narrowband thresholds are near 0. Overall, the sensitivity differences seen between the broadband and 4000Hz conditions was quite small relative to the difference between the broadband and other narrowband frequency conditions. This suggests that wideband performance was largely determined by sensitivity in the higher frequency regions and that modulation in the low frequencies contributed little to the broadband STM sensitivity. Still, performance was better in the broadband than the 4000Hz narrowband case for some rate-scale conditions suggesting that lower frequency regions may have played some role in the broadband STM detection.

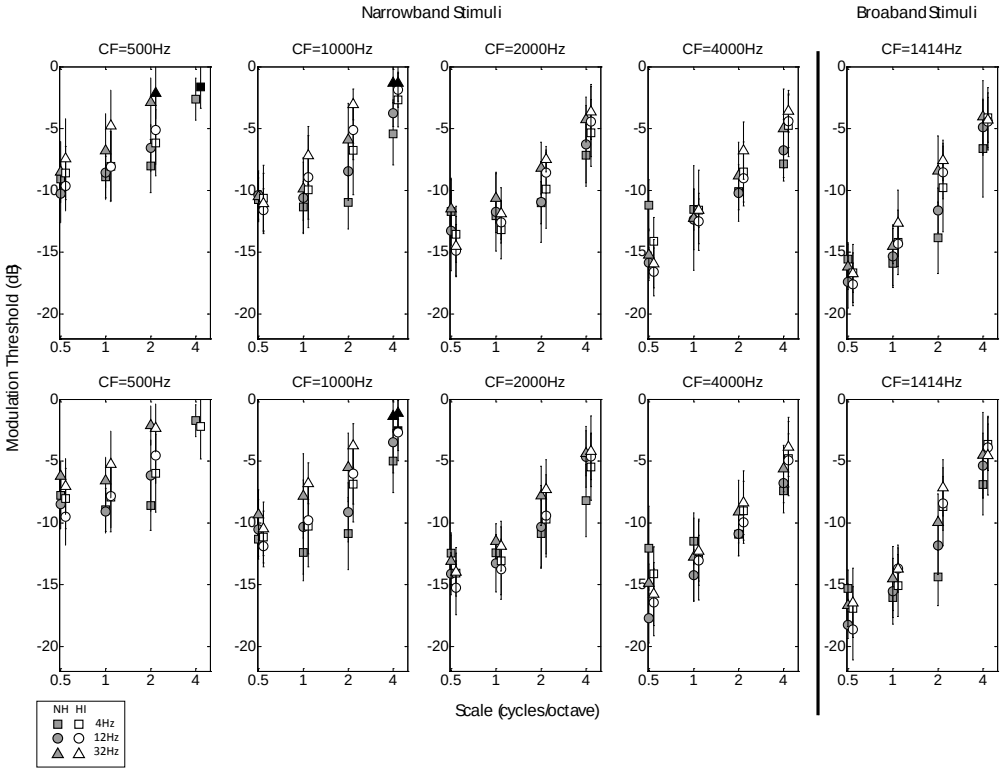

Figure 3: STM data for 12 HI (white) and 8 NH (grey) groups across frequencies. Notice that performance in the 4000Hz region is similar to the performance in the broadband region (last plot). The top panel plots are results for an upward-directed ripple and bottom panel plots are results of a downward-directed ripple. Note that the NH data has been horizontally shifted on the plots for a clearer comparison between the two groups. The black symbols represent conditions where floor effects were present. In addition, missing data from the 500Hz, 4 cyc/oct modulation

combinations indicate the conditions where pitch cues were present specifically <12Hz, 4 cyc/oct> and <32Hz, 4 cyc/oct> in both directions.

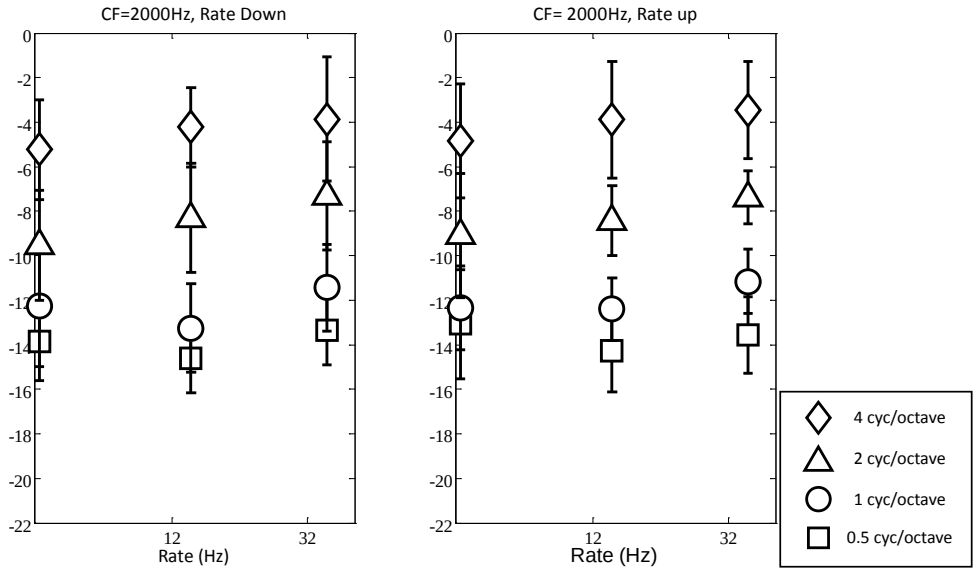

Figure 4: Sample STM data for octave band frequency region centered at 2000Hz for average HI listeners. Data plotted as a function of Rate (x-axis).

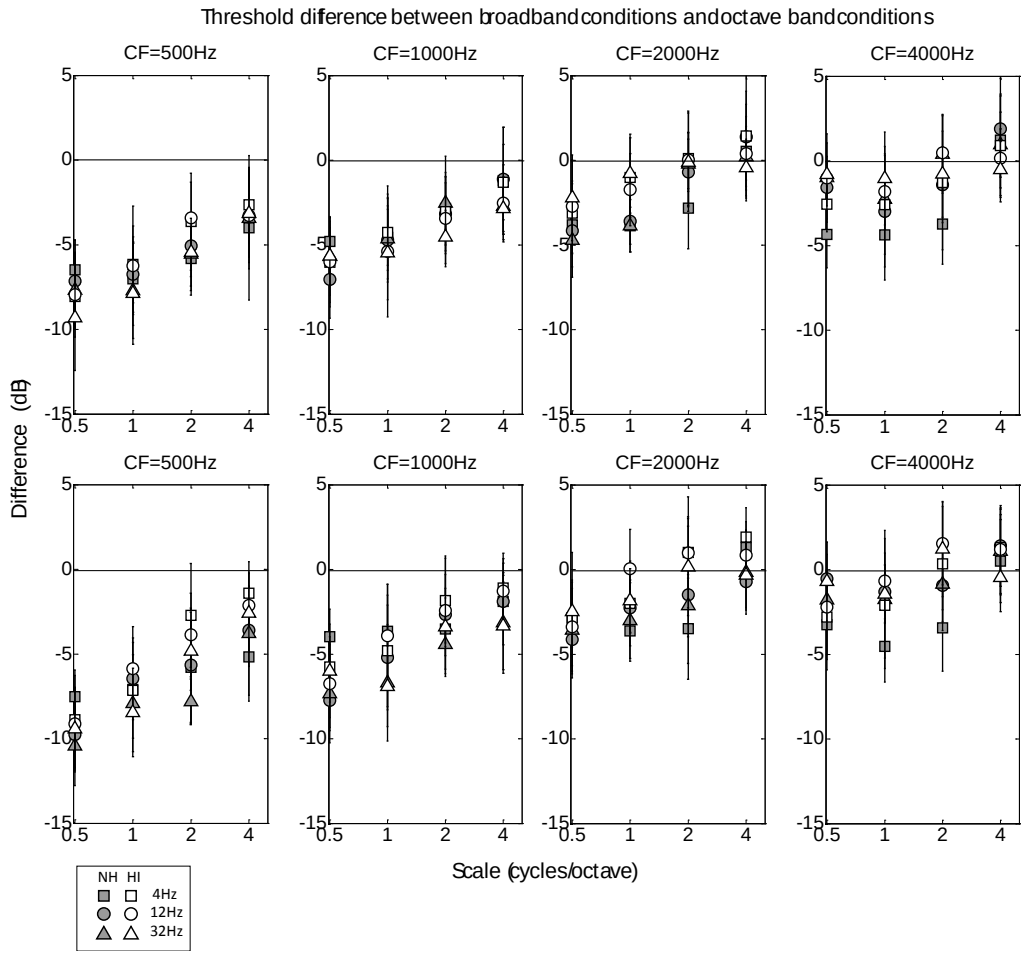

Figure 5: STM threshold difference between the broadband conditions and corresponding octave-band conditions for both NH and HI listeners. The top panel plots are results for an upward-moving ripple and bottom panel plots are results of a downward-moving ripple. Note that the HI data has been horizontally shifted on the plots for a clearer comparison between the two groups. Line through 0 depicts no difference between broadband performance and the octave band performance. Negative values indicate poorer sensitivity in the narrowband case.

Factor	Degree of Freedom	F Value	p-value no scale 4
Scale	1.791	145.475	**p<0.0005**
Rate	2	33.879	**p<0.0005**
Frequency	2.197	75.727	**p<0.0005**
Direction	1	2.257	p=0.150
Hearing Impairment	1	1.096	p=0.309
Hearing Impairment * Frequency	1.852	3.148	p=0.059
Hearing Impairment * Scale	1.471	16.503	**p<0.0005**
Hearing Impairment * Rate	1.8	0.107	p=0.880
Frequency*Scale	6	3.592	**p=0.003**
Frequency*Rate	5.884	26.780	**p<0.0005**
Scale*Rate	4	43.941	**p<0.0005**
Hearing Impairment * Direction	1	4.842	**p=0.041**
Frequency*Scale*Hearing Impairment	6	3.122	**p=0.007**
Frequency*Rate*Hearing Impairment	5.884	0.728	p=0.625
Frequency*Scale*Rate	9.615	2.928	**p=0.002**
Frequency*Direction*Hearing Impairment	2.836	1.790	p=0.163
Scale*Rate*Hearing Impairment	4	1.598	p=0.184
Scale*Rate*Frequency*Hearing Impairment	9.615	1.384	p=0.194

Table 1: ANOVA analysis for the raw STM data. Analysis excludes 4cyc/oct and NH listener 250. Significant effects (p<0.05) are indicated by ***boldfaced font****.*

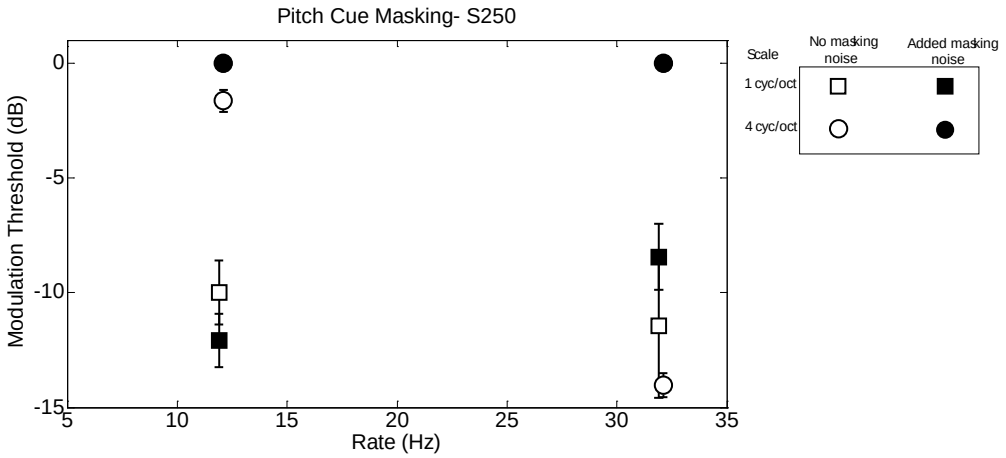

Figure 6: Subject 250 sensitivity measurements of certain ripple conditions at the 500Hz octave region before and after low frequency flanking noise was added to the stimuli. The subject's performance significantly decreases once the extended masking noise is added. The biggest change is seen in the <32Hz,4cyc/oct> condition. The flanking noise was also extended at the octave region centered at 4000Hz; however, no significant change in sensitivity was observed.

