USING MODEL AUDITORY-NERVE RESPONSES TO OPTIMIZE COMPRESSION GAINS FOR HEARING AID AMPLIFICATION SCHEMES

USING MODEL AUDITORY-NERVE RESPONSES TO OPTIMIZE COMPRESSION GAINS FOR HEARING AID AMPLIFICATION SCHEMES

By

FAHEEM DINATH, B.Eng

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Masters of Applied Science

McMaster University

Masters of Applied Science (2008) McMaster University
(Biomedical Engineering) Hamilton, Ontario

TITLE: **Using Model Auditory-Nerve Responses to Optimize Compression Gains for Hearing Aid Amplification Schemes**

AUTHOR: Faheem Dinath, B.Eng (McMaster University)

SUPERVISOR: Dr. Ian C. Bruce

NUMBER OF PAGES: xi, 122

Abstract

Linear and nonlinear amplification schemes for hearing aids have thus far been developed and evaluated based on perceptual criteria such as speech intelligibility, sound comfort, and loudness equalization. Finding amplification schemes that optimize all of these perceptual metrics has proven difficult. Using a physiological model, this thesis investigates the effects of single-band gain adjustments to linear amplification prescriptions. Using a gain optimization strategy, optimal gain adjustments for model auditory-nerve fiber responses to speech sentences were found to be dependent on whether the error metric included the spike timing information (i.e., a time-resolution of several microseconds) or the mean firing rates (i.e., a time-resolution of several milliseconds). Results showed that positive gain adjustments are required to optimize the mean firing rate responses, whereas negative gain adjustments tend to optimize spike timing information responses. The results were examined in more depth using a similar optimization scheme applied to a synthetic vowel /ɛ/. It was found that negative gain adjustments (i.e., below the linear gain prescriptions) minimize the spread of synchrony and deviation of the phase response to vowel formants in responses containing spike-timing information. In contrast, positive gain adjustments (i.e., above the linear gain prescriptions) normalize the distribution of mean discharge rates in the auditory nerve responses. Lastly, optimal compression gains were examined using a synthetic vowel /ɛ/ when the auditory threshold shifts were a result of solely outer or inner hair impairment. Results show that, on the whole, inner hair cell impairment requires less compression gains when the neurogram contained spike-timing information, and greater gains when containing mean-rate response. Outer hair cell impairment showed a need for greater gains in both neurogram types.

Abbreviations

dB	=	decibel
SPL	=	Sound Pressure Level
IHC	=	Inner Hair Cell
OHC	=	Outer Hair Cell
AN	=	Auditory Nerve
NAL-R	=	National Acoustics Laboratory - Revision 1
DSL	=	Desired Sensation Level
CF	=	Centre Frequency
BF	=	Best Frequency
REUG	=	Real Ear Unaided Gain
REAG	=	Real Ear Aided Gain
REIG	=	Real Ear Insertion Gain

Acknowledgements

I would like to express my sincere thanks to my advisor, Dr. Ian C. Bruce, for his unwavering support and patience throughout the course of my thesis, to Mohammed Negm for being a very hospitable and accommodating lab-mate, and most of all, I am grateful for having an extremely supportive and loving family.

Contents

List of Figures

Chapter 1

Introduction

Our sense of hearing is a remarkable feat of biology. The complex organs and structures that make up our ear create a highly specialized mechanism that detects and transduces sound vibrations into neural pulses. It is damage to this system and the resulting inability to hear sounds that defines the term hearing loss. Hearing aids compensate for damage done to our auditory apparatus by amplifying sounds to restore normal hearing sensation. Prescribed amplification gains, however, are not perfect and as a result are undergoing much research to enhance their design.

Amplification prescriptions have a foundation in early empirical studies showing that the most comfortable gain at a particular frequency equals approximately half the hearing threshold shift at the same frequency. This is referred to as the "half-gain" rule. That is, for every 1 dB increase in hearing threshold, a 0.5 dB gain adjustment is needed for tones to be deemed most comfortable (Lybarger, 1978). Today, popular linear hearing aid prescriptions, including the National Acoustic Laboratories' Revision 1 (NAL-R) and the Desired Sensation Level (DSL) prescriptions, are based on modifications of the half-gain rule and on judgments of speech intelligibility, sound comfort, and loudness equalization (Dillon, 2001).

The goal of this thesis is to find and analyze optimal single-band gain adjustments, i.e. compression gains, around the NAL-R and DSL prescribed gains by using models of the neural representation of speech rather than by perceptual feedback.

1.1 Thesis Layout

In the following chapter, Chapter 2, we begin by discussing sounds and the mechanisms of speech production. We describe in depth the fundamental anatomy and physiology of the auditory periphery and how it transduces sounds into neural signals. We also discuss some important audiological topics on compression, loudness recruitment, adaptation, and neural coding. Next, in Chapter 3, we discuss two types of linear hearing aid amplification prescriptions, the DSL and NAL-R, and describe

how they calculate amplification gains for particular hearing loss profiles. In Chapter 4 we present the Zilany and Bruce (2006) auditory periphery model that we use throughout this thesis. Additionally, we give examples of tools we use to analyze the model's neural output. In Chapter 5 we discuss gain optimization strategies used to optimize speech sentences, spoken vowels, and synthetic vowels before summarizing the results. Finally, we conclude with discussions in Chapter 6.

1.2 Related Publications

Some of the work presented in Chapter 5 has been published as a refereed conference paper: Faheem Dinath and Ian C. Bruce, "Hearing aid gain prescriptions balance restoration of auditory nerve mean-rate and spike-timing representations of speech", 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, August 2008. Part of the results in this paper were also published in I. C. Bruce, F. Dinath, and T. J. Zeyl, "Insights into optimal phonemic compression from a computational model of the auditory periphery", International Symposium on Auditory and Audiological Research, August 2007.

Chapter 2

Background

2.1 Acoustics of Auditory Stimuli

The sounds we hear consist of rapid changes in air pressure that are typically presented in one of three forms; as vibrations, as puffs, or as noisy turbulence. At standard temperature (20 °C) and pressure (101.325 kPa), sounds travel at a speed of 331 m/s, or about 1192 km/h, independent of amplitude or frequency. For periodic pressure waves, frequency (Hz) measures the rate of repetitions of a waveform per second. The human ear is able to detect tones in the range of 20 Hz to 20 kHz, although this range varies significantly with age and degree of hearing damage. By our teens, our ear cannot perceive tones close to 20 kHz, and as we age, we progressively lose our high frequency hearing range. However, this loss does not usually hinder our ability to understand speech since most sounds in speech take place at frequencies between 200 Hz to 8000 Hz, well below 20 kHz (Gelfand, 1998).