Effects of Hearing loss

Although there was no significant main effect of hearing loss, there were significant interactions between hearing loss and other variables. This suggests that the HI listeners are impaired, but only for certain combinations of conditions. Hearing impairment affected performance for some frequencies but not others as observed in Fig.3. However, this was not confirmed by a significant interaction between frequency and hearing loss in Table 1. Specifically, differences in sensitivity between the NH and HI group were observed mainly in the lower frequency regions of 500Hz and 1000Hz (Fig. 3). This is unexpected because of the sloping average audiogram of the HI group shown in Fig. 2, with more hearing loss at higher frequencies.

A significant interaction between HI and scale indicates that hearing impairment affected STM sensitivity with certain spectral modulation scales more than others. Furthermore, the three-way interaction between HI, scale, and frequency suggests that the effect of HI on spectral modulation occurs in some frequency regions. In contrast, hearing loss did not differentially affect sensitivity across temporal modulation rates, as indicated by a lack of significant interaction involving hearing loss and rate (Table 1).

Separating out the effects of rate and scale

To further investigate the effects of hearing impairment on STM sensitivity, singular value decomposition (SVD) was implemented to decompose the STM sensitivity data into spectral and temporal dimensions. The SVD expresses the STM sensitivity function as $m=U*\Lambda*V$ where Λ is the eigenvalue matrix and the U,V are the corresponding eigenvectors (Haykins, 1996). If the spectral and temporal sensitivity contributed independently to STM sensitivity, this analysis would yield only one significant eigenvalue. Due to the artifact seen in the raw data because of floor performance, the analysis did not include the scale 4 cyc/oct conditions. Across all listeners and frequencies, all of the non-primary eigenvalues were <19% of the primary eigenvalue suggesting that although there is some interaction between scale and rate (Table 1), most of the STM sensitivity data can be explained in terms of independent contributions from the temporal and spectral modulation sensitivity.

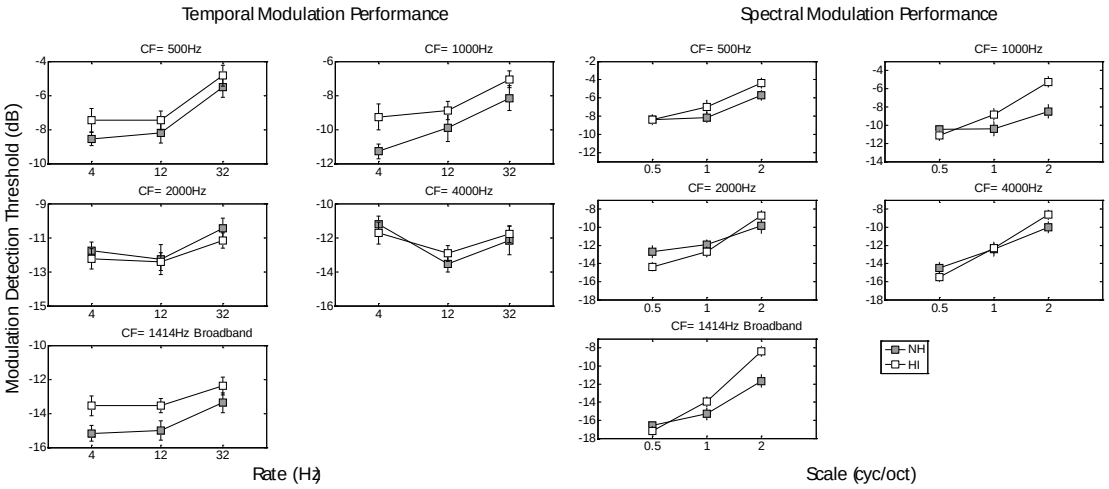

Figure 7: Collapsed STM sensitivity data. (Left panels) Temporal modulation sensitivity. (Right panels) Spectral modulation sensitivity. (no scale 4)

Because the SVD showed that temporal and spectral modulation sensitivity are independent, the STM data was collapsed by averaging the data across scale (Fig.7 left panel) or rate (Fig.7 right panel) to investigate the separate effects. A HI listener with limited frequency or temporal resolution would be expected to show performance that falls off more quickly with increasing scale or rate as they would not be expected to have trouble with relatively slow/broad modulations that fall within the limits of their spectral/temporal resolution abilities. It is only for the scales or rates that exceed their resolution limits where differences would be expected between NH and HI listeners. Therefore, we would expect the performance slopes to be steeper in these cases where HI listeners have reduced resolution. This was generally true for the spectral domain but not the temporal domain.

Comparisons made between the two groups when the STM data is collapsed across rate (Fig. 7, right panel) showed that HI performance is generally worse when compared to NH listeners across most of the frequency regions. Specifically, in the

500 and 1000Hz regions, differences in performance between the two groups became more profound at a high scale of 2cyc/oct demonstrating the spectral resolution limitations of HI listeners. This is consistent with the idea that HI listeners had reduced frequency selectivity in some frequency regions and reconfirms the significant interactions between spectral modulation and HI along with spectral modulation, frequency, and HI in the ANOVA analysis (Table 1). However, at higher frequency regions (2000 and 4000Hz), this trend is not as well defined. In fact, HI listeners are more sensitive to slower spectral modulations (0.5 cyc/oct) than NH listeners. This suggests that HI affects certain spectral modulation conditions more than others in some frequency regions because of poor frequency selectivity.

When the STM data is collapsed over scale (Fig. 7, left panel), HI performance is again seen to be impaired relative to the NH listeners at the lower (500 and 1000Hz) but not the higher frequency regions (2000 and 4000Hz). Within the 500 and 1000Hz region, HI listeners show slightly more impairment relative to NH listeners at a temporal rate of 4Hz than 32Hz relative to the NH group. In contrast, in the 2000 and 4000Hz regions, HI listeners are more sensitive to lower temporal rates than NH listeners. A trend toward poorer performance of HI listeners relative to NH at slow temporal rates in the lower frequency regions was not large enough to be captured by the ANOVA analysis, as there was no significant interaction involving hearing loss and temporal rate (Table 1).

Chapter 4: Model

Modeling Method

To further investigate whether the STM sensitivity results for HI listeners could be explained in terms of reduced frequency selectivity, the Neural System Laboratory auditory model was (Chi et al 1999) used to relate performance in complex spectrotemporal processing to basic peripheral processing in HI and NH individuals. The model consists of two stages: 1) an early auditory portion, which models the transformation of the acoustic signal into neural pattern activity and 2) a central stage that performs a STM analysis.

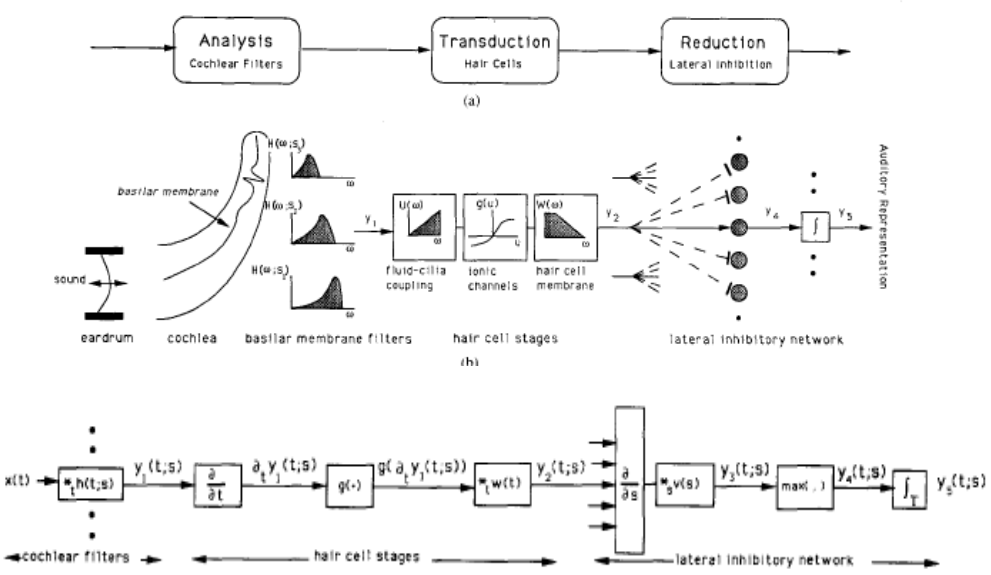

Figure 8: Process of the early stage of the auditory model. This stage consists of the periphery filterbank, the transduction stage and a lateral inhibition process (Wang, Shamma 1992).

Early Auditory Stage

In the peripheral stage of the auditory system, the acoustic signal is transformed into neural pattern activity through three stages; analysis (basilar membrane response), transduction (hair cell response), and reduction (lateral inhibition) stage. The resulting neural pattern of activity is represented in an auditory spectrogram. Figure 8 illustrates this process. Originally, the analysis stage of the model was constructed by 124 asymmetric constant Q bandpass filters equally spaced over a 5-octave frequency range (Chi et al 1999). Because the goal of the modeling study was to match modulation detection performance to estimates of human peripheral tuning, these filters were replaced with a set of 4th order Gamma tone filters that have been shown to provide a good fit to human auditory filter shapes (Patterson et al 1992).

These Gamma tone filters have an impulse response:

$$g(t) = at^{n-1}\cos(2\pi ft + \varphi)e^{-2\pi bt} \qquad (3)$$

where n represents the order of the filter; b is the bandwidth of the filter; a is the amplitude; f is the center frequency; φ is the phase. Filter bandwidths were based on estimates of the equivalent rectangular bandwidth (ERB_N) for normal hearing auditory filters (Glasberg and Moore 1990) described by

$$ERB_N = 24.7(4.37 \cdot 10^{-3} f + 1) \qquad (4)$$

where *f* is the frequency. Fig. 9 shows the relationship of the raw data with the Glasberg and Moore (1990) equivalent rectangular bandwidths (ERBs) filterbank. Because of this modification, the model better represented the broader relative bandwidths of the filters in the lower frequency regions. The original constant Q-filterbank was unable to account for the poorer performance seen in the 500 and 1000Hz frequency regions in the NH (black and grey color 1) data: the sharp filters in the lower frequency regions produced better cortical representation (higher energy), resulting in better model predicted performance compared to the NH data.

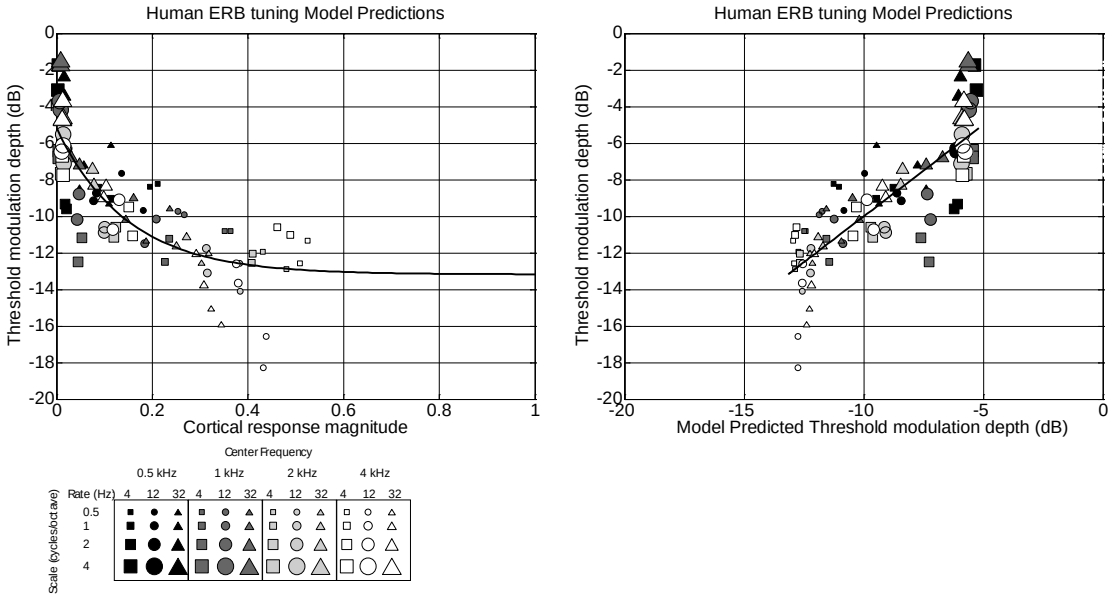

Figure 9: A) The relationship between the psychoacoustic NH STM sensitivity estimates and the corresponding cortical response magnitude of the Gammatone filterbank defined by Glasberg and Moore (1990). Filter ERBs were adjusted based on the notched-noise ERB measurements for the NH listeners. B) The one-to-one relationship between STM data and the predicted STM thresholds based on cortical magnitudes and exponential fit in panel A.

The Gammatone auditory filterbank is defined in such a way that the filter center frequencies are distributed across frequency in proportion to their bandwidth. However, the ERB_N values of the auditory filters are appropriate for sounds presented

at 30-40dB SPL (Glasberg and Moore 1990). To better represent filters for high-level stimuli the bandwidths of the filters at 500, 1000, 2000, and 4000Hz were set based on ERB estimates for NH listeners from the notched-noise data (Table 3). Bandwidth broadening factors were computed at these four frequencies by comparing these ERBs with those determined in equation 4. These factors were linearly interpolated to estimate the ERB factors for the remaining filter center frequencies in the model. The acoustic signal was passed through this modified filterbank producing a complex spatiotemporal pattern of displacements along the basilar membrane of the cochlea described by

$$y(t;s) = h(t;s) * x(t) \qquad (5)$$

where *h(t;s)* represents the impulse response of the cochlear filter at location *s* in the cochlea, *y(t;s)* represents the output of the filter at *s* with input *x(t)* (Wang, 1992). The output of each filter was then passed through a hair cell stage consisting of a high pass filter (fluid cilia coupling); nonlinear compression (ionic channels) and a low pass filter (hair cell membrane). In this stage, the spatiotemporal patterns from the filter outputs were transduced into instantaneous firing rates of the auditory nerve (electrical signal) by

$$y_2(t;s) = g\left(\partial_t y_t(t;s)\right) *_t w(t) \qquad (6)$$

where $\partial_t y_t(t;s)$ is the output of the fluid coupling, *g(.)* is the sigmoidal nonlinearity and *w(t)* is the impulse response of the lowpass filter (Wang 1992). The lateral inhibitory network of the model extracts a spectral estimate of the stimulus from the patterns of auditory nerve responses by rapidly detecting discontinuities along the spatial axis of the auditory nerve patterns and integrating over a few milliseconds (Shamma, 1988). The process involves taking the derivative of the neurons' sound evoked activity with respect to spatial axis of the cochlea. This models the lateral inhibitory influences in the LIN neurons. The half wave rectification of the LIN model represents the threshold non-linearity in the LIN network. The last step of the LIN model involves a long time constant integrator, which accounts for the inability of the central auditory neurons to follow fast temporal modulations (Wang, 1992). Sample outputs (auditory spectrograms) of the peripheral stage of the model in response to STM stimuli are shown in Fig. 1.

Central Auditory Stage

The cortical stage of the model consists of a bank of units that each responds best to a certain combination of rate, scale and frequency. Each unit is tuned to a range of frequencies around the best frequency. In this frequency range, the unit responds best to certain temporal and spectral modulations characterized as spectro-temporal response fields (STRF) (Chi et al, 1999). The central auditory stage analyzes the auditory pattern from the early stage into STM scale-rate plot as shown in Fig. 11. The computation of the scale-rate plots consists of two stages. First, the auditory spectrum is analyzed by the bank of STRFs with varying spectro-temporal

Ω-ω selectivity. The STRFs in the model are tuned to cover a range of best frequencies; best scales (0.25-8 cyc/oct) and best rates (±2 to ±32 Hz). The total output power from the STRFs at each Ω-ω combination is estimated. The ripple spectrogram activates the STRF that matches its outline best (Fig. 10). This is defined as the cortical response of the central stage described in equation (7) where the *STRF()* function is parameterized by its most sensitive spectral and temporal modulations, reflecting the characteristics (i.e. bandwidth) of its excitatory and inhibitory fields (Chi et al 1999) and *y(x,t)* is the auditory spectrogram. Integrating the cortical response described in equation (7) over the whole spectrum yields the scale rate plots shown in Fig. 11B.

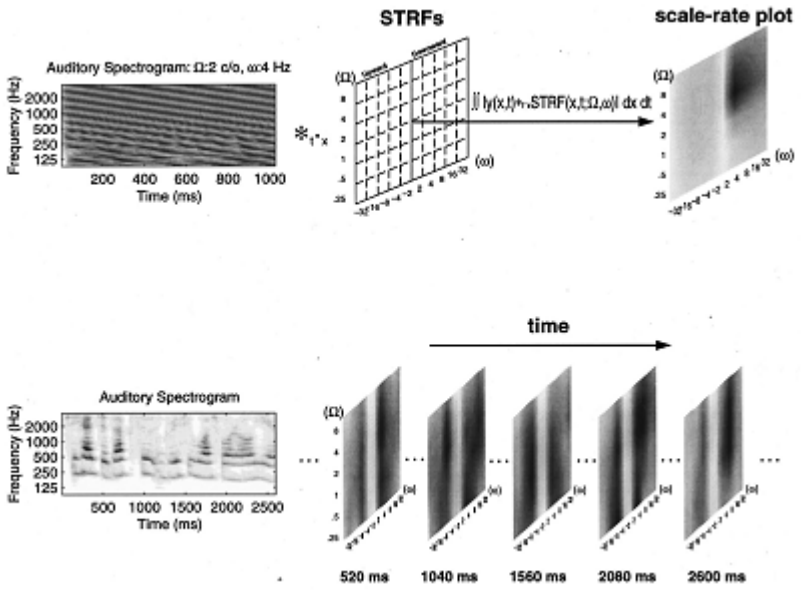

Figure 10: Transformation of auditory spectrogram into plot of STRF in the central stage of the model.

$$r(x,t;\Omega,\omega) = y(x,t) *_{t,x} STRF(x,t;\Omega,\omega) \qquad (7)$$

Fitting Model to Psychoacoustic Data

The cortical response sensitivity of the model for a particular ripple stimulus was characterized by the energy at the appropriate <rate, scale> combination of the scale-rate plot averaged across the appropriate frequency regions: the response of an octave band stimulus was averaged across the frequency channels corresponding to the frequency region of that specific stimulus. Fig. 11 presents the auditory spectrogram and its cortical response plot for a sample <-4Hz, 1 cyc/oct> spectrotemporal combination. As shown in Fig. 11B, the cortical filters tuned near or at <-4Hz, 1cyc/oct> respond best (i.e. most energy) to this stimulus. Fig. 9 plots the cortical response sensitivity plotted against the mean psychoacoustic STM sensitivity data for NH listeners. The model is able to capture the general behaviors of the psychoacoustic data where the cortical response is weaker at higher scales (larger symbols) and lower frequency regions (smaller symbols), corresponding to poorer performance in the data. The relationship between the model response and the NH sensitivity data (Fig.9) was fit with an exponential function with three free parameters (equation 8). The best fitting parameters were a=8.2555, b= -6.8685, and c=13.2558. Although this function best describes the relationship between the model and the NH data (Fig. 9), it was unable to capture listener performance seen at 0.5 cyc/oct conditions at 4000Hz (white small shapes): the NH listeners had high sensitivity to these conditions than the cortical responses that were predicted by the model for the same conditions. In addition, the model did not represent the 4cyc/oct stimuli clearly as seen in Figures 9, 12, 13. The cortical representation of the high scale conditions

hit a floor in the model shown in Fig. 9, suggesting that the bandwidth of the NH filters were too broad to be able to represent the 4 cyc/oct stimuli. Perhaps, because the cortical representations were presented on a linear scale, the small cortical response differences in the 4cyc/oct conditions were unclear. To represent these small differences more clearly, a log representation of the cortical responses should be used in future analyses.

This function describing the relationship between the model output and STM sensitivity was assumed to be fixed across all NH and HI listeners to test the hypothesis that decreased STM sensitivity for HI listeners may be explained by peripheral functions alone.

$$y = ae^{bx} + c \qquad (8)$$

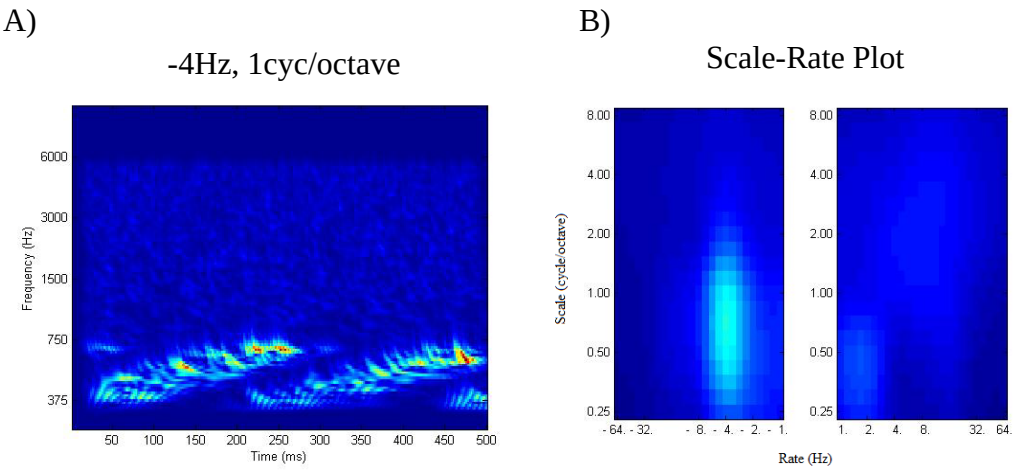

Figure 11: a) Auditory spectrogram of ripples 4Hz, 1cyc/oct, upward direction at CF=500Hz BW=1 octave. b) Scale-rate plot of the ripple at the cortical stage. Note that negative value of the rate in the scale rate plot refers to the upward direction of the ripple in the model.