The instantaneous magnitude of sound pressure is measured in SI units of Pascal (Pa). The range of sound pressures the human ear is able to detect is enormous. By the time a sound becomes uncomfortable, it may be around 10 million times the level at which it is barely audible (Gelfand, 1998). To represent such a large range of values in an easy to interpret manner, sound pressure is measured on the decibel scale. The decibel scale is a logarithmic scale, in the base power 10, of a particular sound pressure with respect to a reference sound pressure (Warren, 1999). The level in dB is referred to as the sound pressure level (SPL) and is defined as:

$$\text{dB} = 20 \cdot \log_{10}(\frac{P_1}{P_2}) \tag{2.1}$$

where P_1 is the measured sound pressure level and P_2 is the reference sound pressure level (Warren, 1999). The standard reference pressure 2^{-5} Nm^{-2}, or equivalently 20 µPa, is the quietest sound humans can hear. It is important to remember that decibel is a relative measurement. A 0 dB SPL does not indicate that no sound exists,

rather that the sound is comparable to the reference quantity of 20 µPa. To help give an idea for sounds measured in terms of dB, a quiet whisper in a library has a sound pressure level of about 30 dB, a normal conversation at a meter is about 40–60 dB, traffic noise on a major road is approximately 80–90 dB, and some sounds, such as gun shots, can reach levels of 100–120 dB or more (Fig. 2.1). Sound pressure levels greater than 100 dB are painful and can cause permanent hearing damage even for relatively brief exposures.

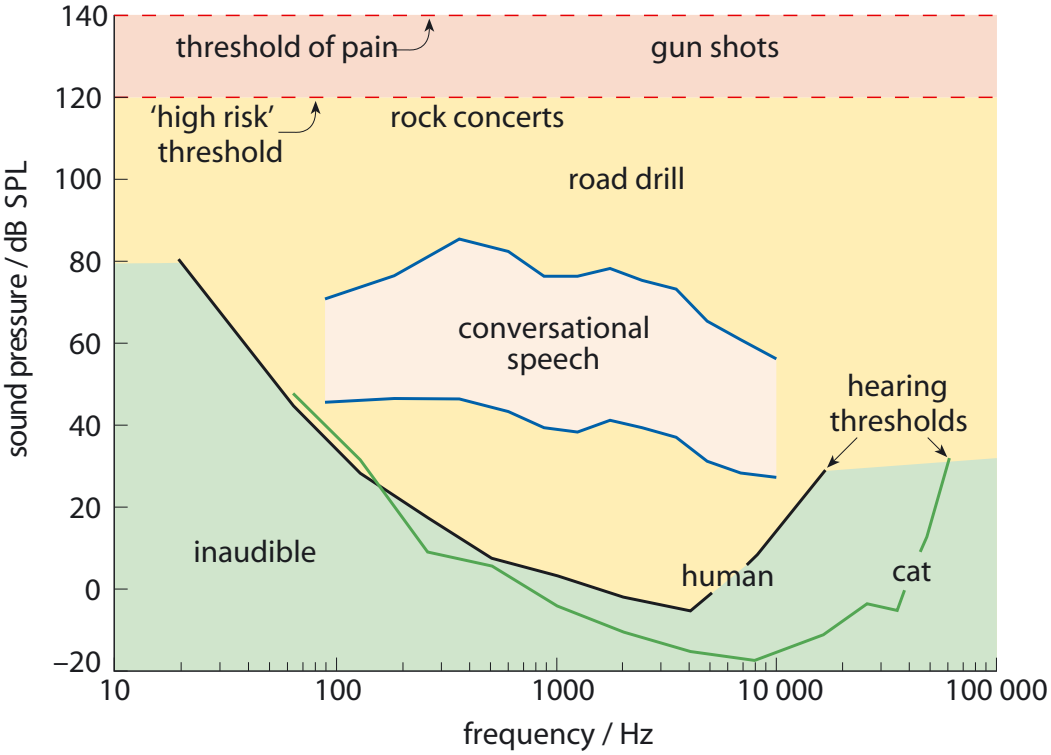


Figure 2.1: Conversational speech sound pressure levels. Human auditory threshold curve shown in black and for comparison, cat auditory threshold curve shown in green. Taken from Roberts (2002).

2.2 Speech Production and Characteristics

The mechanism for speech production consists of the vocal cords, larynx, nasal cavity, tongue, teeth, and lips, in addition to the lungs and diaphragm. The mouth, tongue, lips, and nasal cavity are often referred to collectively as the vocal tract.

Production of speech signals can be classified as either voiced or unvoiced depending on the nature of the excitation of the vocal tract. In voiced speech, the

surrounding muscles of the vocal cords pull them taut while the lungs, providing a steady flow of air, cause the cords to vibrate. The effect is much like a blowing through the mouth piece of a reed instrument; it sounds like a buzz. The more taut the cords, the higher the buzz pitch. Typical pitch values of normal adults lie between 50–500 Hz, with females generally on the higher end, and males on the lower (Warren, 1999). The resulting pressure pulses from the vocal cords propagate through the vocal tract where they are filtered to produce voiced speech. The pitch value of the vocal tract determines the fundamental frequency of voiced speech.

Voiced speech is characterized by dominant peaks in the spectral envelope called formants (Fig. 2.3) that indicate resonance frequencies of the vocal tract. We can manipulate the position or existence of formant frequencies by articulating the vocal tract. That is, by contracting and relaxing the surrounding musculature, we can change the shape of the vocal tract and therefore its resonance points. The first three formants, F1, F2, and F3 are usually the most dominant, with F1 generally the largest, followed by the second and third formants, respectively (Warren, 1999). These formants are important in vowel production and intelligibility because we can produce different vowels depending on the position of formant frequencies. Studies by Patterson and Barney (1952) have shown that vowels may be identified solely by differences in the first and second formant frequencies (see Fig. 2.2). By plotting the first two formants on an F1 vs F2 plane for a number of different vowels, we see that vowel's of the same type can be clearly distinguished as they tend to congregate in the same area with little overlap of adjacent areas.

In unvoiced speech, relaxation of the surrounding vocal cord musculature loosens the cords allowing a steady turbulent flow of air to pass from the lungs. When this turbulence radiates out from the vocal tract mostly unchanged, it is known as aspiration. However, the vocal tract can add further turbulence to aspired flow causing the resulting speech to be quite noisy. Examples of such type of speech are consonants such as “sh”, “th”, “ss”. Finally, some consonants such as “Zz” are classified as voiced fricatives. They are a result of mixed excitation whereby unvoiced turbulence is produced in addition to voiced periodic waveforms by the vocal cords.

2.3 Anatomy and Physiology of the Auditory Periphery

2.3.1 Outer Ear

The anatomical structure of the ear (Fig. 2.4) can be divided into three sections; the outer, middle ear, and inner ear. The outer ear, also called the auricle, consists of a cartilaginous, conical structure that acts to funnel sounds into the auditory canal and to the tympanic membrane (commonly called the ear drum). The shape and

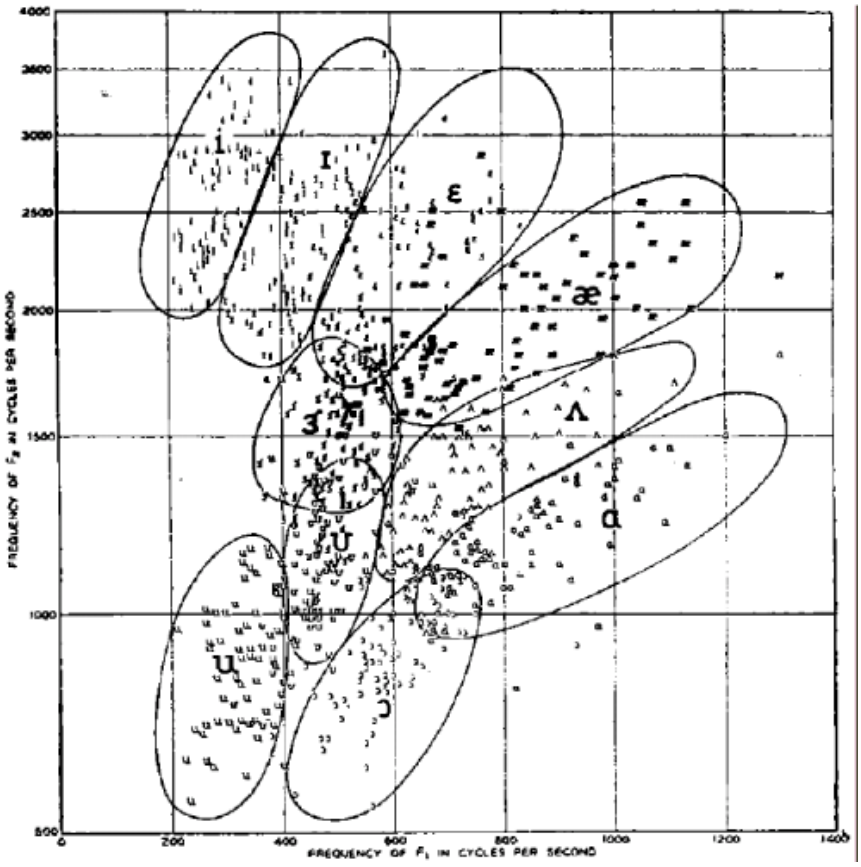
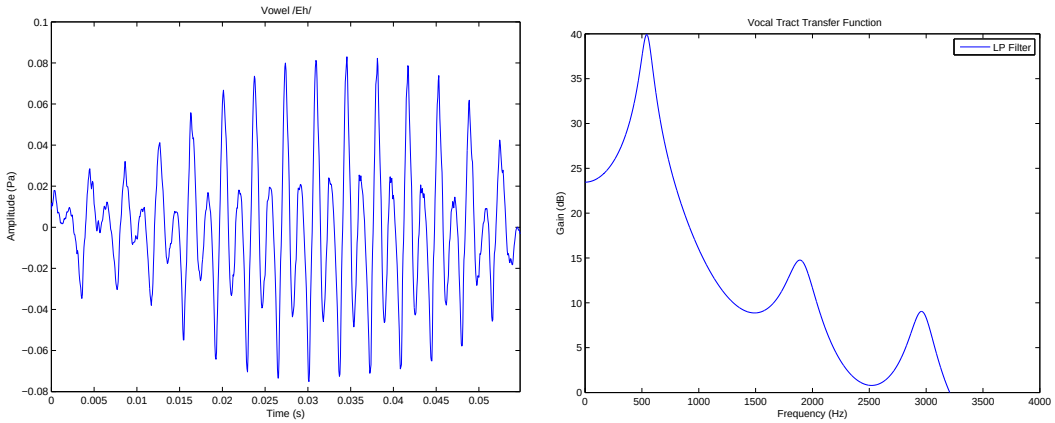


Figure 2.2: Formant F2 versus F1 for 10 vowels by 76 speakers. Speakers included 33 men, 28 women, and 15 children speaking American English. Vowel regions are marked by: [i] heed, [I] hid, [ɛ] head, [æ] had, [a] father, [⊃] ball, [U] hood, [u] who'd, [ʌ] hud, [ɚ] heard. Figure taken from Patterson and Barney (1952).

size of the auditory canal varies widely from person to person but the basic structure remains the same; it is characterized as a tube opened at one end and closed at the other.

The function of the canal is to spectrally enhance incoming sounds at a frequency equal to its resonance frequency - at a wavelength about four times the length of the canal. Since the canal is roughly 2.5 cm long in humans, the resonance frequency should occur at a wavelength of 10 cm or about 3.3 kHz. The peak resonance, however, is slightly different due to the non-uniform shape of the canal. Studies of the ear canal transfer function (Wiener and Ross, 1946; Shaw, 1974) show gains greater than 3 dB between 1–6 kHz and a maximal gain of 17.5 dB at roughly 2.5 kHz (Fig. 2.5)



(a) (b)

Figure 2.3: An example of the temporal waveform of a voiced /ε/ vowel shown in (a). Transfer function of the vocal tract that produced the /ε/ vowel is shown in (b). Note that the peaks of the transfer function coincide with locations of formant frequencies. The vocal tract transfer function was estimated from the /ε/ vowel using linear predictive coding.

2.3.2 Middle ear

Pressure changes (rarefactions and compressions) at the end of the auditory canal vibrate the ear drum, a thin, pliable membrane that separates the outer ear from the air-filled, middle ear cavity. Vibrations are transmitted to three tiny bones in the middle ear called ossicles, the malleus, incus, and stapes. The malleus, which is directly attached to the tympanic membrane, transmits mechanical vibrations to the incus and then to the stapes. The footplate of the stapes is attached to the oval window of the cochlea, another thin membrane although here it covers the entrance to the fluid-filled cochlea (see Figs. 2.4 and 2.6). Sound pressure waves are transmitted from the tympanic membrane to the fluid-filled cochlea in a rather complex manner. Since sound pressure waves do not readily cross air/water interfaces, and greater than 99 percent of the incident pressure waves are reflected back into the air medium, therefore if the ear consisted simply of a membrane covering the entrance to the cochlea, hardly any sound energy from air would travel into the fluid filled cochlea. In order to create sufficient fluid movement in the cochlea, much greater pressures are needed at the oval window than what are seen at the tympanic membrane. Because of mechanical leverage of the ossicle system and the reduced area of the stapes foot plate in comparison to the ear drum (a factor of about 17), the middle ear thereby converts small pressures at the ear drum into large pressures at the oval window by increasing forces over a smaller area and reducing the displacement (Roberts, 2002). The stapes, connected to a membrane over the oval window, plunges back and