The results obtained by the ANOVA analysis and averaged STM data over rate and scale suggested that STM sensitivity for HI listeners was mainly affected by impairments in spectral processing. To investigate how much of a reduction in spectral sensitivity would be needed to explain the STM sensitivity results for HI listeners, the effects of reduced frequency selectivity were modeled by broadening the peripheral filters relative to the NH filters. The approach was to produce model predictions for a range of filter broadening factors (ERB factors), then determine which set of factors (one for each frequency region) produced the best fit (lowest mean square error) to the octave-band STM sensitivity data for each individual HI listener. An interpolation was performed based on each set of factors for individual HI listeners to determine the broadening factors for frequency regions that were not parameters in the STM detection experiment. For each HI listener, the periphery filters were broadened based on estimated filter bandwidth factor sets. This resulted in an optimal fit of the model to a HI listener's STM data based on frequency selectivity alone.

Indeed, the broadening of the filters yielded improved model fits for the individual HI data. Fig. 12 shows this improvement for the average HI STM sensitivity data where Fig. 12A describes the relationship between model predictions and the data using the NH filter bandwidths and Fig. 12B shows the model fit after the periphery filters had been broadened to fit the data. The adjustments of the filters for the average HI STM sensitivity reduced the RMS error about 14.5%.

Individual model fits to each HI subject's STM sensitivity data exhibited more of a significant improvement than the overall model fit to the average HI measurements. Fig. 13 shows the fitting for HI listener 15. The filter adjustments reduced the RMS error by 22.6% for this particular listener's STM sensitivity. While these estimated ERB factors shown in Table 2 improved the model fits for the individual HI data, frequency selectivity was unable to capture all the characteristics of the psychoacoustic data. In particular, the model was unable to predict STM sensitivity for most of the 4 cyc/oct conditions (Fig. 9,12,13). Again, plotting the cortical output on a logarithmic scale (Fig.9A) might have yielded a better fit, especially to the 4 cyc/oct.

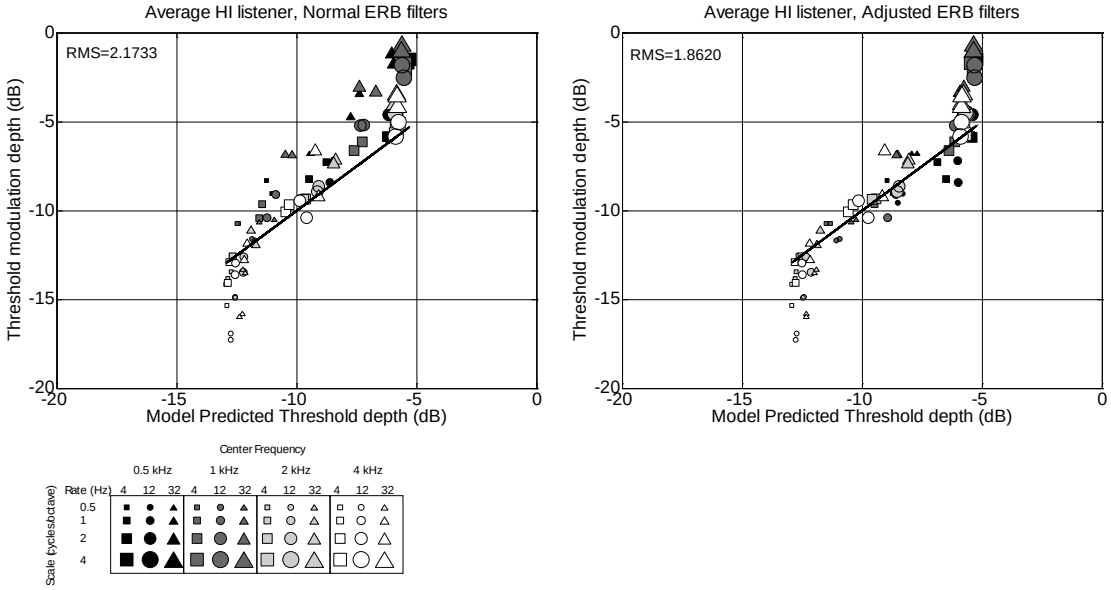

Figure 12: Comparison of average raw data with model for the HI group. (Left panel): Comparison of the STM sensitivity data with predicted thresholds based on the NH model peripheral filters. (Right panel): Comparison of data and model predictions with the bandwidths of the peripheral filters adjusted (i.e. broadened) to fit the data.

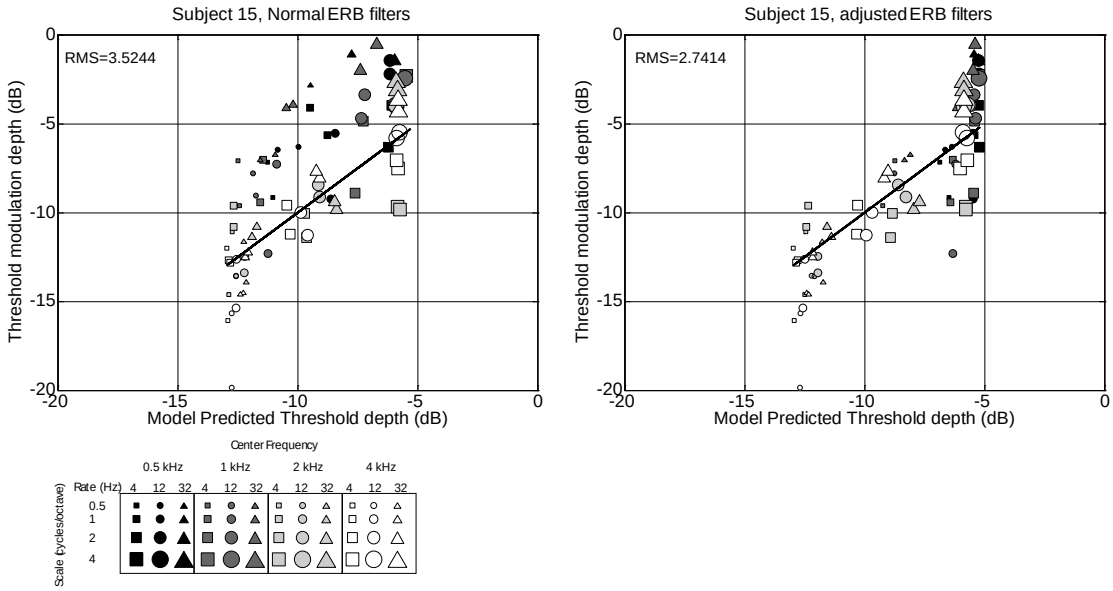

Figure 13: Comparison of raw data with model for HI subject 15. (Left panel): Comparison of the STM sensitivity data with predicted thresholds based on the NH model peripheral filters. (Right panel): Comparison of data and model predictions with the bandwidths of the peripheral filters adjusted (i.e. broadened) to fit the data.

Subject	ERB factor @ CF= 500Hz	ERB factor @ CF= 1000Hz	ERB factor @ CF=2000Hz	ERB factor @CF=4000Hz
S13	1	1	0.5	0.75
HI07	0.5	0.75	0.5	1
HI49	0.75	1	1	1.2
HI17	2.8	2	1	1.4
S17	1.4	1.8	1	1
S18	1.4	1.2	0.75	1
S14	1.2	1.2	1.2	1.4
S6	1.4	1.4	1.4	1.2
S16	2.4	1.4	1	1.2
S15	2.6	2.2	0.75	1
S19	3	3	1.2	2.8
S10	1	1.2	1	1.2

Table 2: Model Predicted ERB factors for each HI subject at each frequency region.

Chapter 5: Relationships to other psychoacoustic measures and speech intelligibility

The STM sensitivity data was compared to data collected for the same HI listeners in other laboratories at Walter Reed Army Medical Center by Van Summers, Matthew Makashay and Sandeep Phatak and at the Portland-VA site by Marjorie Leek, Sarah Melamed and Frederick Gallun.

STM Data

The HI STM data were processed to reduce the number of variables in the correlation analysis. The general approach was to estimate the scale or rate required to yield an STM detection threshold at a fixed value of -6 dB. Although there was a general trend for performance to decrease with increasing modulation rate, STM sensitivity was often non-monotonic as a function of rate for individual listeners. STM sensitivity was more reliably monotonic as a function of scale. Therefore, STM sensitivity was characterized for each combination of frequency, rate and direction by fitting a line to the STM vs. log-scale data and estimating the scale needed to yield an STM detection threshold of -6 dB. STM sensitivity for each frequency and listener was characterized as the mean of the six (three rates x two directions) log-scale estimates.

Speech intelligibility data

The speech intelligibility data were also processed to reduce the number of variables in the correlation analysis. Intelligibility was measured for IEEE (1969) sentences presented in stationary and modulated noise at four SNRs (-6, -3, 0 and +3 dB). The target speech level was presented at a high level of 92 dB SPL in an attempt to overcome audibility limitations. Comparisons with various SNRs and modulated and stationary noise showed qualitatively similar results, although correlations were strongest for -3 and 0 dB SNR where there were no floor or ceiling effects. Figure 14 shows the comparison between the STM data at various frequencies and the speech intelligibility scores for stationary noise with SNR=0dB. The comparison reveals significant correlations between STM sensitivity and speech for all absolute frequency regions with the exception of 4000Hz (Fig.14). Highest correlation was found for STM sensitivity in the 500Hz frequency region (r^2=0.30). These correlations confirm the hypothesis that speech intelligibility performance in background noise is dependent on STM sensitivity.