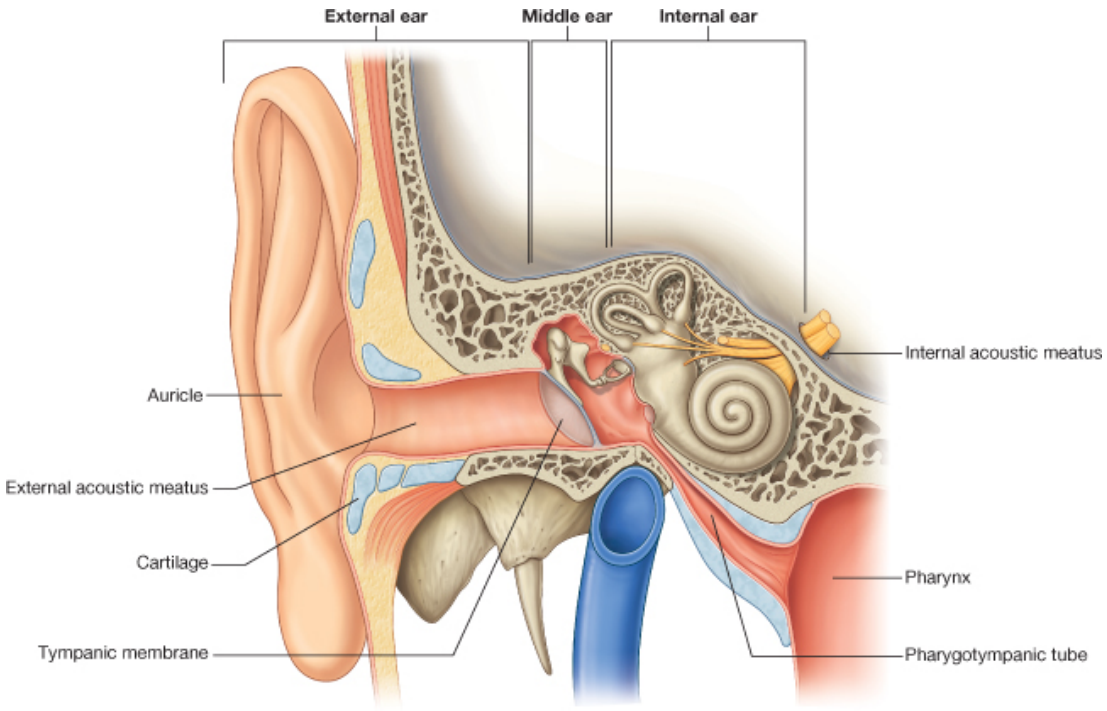


Figure 2.4: Cross sectional view of the ear. Figure taken from Drake et al. (2004).

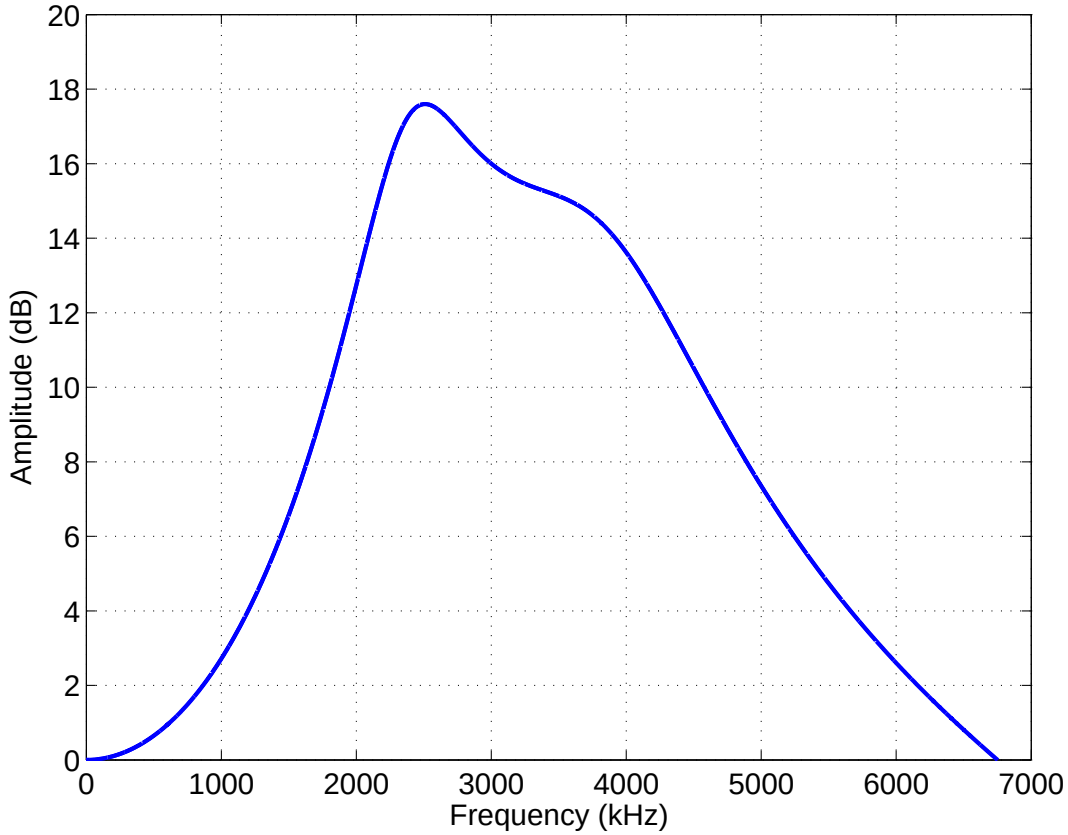


Figure 2.5: Head-related transfer function (HRTF) from Wiener and Ross (1946)

forth creating sufficient pressure waves that travel through the fluid-filled cochlea and ultimately stimulate the auditory nerve. Additionally, the complex mechanical coupling in the middle ear is essential to ensure maximum energy transfer from air to cochlea fluid over a wide range of frequencies (Fettiplace, 2006)

2.3.3 Inner ear

The inner ear is an enclosed cavity formed in the temporal bone. It consists of the membranous labyrinth, the organ responsible for orientation and acceleration, and the cochlea, the organ responsible for hearing (see Fig. 2.6). The cochlea, which resembles that of a snail shell, is a tube roughly 3.5 cm long, coiled in about 2.75 turns. The tube is partitioned down its length by Reissner's membrane and the basilar membrane to form three fluid-filled canals called the scala vestibule, scala media, scala tympani. At the apex, the scala vestibule and scala tympani are joined through a small opening called the helicotrema. The fluid contained in these two scalae is called perilymph and the fluid in the scala media called endolymph. The two partitioning membranes are highly flexible, and pressure waves sent down the scala vestibule are transmitted through the scala media, into the scala tympani, before finally being relieved at the round window. In this process Reissner's membrane and the basilar membrane both vibrate, but in regard to sound sensation, it is the movement of the basilar membrane that is most important. This is because attached to the basilar membrane is the organ of Corti, the organ containing hair cells responsible for converting mechanical vibrations into electrical signals, i.e., performing the sensory transduction.

2.3.4 The Basilar Membrane

The basilar membrane (BM) is a wedge-shaped structure with a width of about 0.1 mm at its base (by the oval window) and about 0.5 mm at the apex (see Fig. 2.7). As the BM widens, it also becomes more flexible, with the apex being nearly 100 times more flexible than near the base (Gelfand, 1998). It is because of this non-uniformity that a gradient of resonance frequencies is established along the length of the membrane. As pressure waves move along the basilar membrane, high frequency components maximally displace regions on the basilar membrane closer to the base, and low frequency components maximally displace regions closer to the apex (Fig. 2.7). The basilar membrane is organized tonotopically, that is, each point along the basilar membrane responds maximally (resonates) to one particular frequency called the characteristic frequency (CF). Alternatively, one could think of each point on the basilar membrane having a band-pass filter with a narrow pass-band at the CF.

Research conducted by von Bksy (1960) describes the mapping of CF's along the length of the basilar membrane as being:

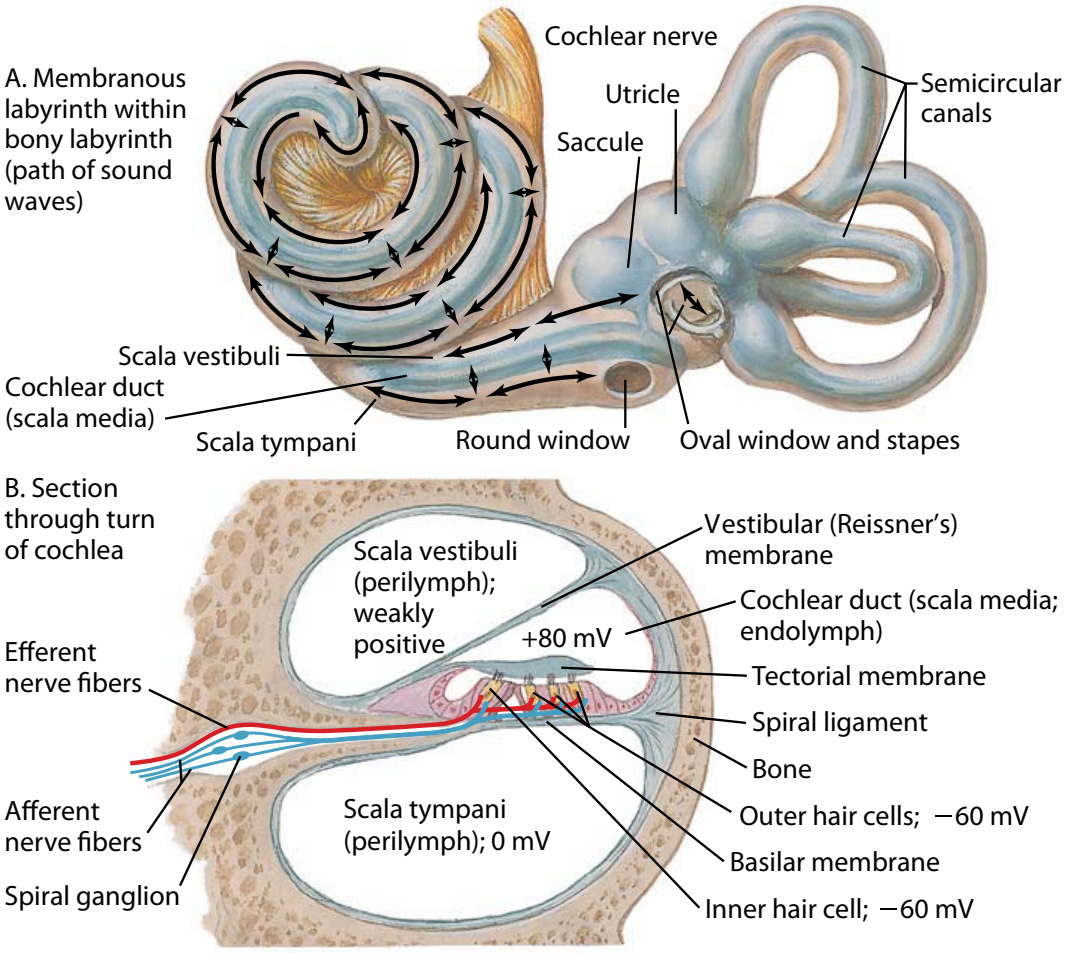


Figure 2.6: Sagittal and cross-sectional view of the cochlea. Figures taken from Netter (2002).

$$\mathrm{CF} = 10^{x\alpha} - k \tag{2.2}$$

where CF is expressed in kHz. The variable x is the distance from the apex as a proportion of basilar membrane length, and ranges from 0 to 1. The constant α is typically 2.1 and remains the same in many cochleae whereas k varies between 0.8 and 1 (typically 0.85) between mammalian species. Interestingly, equation 2.2 expresses a relationship between the linear basilar membrane and the logarithmic CF. The result is that sections at the base of the basilar membrane have a coarse frequency resolution, becoming increasingly finer in sections further towards the apex.

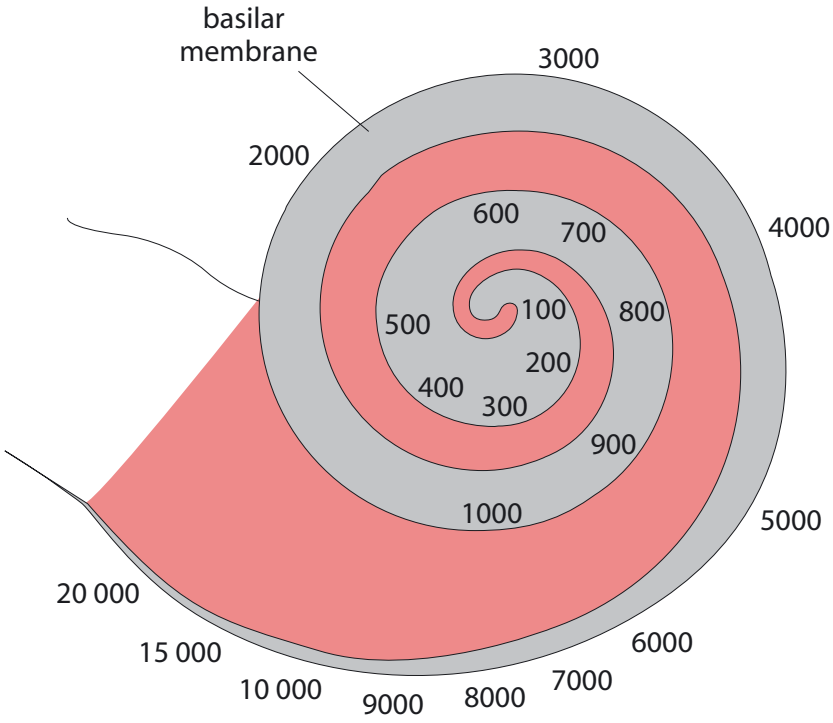


Figure 2.7: Frequency distribution along the basilar membrane. The basilar membrane is narrow and stiff at the base where it is most sensitive to high frequency tones. At the apex, the basilar membrane is wide and flexible where it responds maximally to low frequency tones. Taken from Roberts (2002).