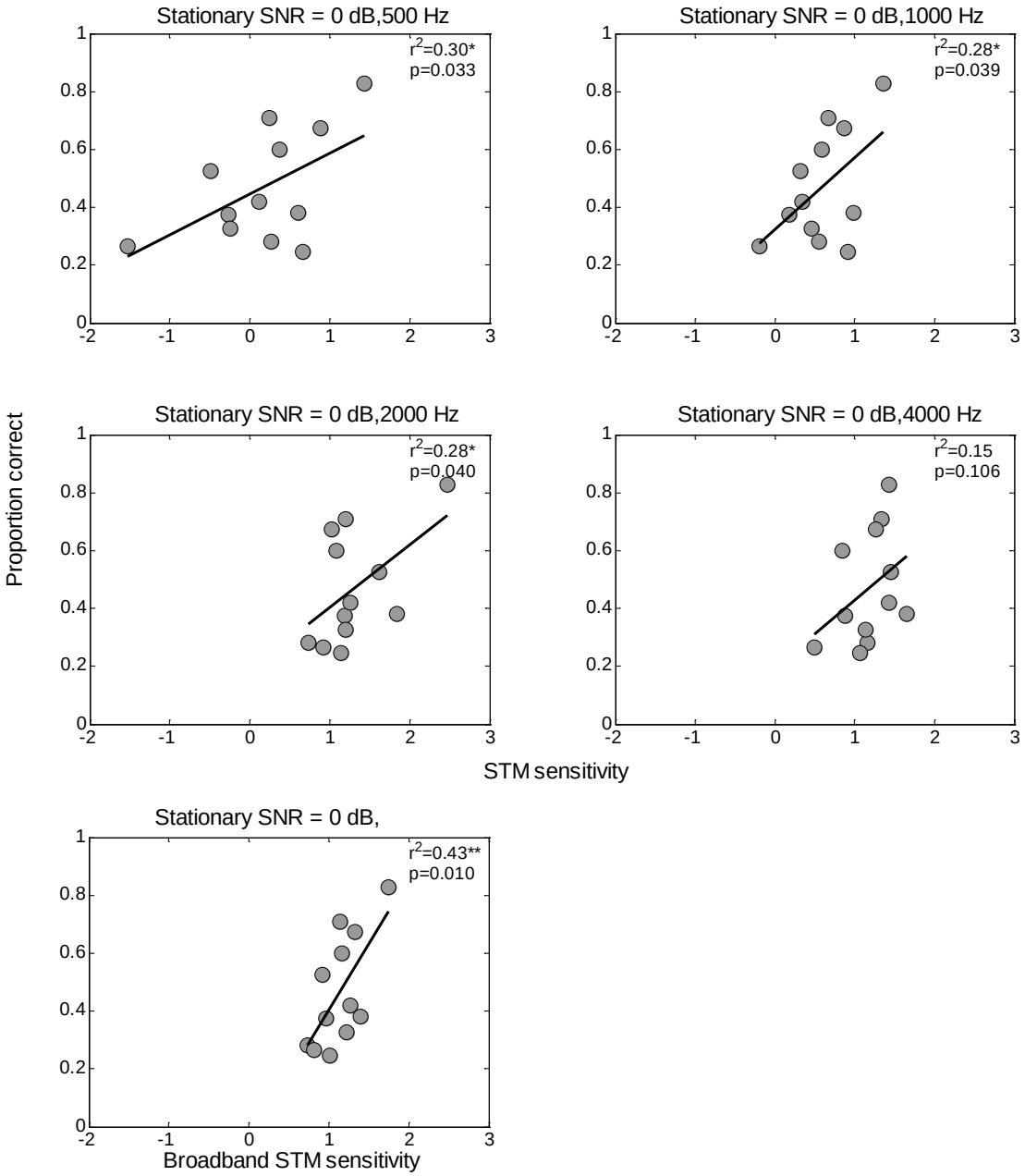

Figure 14: Comparison of Speech Intelligibility scores and STM sensitivity across absolute frequency. Speech was presented in stationary noise with a SNR of 0dB. The p values listed in each panel are one-tailed p values. It was assumed a priori that the correlations can only go one way - listeners who are worse at one task will also be worse at the other. Last plot compares broadband STM sensitivity to Speech intelligibility scores.

Frequency selectivity data

Masked thresholds were determined for sinusoidal signal tones of 500, 1000, and 4000Hz in the presence of notched-noise maskers with variable notch widths. The notches were placed both symmetrically and asymmetrically about the signal frequency and the noise level was varied to determine the level needed just to mask the pure-tone. Frequency selectivity was estimated by fitting a roex filter to the data (Baker and Rosen, 1994) and calculating the filter's equivalent rectangular bandwidth (ERB). For most listeners and frequencies the tone level was set at 70 dB SPL, but in some cases where listeners yielded erratic results, a higher tone level was selected. Table 3 shows the notched-noise ERB estimates for both the NH and HI listeners tested at Walter Reed.

Subject #	ERB@ 500Hz	ERB@1000Hz	ERB@2000Hz	ERB@4000Hz
251	155.1	292	530.7	827.7
217	153	198.9	539.5	906.9
218	153.4	271.5	497	1196.5
254	107.1	193.4	504.8	987.9
17	146.5	251.2	1114.5	1179.1
13	169	332.6	1241.8	2520.4
18	97.7	204.3	688.4	2471.4
16	178.2	637	1048.8	2112.1
15	110.8	270.8	997.6	1820.7
10	90.9	306.8	611.3	870
19	89.3	281.4	1100.8	1168
6	157.8	300.7	1053.1	5286.5

Table 3: Notch noise ERB estimates for NH and HI listeners @ 70dB SPL.

The bandwidth widening factors that yielded the best fit of the cortical model prediction to the STM sensitivity data are shown for each HI listener that had

completed testing at the time the model was fit in Table 2. These factors estimated by the STM model indicate the reductions in frequency selectivity that would be required to explain the STM sensitivity data if the reduced performance for HI listeners could be described simply in terms of filter broadening that accompanies hearing loss. If the reduced STM sensitivity for HI listeners were caused by reduced frequency selectivity, then these best fitting bandwidths should correlate with auditory filter bandwidths estimated using the notch-noise method. However, the comparison of notch-noise ERB estimates (Table 3) with the model predicted ERBs for the HI listeners (Table 2) yielded no significant correlation. Fig. 15 exhibits the lack of correlation between the notched-noise and model predicted ERB factors for each frequency region. In the higher frequency regions (2000 and 4000Hz) the comparison results in a horizontal function (Fig.15) indicating impairments in notched-noise estimates but not STM. This is in contrast with the comparisons at 500Hz and 1000Hz where the data is vertical, suggesting impairment in the STM detection and not the notched-noise estimates. The large almost-significant correlation seen at center frequency of 1000Hz is in fact negative and thus, does not support the hypothesis. Similarly, the comparison of the mean HI notched-noise ERBs and the model fit to the mean HI STM sensitivity data (model predicted ERBs) for each frequency region reiterated the lack of correlation (Fig.16). The comparison shows a complete disassociation between the frequency regions where HI listeners are having difficulty with STM detection (500 and 1000Hz) and the frequency regions where they show broader filters (2000 and 4000Hz). In addition, no significant correlations were found across the HI listeners between the notched-noise derived ERBs and STM

sensitivity at any absolute frequency, corroborating the comparison that are plotted in Figs. 15 and 16. However, the data for this comparison was not included. The lack of correlation suggests that reduced STM performance for HI listeners is not explained by a reduction in frequency selectivity.

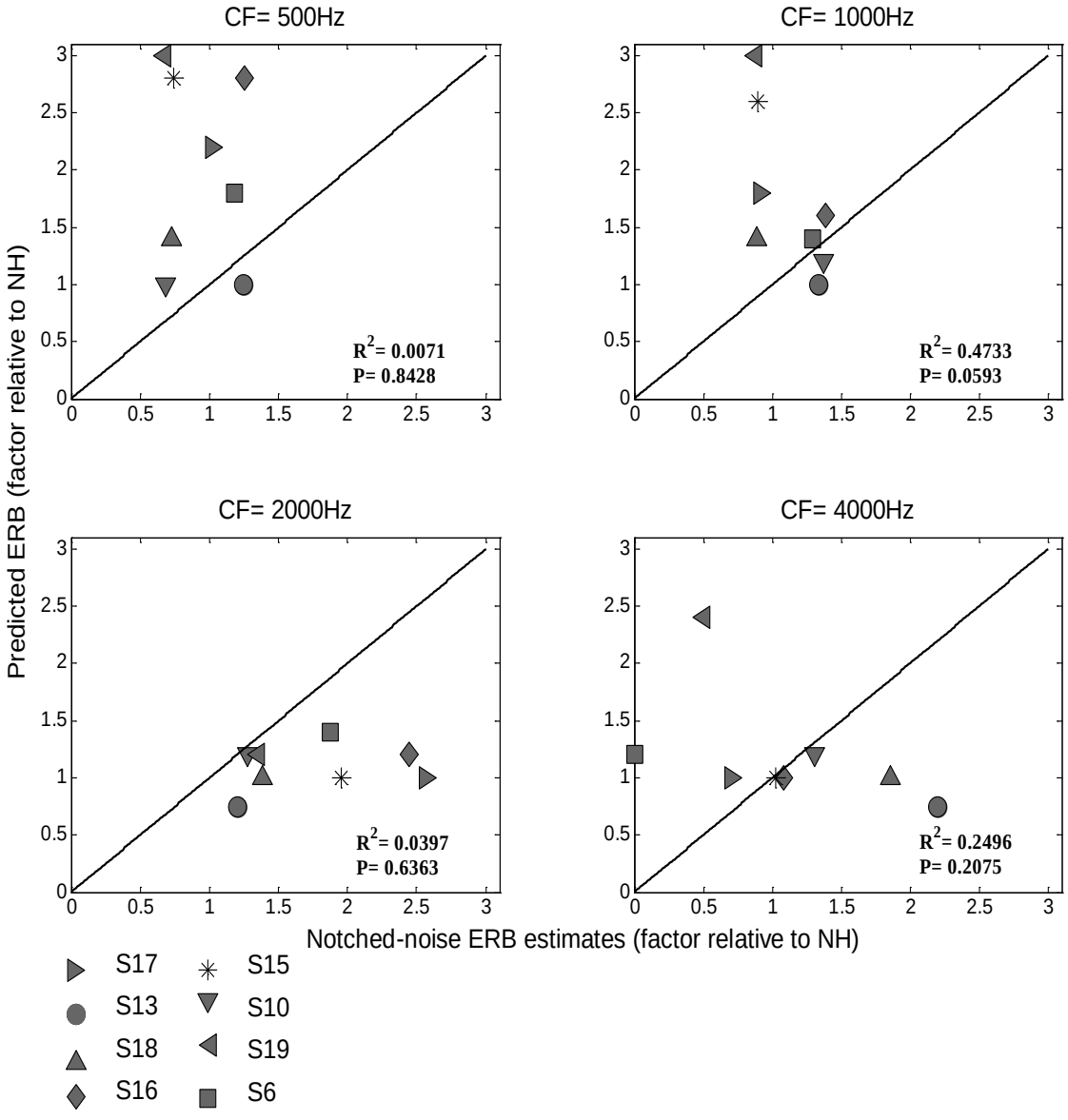

Figure 15: Comparison of model predicted ERB estimate to notched-noise ERB estimated for each HI listeners at each frequency region.