2.3.5 The Organ of Corti

The organ of Corti runs the complete length of the basilar membrane. It consists of a single row of about 3500 inner hair cells (IHC) and three rows of about 12000 outer hair cells (OHC) separated by a triangular tunnel called the tunnel of Corti (see

Fig. 2.8). Hair cells are the sensory cells that act as mechano-electrical transducers to convert sound-induced mechanical vibrations into electrical signals. The top of each hair cell consists of a circular plate containing bundles of tiny finger-like projections called stereocilia. Each bundle is typically organized in a staircase pattern, with the longest stereocilia in each bundle connected to the tectorial membrane (Gelfand, 1998). Vibration of the basilar membrane exert shearing forces that tilt stereocilia back and forth.

Depending on the direction of tilt, stereocilia can either hyperpolarize or depolarize the cell owing to protein cross-links which connect from the side of one stereocilia to ion channels gating proteins in the tip of shorter, adjacent stereocilia. Positive tilting of stereocilia causes a cascade of pulls that open ions channels of shorter adjacent stereocilia, thereby causing an influx of positive ions which help depolarize the hair cell (See Fig 2.9). That in turn helps generate action potentials in auditory nerve fibers synapsed to the hair cell (Robles and Ruggero, 2001). Furthermore, given that about 10 percent of ions channels are open at a time, negative tilting aids in closing these open channels leading to a reduction in the hair cell membrane potential (Roberts, 2002).

Hair cell potentials are graded, that is, they do not produce action potentials themselves owing to a lack of voltage gated sodium ion channels. Rather, hair cells change their membrane potential depending on the magnitude of the stimulus and the resulting influx of ions. In sum, a positive deflection of the basilar membrane results in an increase in hair cell potential and a negative membrane deflection results in a decrease in hair cell potential (Roberts, 2002).

Finally, it is important to note that damage to stereocilia due to prolonged exposure to loud sounds is one of the most common causes of sensory hearing loss. Stereocilia are delicate structures that are easily susceptible to damage by excessive movement of the basilar membrane. Most stereocilia frequently survive severe noise exposure but may become disarrayed and floppy. Over time, excessive noise exposure can lead to progressive, irreversible, degeneration of stereocilia (Henderson and Hamernik, 1995).

2.3.6 Cochlear Amplifier

Hair cells of the organ of Corti transmit and receive abundant neural information along the auditory nerve. Interestingly, almost all the information going to higher levels of the auditory system is derived from the IHCs and almost none from OHCs although there are almost 4 times as many OHCs as there are IHCs (Roberts, 2002). Ongoing investigation with OHCs have revealed that they play a unique and important role in the basilar membrane mechanics. OHCs contain motor proteins which increase and decrease the length of the cell in response to changes in cell's potential, the purpose of which is to oppose the damping effects of the endolymph on the basilar

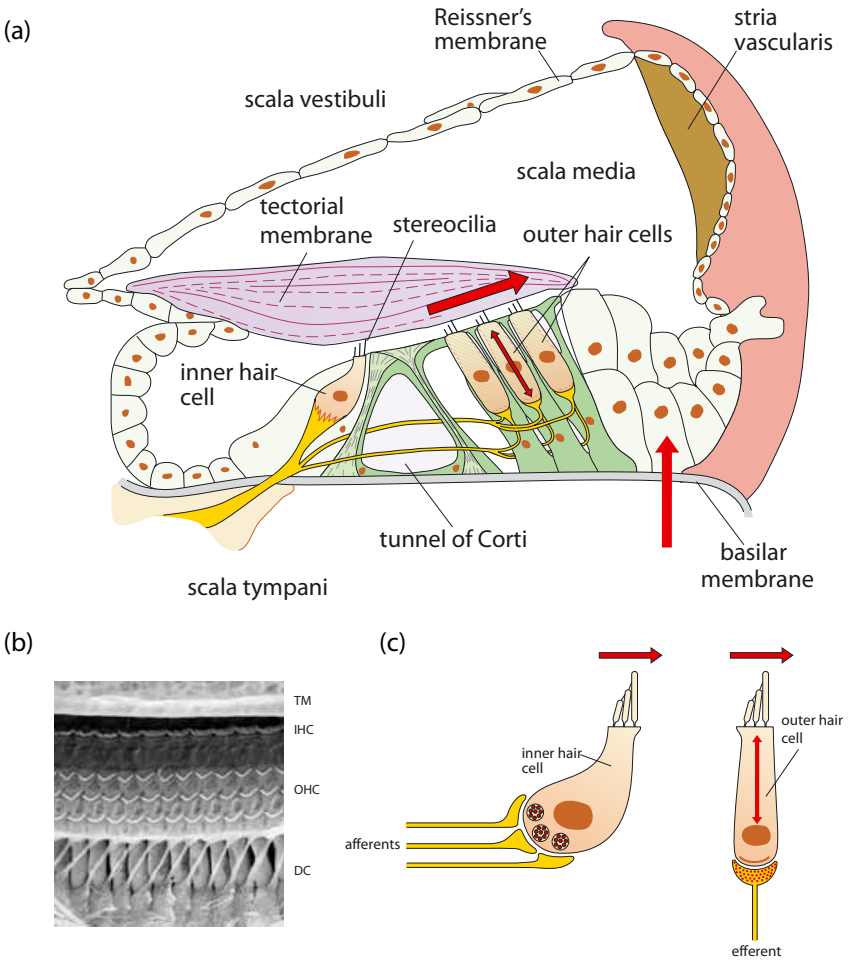


Figure 2.8: Shown in (a): a cross-sectional view of the cochlea, with the major components indicated. Red arrows indicate forces exerted during speech stimuli. Note the presence of the force generated by the outer hair cell. In (b): a scanning electron micrograph of the organ of Corti showing 1 row of inner hair cells and 3 rows of outer hair cells. In (c): a closer view of the inner and outer hair cells. Figures modified from Roberts (2002).