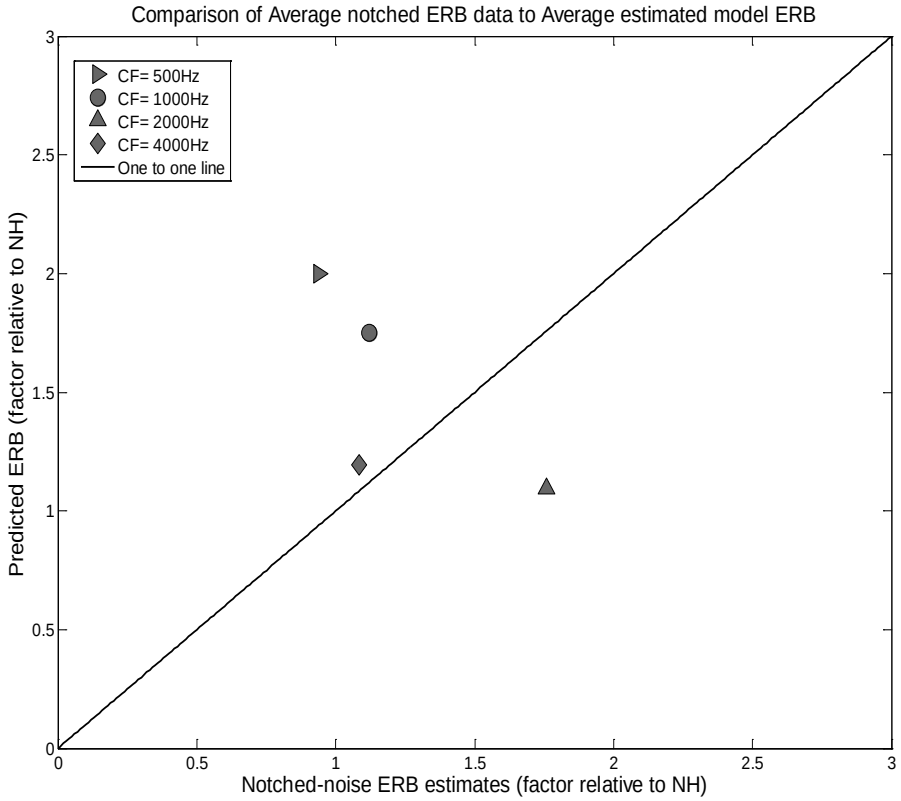

Figure 16: Comparison of model predicted ERB estimate to notched-noise ERB estimated for average HI listener.

Frequency Modulation detection data

Estimates of frequency modulation sensitivity are thought to reflect the ability to use temporal fine structure (TFS) information (Moore and Sek, 1996). Listeners were tested in their ability to detect frequency modulation applied to a tone carrier (500, 1000, 2000 or 4000 Hz) presented at 85 dB SPL. Random amplitude modulation was added to both the reference and signal intervals to reduce the usefulness of induced amplitude-modulation cues. FM detection sensitivity measured by Van Summers and Matt Makashay was reported as the logarithm of the minimum modulation depth (in Hz) required for detection. Comparison of the STM and FM

detection sensitivity demonstrated significant correlations in the low frequency regions of 500 and 1000Hz (Fig.17). Interestingly, the comparison of sensitivity of broadband STM and FM also showed a strong correlation when the FM carrier frequency were low (500 and 1000Hz) (Fig.18). The implication of the relationship found between STM sensitivity at low absolute frequencies and FM detection with low carrier frequencies is that STM and FM sensitivity at these low absolute frequencies are based on similar underlying mechanisms, possibly temporal fine structure (TFS) processing

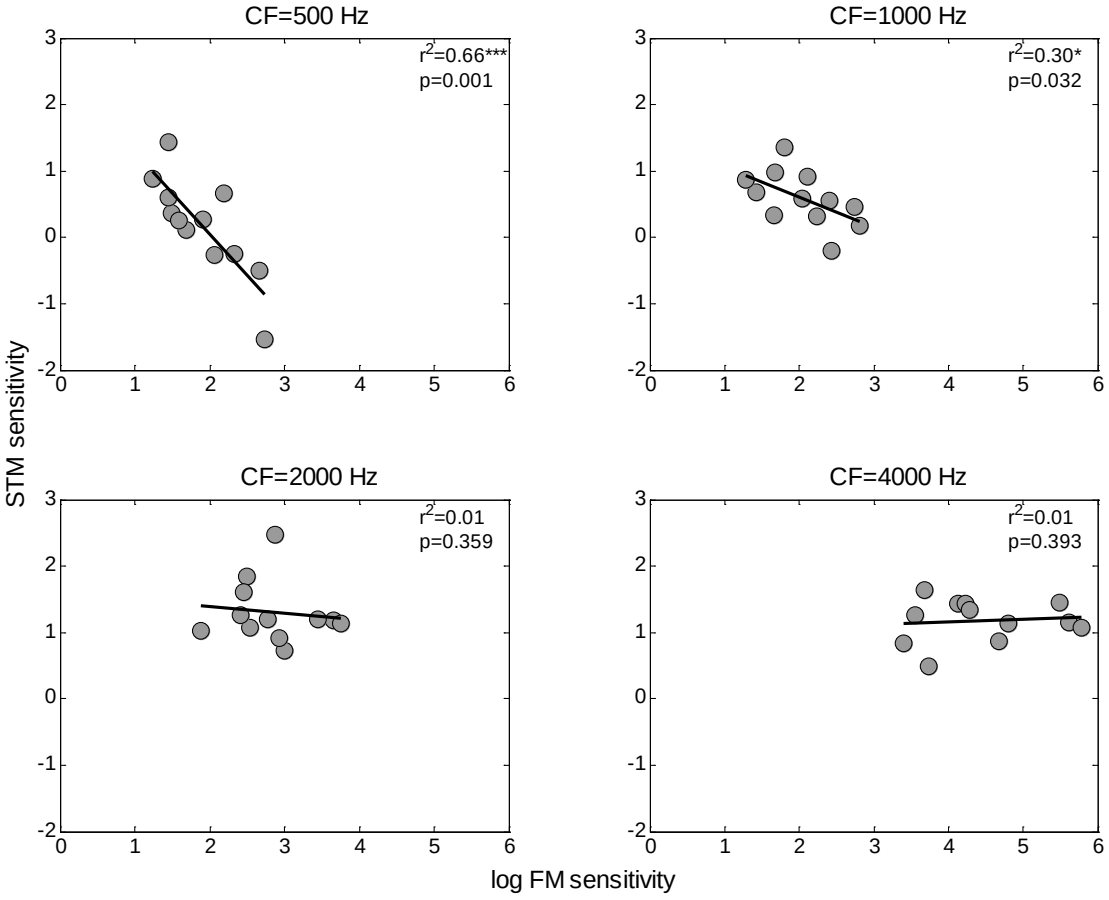

Figure 17: A comparison between STM sensitivity and FM detection. Each plot compares the STM data for that absolute frequency region with the FM data that uses the corresponding carrier frequency.

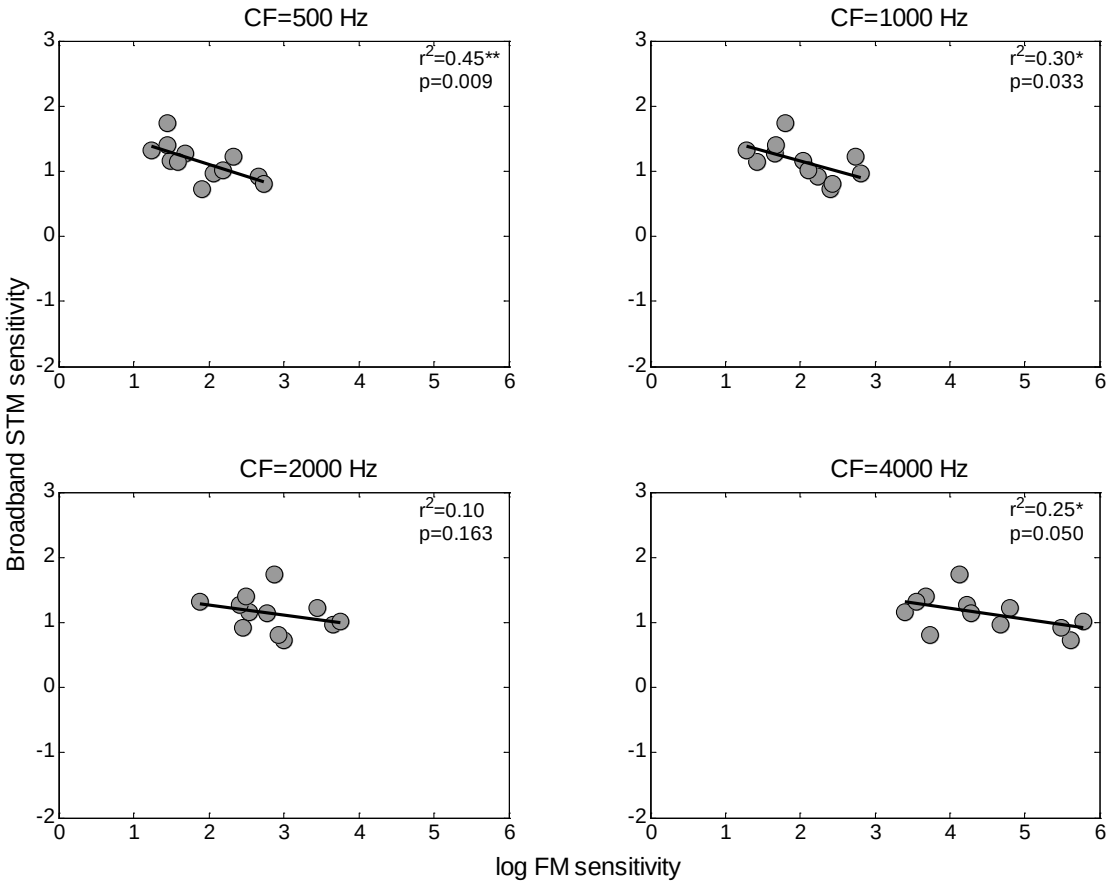

Figure 18: A comparison between broadband STM sensitivity and FM detection. Each plot corresponds to a different FM carrier frequency.

Chapter 6: Discussion

General Trends

The general form of the STM modulation transfer measured in NH and HI listeners in this study agrees with STMTFs measured in previous studies (Chi et al 1999). The deterioration of sensitivity with increasing temporal modulation rate or spectral modulation seen in both hearing groups exhibits the typical limitation of the spectral and temporal resolutions of the auditory system as seen in studies measuring purely spectral (Green, 1986; Hillier, 1991; Eddins & Bero 2006) and temporal modulation sensitivity (Vemiester 1979; Yost ad Moore, 1987; Van Zanten and Senten, 1983).

There were, however, some differences in the details: the minimum detection threshold occurred for 0.5-1cyc/oct stimuli (Fig.3) in agreement with results from Chi et al. (1999). However, this differs from the focus of best sensitivity in the 2 -4cyc/oct range reported in other studies (Eddins & Bero, 2006; Summers and Leek, 1994) that have looked at the detection of spectral modulation alone for octave-band stimuli. Furthermore, the comparison of the STM sensitivity in this study to the spectral transfer functions measured in Eddins and Bero (2006) showed performance deterioration about 2 times faster from peak performance with increasing spectral modulation scale. A possible explanation for this fast drop of sensitivity in this study is when measuring spectral modulation sensitivity alone, the multiple spectral peaks fall within the same peripheral filter at higher scales resulting in induced temporal

modulation that could be used to detect these high spectrally modulated signals. However, when there is also temporal modulation in the signal, as in the STM task, temporal cues are also available for low spectral modulation scales, improving performance relative to the spectral-modulation alone conditions. At high scales, the additional temporal modulation might not benefit performance as much because the induced temporal modulation is already available in the STM stimuli. Because of the possible presence of temporal cues in pure spectral modulation detection, true spectral modulation sensitivity may not be measured in these studies. To prevent the use of the arising temporal cue, random temporal modulation could be added to both the spectrally modulated signal and comparison signal in the identification task.

The frequency effect in STM data showed that performance in both HI and NH listeners was better at higher frequency regions than at lower frequency regions. The absolute frequency effects may be attributable to the tuning characteristics of the peripheral auditory system, with higher-frequency filters more sharply tuned (relative to CF) than lower frequency filters. Interestingly, the comparison of STM across frequencies further showed that wideband STM detection performance was roughly equal to performance for the best narrowband condition at 4000Hz. It is therefore important to characterize STM sensitivity across a range of carrier frequencies rather than just broadband alone like in previous studies (Chi et al 1999; Henry et al 2005; Won et al 2007). Supin et al (1997) identified a similar absolute frequency effect for spectral ripple resolution using a phase-reversal test. They reported a gradual increase of spectral resolution from 500 to 8000Hz. However, Eddins and Bero (2006) reported that spectral modulation transfer functions (SMTF) are not strongly

dependent on carrier frequency region ranging from 200 to 12,800 Hz, with the exception of carrier bands restricted to very low audio frequencies (e.g., 200–400Hz). They explained that the significant poorer performance at 200-400Hz octave band condition might be due to the possibility that the perception of global spectral shape is not as good at low audio frequencies as it is at higher audio frequencies. However, the insignificant dependence of spectral modulation detection on carrier frequency in Eddins and Bero's study (2006) may have arose as a result of other possible cues (i.e. peripherally-induced temporal fluctuations) that could have improved the spectral modulation detection. These cues might have been especially salient at lower frequency regions. Because low frequency auditory filters are broader on a logarithmic scale, more spectral peaks would have fitted within an auditory filter for a given spectral modulation scale. Furthermore, these peaks would be more closely spaced on an absolute frequency scale, yielding lower-rate induced modulation that would be easily detectable.

Alternatively, the absolute frequency effects may also be attributed to the extent to which temporal modulation sensitivity varies with absolute frequency. Eddins (1993; 1999) investigated temporal modulation sensitivity as a function of bandwidth and frequency region and reported that temporal modulation sensitivity was not greatly affected by absolute frequency region when comparisons were made between stimuli with the same absolute bandwidth (in Hz). The comparison of the estimates of temporal modulation sensitivity of Eddins (1999) for conditions with the same *relative* bandwidth of one octave showed very little difference as a function of absolute frequency. The comparison showed an improvement of sensitivity of only

3dB max from 800-6400Hz at a frequency modulation of 16Hz with an average improvement of 1dB/octave. In contrast, the comparison of the data with the same relative bandwidth of 0.415 octaves across absolute frequency regions exhibited a more significant improvement of temporal acuity from 800-12800Hz (max of 11dB at fm=32Hz). The increase of temporal modulation sensitivity with frequency may be due to the possible interference of inherent fluctuations created by the filters in the lower frequency regions. The absolute filter bandwidths at the lower frequency regions are narrower than the bandwidths of higher frequency filters and as a result, lower rate inherent fluctuations are created. Thus, stimuli in the lower frequency regions are difficult to distinguish because of the interfering fluctuations in the noise carrier, limiting modulation detection.

Based on frequency selectivity, the model was generally able to capture the deterioration of STM sensitivity with decreasing absolute frequency once the peripheral filter bandwidths were set to reflect the non constant-Q nature of human frequency tuning (Glasberg and Moore, 1990), supporting the implications of the frequency effect in the STM data. Furthermore, although the model takes into account the temporal and spectral effects seen in the STM data when estimating the ERBs, it failed to capture the effect of the noise carrier on STM sensitivity which may have resulted in the disconnect between the notched- noise and model ERB estimates. When the noise carrier was filtered, modulations near the frequency of the filter bandwidth arose. These inherent modulations can mask the actual STM at lower frequencies more than higher frequencies as the modulations are slower in the lower frequencies due to smaller absolute bandwidth. This effect was not accounted for

when estimating the filter ERBS because the noise carrier was not passed through the model. However, the effect of the noise carrier cannot be so profound to cause this lack of correlation.

Effects of Hearing loss

The current study also found that hearing loss has little effect on STM sensitivity: the HI listeners were only slightly impaired in the ability to process the complex STM signals when compared to NH listeners. Furthermore, there was dissociation between the frequency regions where HI listeners were more impaired in STM detection (low frequencies) and where bandwidths estimates using notched-noise were broader (high frequencies). Several possible contributing factors might help to explain the slight impairment seen in the HI listeners relative to the normals. One possibility is that the high stimulus level of 80dB/octave could have reduced the performance of the NH group. Many studies have documented the broadening of the auditory filters with increasing stimulus level (Weber, 1977; Pick, 1980; Lutfi and Patterson, 1984; Glasberg and Moore, 1990; Rosen and Stock, 1992; Rosen and Baker, 1994; Moore and Glasberg, 1986; Rosen *et al.*, 1998; Hicks and Bacon, 1999; Glasberg and Moore, 2000). Thus, because of this behavior, the broader filters at the high stimulus level may have yielded a poorer spectral resolution resulting in reduced performance in the NH listeners. However, there was no relationship between STM sensitivity and auditory filter bandwidth across frequency even when performance was compared between groups with both tasks at similar high levels (Fig. 15). Nevertheless, where HI listeners showed impairment, performance was affected in the

spectral domain more than the temporal domain suggesting that STM sensitivity deficits reflect differences in spectral resolution. This is consistent with previous results that have shown impaired spectral modulation sensitivity (e.g., Henry et al., 2005) but not impaired temporal modulation sensitivity (e.g., Moore and Glasberg, 2001). However, there was little relationship between the resulting estimates of frequency selectivity based on the STM sensitivity data and ERB data derived using the notched-noised method in the same listeners. Notched-noise ERB estimates indicate wider filters in the higher frequency regions corresponding to the high-frequency loss seen in the audiograms of the HI listeners. This contrasts the STM data where HI listeners show impairment at the low frequencies and as such, have broader filters at the lower frequencies. HI listeners may compensate for reduced peripheral selectivity at more central processing stages. Alternatively, STM sensitivity may not have been truly measured at the low frequencies. Overall, this disconnect between the ERB estimates needs further analysis.

Another possible explanation is that the poorer frequency selectivity seen in the HI listeners through notched-noise ERB measurements at the higher frequencies was perhaps offset by better temporal resolution in STM detection. Previous studies have shown that HI listeners can perform quasi-frequency modulation (QFM) detection out to higher modulation frequencies than NH listeners (Nelson and Schroder, 1995; Bernstein and Oxenham, 2006). This can be explained by the wider peripheral filters associated with hearing loss. In the normal periphery, sharp filter characteristics attenuate the sidebands associated with the introduction of temporal modulation, thereby reducing modulation depth (i.e. peaks and values of modulation

become less distinguishable) when the modulation frequency exceeds the bandwidth of the filter. With the wider auditory filter bandwidths associated with hearing impairment, the attenuation of the sidebands would be reduced. Thus, wider peripheral filters for HI listeners might increase the peripheral interaction between components, thereby yielding stronger temporal modulations and compensating for the poor frequency selectivity.

Although these explanations describe possible reasons for the only slightly impaired STM sensitivity exhibited by HI listeners, they do not account for the observation that STM sensitivity was more impaired at low than at high frequencies in contrast to the filter bandwidth data that showed the opposite trend. The STM detection task in this study is similar to an FM detection task where a frequency of a single tone is being varied continuously over time, except that in this study, the frequency of the spectral peaks of the STM signal are being varied over time (Fig.1). Based on this notion that STM detection is similar to pure-tone FM detection, this interaction between the effects of hearing loss and absolute frequency might be explained by the possible influence of TFS. Moore and Sek (1996) suggested that FM detection depends on both TFS and changes in the excitation pattern [induced amplitude modulation (AM) cues]. They investigated the dependence of FM detection on TFS and AM. They showed that when AM was added to both intervals of a forced-choice trial to disrupt FM detection based on excitation pattern cues, detection ability worsened for high carrier frequencies but not for low carrier frequencies, especially for low modulation rates. They suggested that at low modulation rates and carrier frequencies, FM detection depends on TFS (phase locking) whereas FM

detection at higher modulation rates and carrier frequencies depends on AM cues where TFS processing is not reliable (Moore and Sek 1996; Moore and Skrodzka, 2002). This effect was found in this study data when the data was collapsed over scale (Fig.7 left panel) to show the effects of temporal modulation on STM sensitivity. This representation showed a non-significant trend where relative to normal, HI listeners had slightly more impaired STM sensitivity at lower (4 Hz) than higher (32Hz) temporal rates in lower frequency regions of 500 and 1000Hz. The STM data might be explained in terms of TFS processing deficits: similar to FM detection tasks, at higher frequency regions of 2000 and 4000Hz, both NH and HI listeners might have had a reduced ability to use TFS information due to phase locking limitations at higher frequencies and instead used AM cues, which may not be affected by hearing loss. In the 500 and 1000Hz frequency regions where distinct differences between NH and HI group were found, NH listeners might have used TFS information to detect the STM signals at the different temporal rates; however, because of possible damage to TFS processing, HI listeners might have used the TFS information poorly or not at all.

To the extent that speech intelligibility is mediated by STM sensitivity these results corroborate the finding of Buss et al (2004) showing that differences in TFS sensitivity but not frequency selectivity can explain differences in speech intelligibility. Indeed, the strong correlation found between STM sensitivity at low absolute frequency regions and FM detection with low carrier frequency is consistent with the idea that TFS processing is important for STM detection: HI impaired listeners possibly have impaired TFS processing and as such, are unable to detect

STM signals at low absolute frequencies despite their normal frequency selectivity in these regions. In addition to the relationship between STM and FM detection, the significant correlation found between STM sensitivity and speech intelligibility supports the notion that differences in TFS sensitivity might explain differences in speech intelligibility. Lorenzi et al (2006) showed that speech performance of HI listeners was affected only when the TFS of the speech was presented suggesting that HI listeners had TFS processing deficiencies affecting their speech intelligibility performance.

Chapter 7: Future Work

The current findings of STM detection in NH listeners corroborate the previous STM trends seen in Chi et al (1999) showing low pass behavior for both temporal and spectral modulation. The current study extends this finding by establishing relationships between STM sensitivity and hearing impairment and absolute frequency, respectively. In this study, it was hypothesized that impaired frequency selectivity would cause poor STM sensitivity in HI listeners as it has been shown in previous studies that frequency selectivity (Chi et al 1999) but not temporal resolution that is significantly impaired with sensorineural hearing loss (Bacon and Viemeister, 1985; Bacon Gleitman, 1992; Moore et al, 1992d). As a result, the possibility of TFS affecting the STM sensitivity was not explored in the model. The adjustments of the peripheral filters in the model only accounted for a small percentage of the STM sensitivity behavior suggesting that other factors such as TFS or temporal resolution deficits may have contributed to the STM sensitivity. As such, it is necessary to measure purely temporal and purely spectral modulation across frequencies to see if temporal resolution is possibly impaired at particular frequencies regions in HI listeners. In addition, the measurements could show more directly whether impairment in temporal or spectral modulation is causing the reduced performance of the STM.