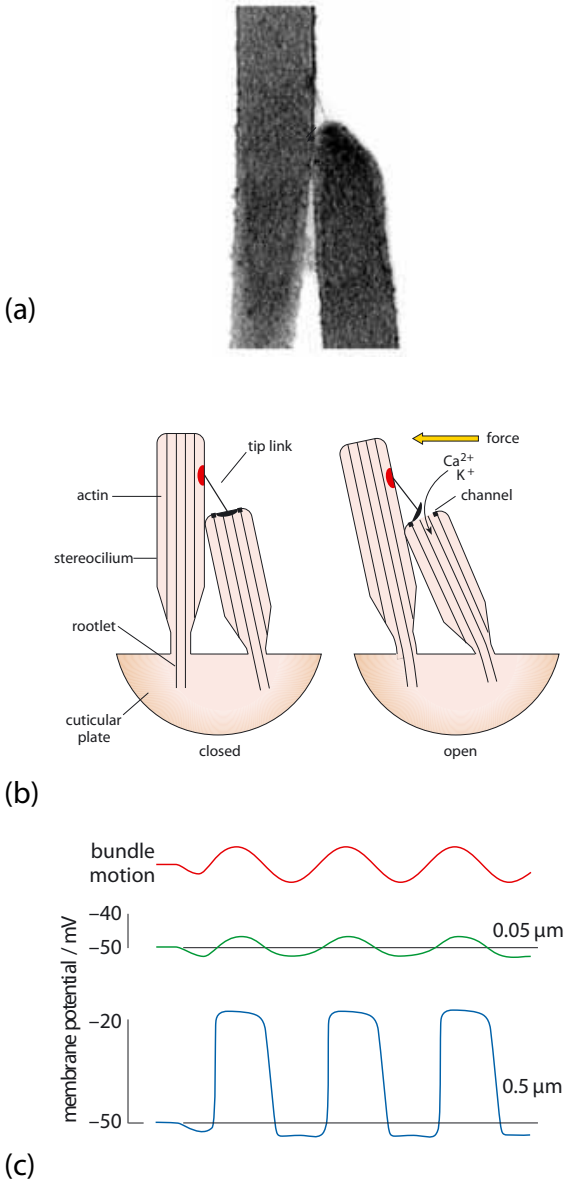


Figure 2.9: Shown in (a): an electron micrograph of two stereocilia connected with a cross-link. In (b): a force applied (arrow) to the stereocilia causes the cross-link to pull open the ion channel which allows positive ions to flow into and depolarize the hair cell. In (c): the membrane potential as the stereocilia deflect by a small 0.05 µm distance (green) and a large 0.5 µm distance (blue). Figures modified from Roberts (2002).

membrane (refer to Fig. 2.8). This allows for greater amplification due to resonance in the basilar membrane in response to tones, an effect known as the 'cochlear amplifier' (see Fig. 2.10). Without the cochlear amplifier, sounds would have to be at least 40dB greater in order for the basilar membrane to respond the same as it does normally at 0 dB SPL (Roberts, 2002).

2.4 Nonlinearities of the Auditory Periphery

2.4.1 Frequency Tuning Curves

Auditory nerve fibers at characteristic frequency locations along the basilar membrane generally exhibit greatest sensitivity to that frequency as well. This most sensitive frequency is labeled the best frequency (BF) of the auditory nerve fiber, and is essentially synonymous to CF. The response of an auditory nerve fiber given a range of pure tones is often shown as a tuning curve such as that in Figure 2.11A. The general form of the tuning curve has a gentle slope on the low frequency side and a steep slope on the high frequency side, but there are also tuning curves that are quite symmetrical on either side of the BF (Roberts, 2002). The measure of sharpness of the tuning curve is described by the Q10db parameter which is defined as the BF divided by the bandwidth of the tuning curve 10dB above the threshold at BF. Thus the sharper the tuning curve and higher the BF, the greater the Q10db (Sachs et al., 2002).

Normal neural response depends on the well-being of both the inner and outer hair cells. Damage to OHCs causes the response of the basilar membrane to diminish which in turn affects the response of the IHCs. In this case, the tip of the tuning curve exhibits a broadening and shallowing, with an associated decrease in Q10dB. Damage to the IHCs alone primarily causes an elevation in tuning curve, but generally does not affect the Q10dB (see Fig. 2.12). With OHC damage in particular, the result is that the auditory nerve fiber becomes less tuned to its best frequency, and in both cases of IHC or OHC damage, the auditory nerve fiber may exhibit difficulty responding solely to its best frequency.

Figure 2.13 shows the effect of noise-induced hearing loss in the cat auditory periphery. Sensory hearing loss was introduced by exposing the cats to high level noise, about 110 dB SPL centered at 2 kHz. Tuning curves exhibit elevated thresholds and broadened bandwidth. Subfigure B shows an apparent elevation in the best threshold curve (BTC) around the 2 kHz region.

2.4.2 Compression

We have mentioned that the cochlear amplifier aids in displacing the basilar membrane maximally at particular locations along its length in response to resonance