The similar performance between NH and HI listeners in STM detection task could be due, at least in part, to an effect of level that equalized the performance of the NH and HI listeners. To verify this explanation, STM sensitivity in NH listeners

needs to be measured at various signal levels to determine if presentation level has an impact on STM sensitivity. This measurement could additionally tell us, in a more controlled way, about the relationship between frequency selectivity and STM processing, since frequency selectivity is known to vary systematically with level in NH listeners. Additionally, the possible effect of level on STM detection in NH listeners can potentially be accounted for in the modeling of the data by incorporating a more realistic nonlinear peripheral model such as the dual-resonance nonlinear (DRNL) model (Lopez and Meddis, 2001).

Chapter 8: Conclusion

STM sensitivity in the spectral and temporal dimensions adminstered in this study demonstrated the lowpass characteristics shown in previous studies (Chi et al 1999). The STM detection performance showed a significant dependence on absolute frequency. It is therefore critically important to estimate STM sensitivity across the audible range because estimates of broadband STM sensitivity alone do not adequately characterize sensitivity performance at low frequencies.

HI listeners were only mildly impaired in their ability to detect STM, and unexpectedly, mainly at low frequencies where they are audiometrically more similar to NH listeners. The mild impairment in HI STM performance could have been because of impaired temporal fine structure processing that affected performance mainly at lower frequencies. Furthermore, HI listeners may have been able to compensate for impaired frequency selectivity in the periphery by better temporal resolution. Estimates of frequency selectivity in HI listeners derived from the STM data did not correspond to more traditional notched-noise based estimates. The STM results might instead be explained by impaired TFS processing in HI listeners at low absolute frequencies. Significant correlations between STM and FM sensitivity and speech intelligibility suggest that impaired TFS processing may affect STM sensitivity and speech intelligibility.

Glossary

1. HI- Hearing Impaired (Impairment)
2. NH- Normal Hearing
3. STM- Spectrotemporal Modulation
4. TFS- Temporal Fine Structure
5. TMTF- Temporal Modulation Transfer Function
6. AI- Auditory Index
7. SII- Speech Intelligibility Index
8. STMTF- Spectrotemporal Modulation Transfer Function
9. ANOVA- Analysis of Variance
10. SVD- Singular Value Decomposition
11. LIN-Lateral Inhibition Network
12. STRF-Spectro-Temporal Response Field
13. SNR- Signal to Noise Ratio
14. FM- Frequency Modulation
15. ERB- Equivalent Rectangular Bandwidth
16. QFM-Quasi-Frequency Modulation
17. AM- Amplitude Modulation

Bibliography

1. Amagai, S., Dooling, R., Shamma, S., Kidd, T., and Lohr, B. (1999). ‘‘Perception of rippled spectra in the parakeet and zebra finches,’’ *J. Acoust. Soc. Am.*

2. ANSI S3.5-1997, American National Standards Institute, New York

3. Bacon, S.P. and Viemeister, N.F. (1985). “Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners”. *Audiology*, 24, 117-134.

4. Bacon S. P., and Gleitman R. M. (1992.). “Modulation detection in subjects with relatively flat hearing losses,” *J. Speech Hear. Res*. 35, 642–653

5. Baer, T, and Moore, C. J. (1993). “Effects of spectral smearing on the intelligibility of sentences in noise.” *J. Acoust. Soc. Am*. 94, 1229

6. Baer, T, and Moore, C. J. (1994). “Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech.” *J. Acoust. Soc. Am*. 95, 2277

7. Bernstein, L.R. and Green,D. M, “Detection of simple and complex changes of spectral shape”, *J. Acoust. Soc. Am*. 82

8. Bernstein L, . R., and Green, D. M. (1987). "The profile analysis bandwidth," *J. Acoust. Soc. Am*. 81, 1888-1895. (1987)

9. Bernstein , J. G. W. and Oxenham, A. J. (2006), “The relationship between frequency selectivity and pitch discrimination: Sensorineural hearing loss”, *J. Acoust. Soc. Am*. 120, 3929

10. Buss E., Hall J. W., and Grose J.H (2004), “Spectral integration of synchronous and asynchronous cues to consonant identification”, *J. Acoust. Soc. Am*. 115, 2278

11. Drullman R, Festen JM, and Plomp R. (1994a). “Effect of temporal envelope smearing on speech reception”, *J Acoust Soc Am* , 95: 1053–1064.

12. Drullman R, Festen JM, and Plomp R (1994b). “Effect of reducing slow temporal modulations on speech reception.”, *J Acoust Soc Am* , 95: 2670–2680.

13. Eddins, D. A. (1993), “Amplitude modulation detection of narrow-band noise: Effects of absolute bandwidth and frequency region”, *J. Acoust. Soc. Am.* 93, 470

14. Eddins, D. A. (1999), “Amplitude-modulation detection at low- and high-audio frequencies”, *J. Acoust. Soc. Am*. 105, 829.

15. Eddins, D. A and Bero, E. M (2006). “Spectral modulation detection as a function of modulation frequency, carrier bandwidth, and carrier frequency region.” *J. Acoust. Soc. Am*. 121 363

16. Elhilali, M., Taishih, C., and Shamma, S. A. (2003). “A spectro-temporal modulation index(STMI) for assessment of speech intelligibility.” *Speech Communication*, 41, 331-348.

17. Glasberg, B. and Moore, B.C. J. (1986). “Auditory filter shapes in subjects w ith unilateral and bilateral cochlear impairments”, *J . Acoust.Soc. Am*. 79, 1020-1033.

18. Glasberg, B., and Moore, B.C. J. (1990). "Derivation of auditory filter shapes from notched-noise data," *Hear. Res*. 47, 103-138.

19. Glasberg, B. R. and Moore, B. C. J. (2000), “Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise”, *J. Acoust. Soc. Am*. 108, 2318

20. Green, D. M and Forrest T.G (1986), “Profile analysis and background noise”, *J. Acoust. Soc. Am*. 80, 416

21. Haykin, S. (1996). “Adaptive Filter Theory”, Prentice–Hall, Englewood Cliffs, NJ.

22. Henry, B. A, Turner, C. W., and Behrens, A (2005) “Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners” J. Acoust. Soc. Am. 118, 1111 Hicks, M.L and Bacon , S.P. (1999), “Psychophysical measures of auditory nonlinearities as a function of frequency in individuals with normal hearing”, *J. Acoust. Soc. Am*. 105, 326

23. Hillier, D. A. (1991). “Auditory processing of sinusoidal spectral envelopes” (unpublished doctoral dissertation) Washington University, St.Louis).

24. Houtgast, T. and Steeneken, H.J.M. (1985). “A review of the MTF-concept in room acoustics”, *J.Acoust. Soc. Am*. 77, 1069 – 1077.

25. Keurs, M., Feston, J. M. and Plomp, R. (1992). “Effect of spectral envelope smearing on speech reception. I.” *J. Acoust. Soc. Am*. 91, 2872.

26. Keurs, M., Feston, J. M. and Plomp, R. (1993), “Effect of spectral envelope smearing on speech reception. II.” *J. Acoust. Soc. Am*. 93, 1547

27. Levitt, H. (1971). “Transformed up-down methods in psychoacoustics.” *J. Acoust. Soc. Am*, 49(2):467-477.

28. Lopez-Poveda, E. A. and Meddis, R. (2001). "A human nonlinear cochlear filterbank”, *J. Acoust. Soc. Am* 110 (6): 3107-3118.

29. Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (2006).“Speech perception problems of the hearing impaired reflect inability to use temporal fine structure,” *Proc. Natl. Acad. Sci.* U.S.A. 103: 18866–18869.

30. Lutfi, R. A. and Patterson, R. D. (1984), “On the growth of masking asymmetry with stimulus intensity.”, *J. Acoust. Soc. Am*. 76, 739.

31. Moore, B.C.J, and Glasberg, B.R. (2001). “Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners.” *J. Acoust Soc. Am*. 110: 1067-1073.

32. Moore, B. C. J. and Sek, A. (1996) “Detection of frequency modulation at low modulation rates: Evidence for a mechanism based on phase locking .” *J. Acoust. Soc. Am*. 100, 2320.

33. Moore, B. C. J., Shailer, M. T., Schooneverldt, G. P. (1992), “Temporal modulation transfer functions for band-limited noise in subjects with cochlear hearing loss. *Br. J Audiol*. 4: 229-237.

34. Moore B. C,. J. and Skrodzka, E . (2002), “Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation”, *J. Acoust. Soc. Am*. 111, 327

35. Nelson, D. A. and Schroder, A. C. (1995),”Critical bandwidth for phase discrimination in hearing-impaired listeners, *J. Acoust. Soc. Am*. 98, 1969.

36. Patterson, R.D., Robinson, K., Holdsworth, J., McKeown, D., Zhang, C. and Allerhand M. (1992) Complex sounds and auditory images. In: Cazals, Y.,

Demany, L., Horner, K. (Eds), *Auditory physiology and perception, Proc. 9th International Symposium on Hearing*. Pergamon, Oxford, pp. 123-177.

37. Pick, G. F. (1980), "Level dependence of psychophysical frequency resolution and auditory filter shape", *J. Acoust. Soc. Am*. 68, 1085

38. Rosen, S. and Baker, R. J. (1994). "Characterizing auditory filter nonlinearity." *Hear. Res*.73: 231-243.

39. Rosen, S. , Baker, R. J. and Darling, A (1998), "Auditory filter nonlinearity at 2 kHz in normal hearing listeners.", *J. Acoust. Soc. Am*. 103, 2539.

40. Rosen, S. and Stock D. (1992) "Auditory filter bandwidths as a function of level at low frequencies (125 Hz--1 kHz)", *J. Acoust. Soc. Am*. 92, 773

41. Shamma, S. A. (1988), "The acoustic features of speech sounds in a model of auditory processing: vowels and voiceless fricatives," *J. Phonetics*, vol.16, 77-91.

42. Steeneken, H. J. M. and Houtgast, T. (1980) "A physical method for measuring speech-transmission quality." *J. Acoust. Soc. Am*. 67, 318

43. Summers, V., and Leek, M.R. (1994). "The internal representation of spectral contrast in hearing impaired listeners." *J. Acoust. Soc. Am*., 95, 3518-3528.

44. Supin, A., Popov, V. V., Milekhina, O. N. and Tarakanov, M. B. (1998), "Ripple density resolution for various rippled-noise patterns.", *J. Acoust. Soc. Am*. 103, 2042

45. Viemeister, N. F. (1979). "Temporal modulation transfer functions based upon modulation thresholds," *J. Acoust. Soc. Am*. 66, 1364–1380.

46. Wang, K., and Shamma, S. A. (1994). ''Self-normalization and noise robustness in early auditory representations,'' *IEEE Trans. Speech Audio Process*. 2, 421–435.

47. Weber, D. L. (1977), "Growth of masking and the auditory filter", *J. Acoust. Soc. Am*. 62, 424.

48. Won J. H, Drennan, W. R. and Rubinstein, J. T. (2007), "Spectral-Ripple Resolution Correlates with Speech Reception in Noise in Cochlear Implant Users". *JARO*, 8: 384–392.

49. Yost, W., and Moore, M. (1987). ''Temporal changes in a complex spectral profile,'' *J. Acoust. Soc. Am*. 81, 1896–1905.

50. Zanten, G. A. and Senten, C. J. J. (1983), “Spectro-Temporal Modulation Transfer Function (STMTF) for various types of temporal modulation and a peak distance of 200 Hz”, *J. Acoust. Soc. Am*. 74, 52.