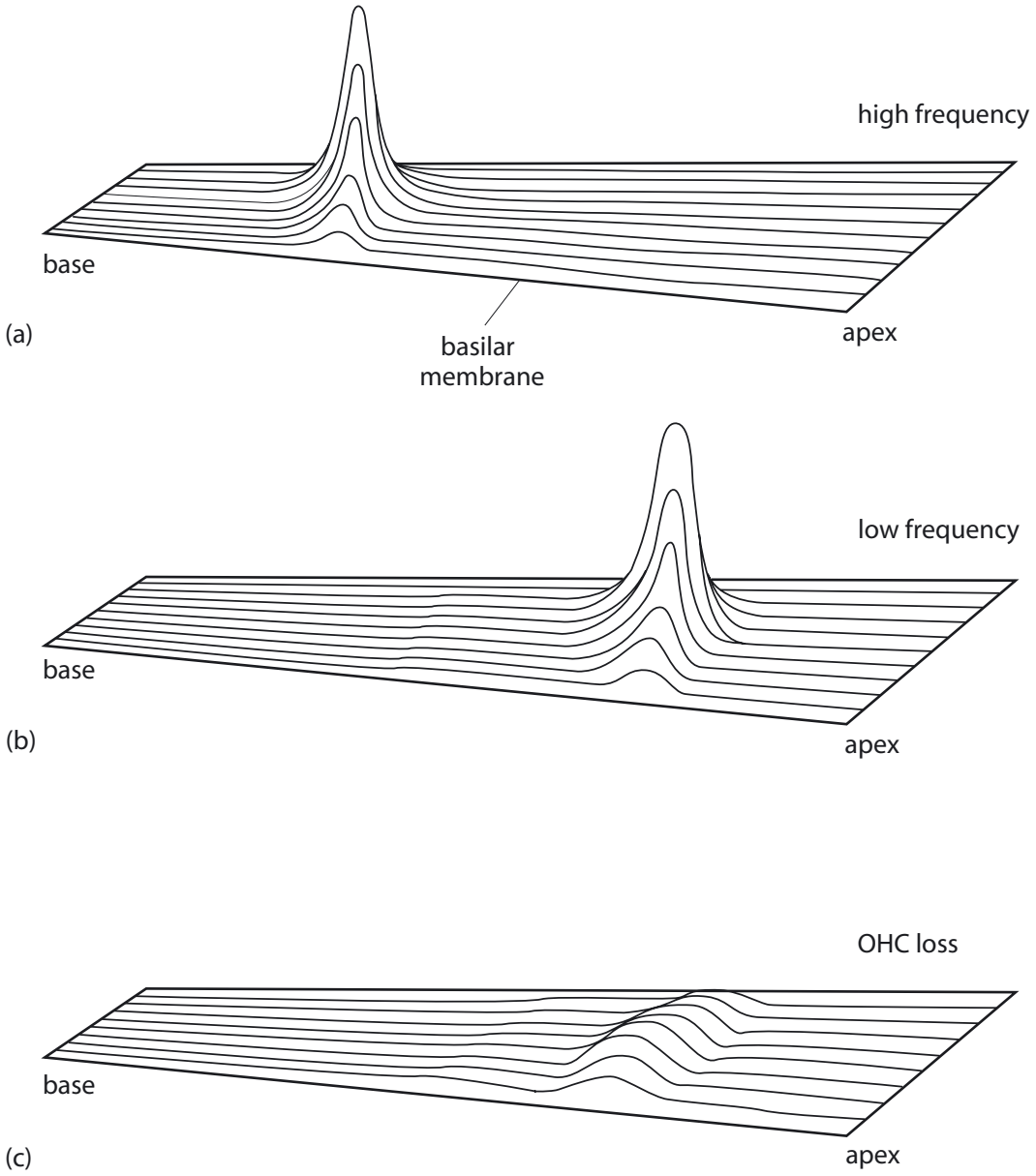


Figure 2.10: Shown in (a) and (b): the envelope of maximum oscillation of the basilar membrane in response to a high and low frequency tone. In (c); the same oscillatory response to a low frequency tone in absence of the OHCs cochlear amplifier effect. Figures taken from Roberts (2002).

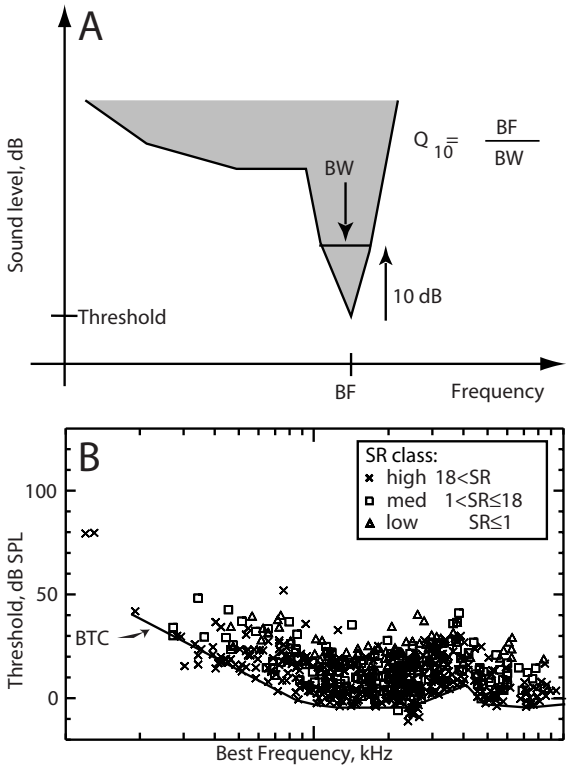


Figure 2.11: In A: tuning curve at a specific CF in the normal cat auditory. SPL is shown on the ordinate and frequency is shown on the abscissa. Tones with amplitudes in the grey area increase the fiber's spontaneous rate. Note that the tuning curve dips to a minimum SPL threshold value at the best frequency indicating greatest sensitivity to that tone. in B: plot of best frequency thresholds along the BM. The solid line supporting the points represents the best threshold curve for animals used in the study. Points are classified by spontaneous rates as shown in legend. Figure taken from Sachs et al. (2002).

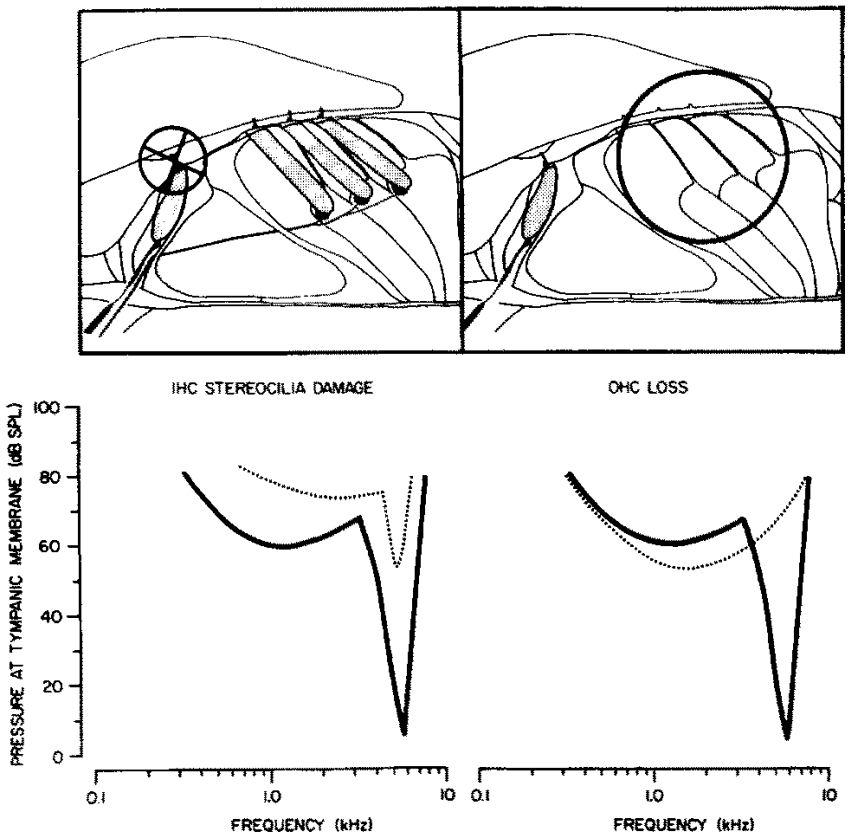


Figure 2.12: Effect of hair cell loss on the tuning curve. Normal tuning curve shown in the solid black line. Left: Dotted line shows IHC impairment due to acoustic trauma. Right: Dotted line shows OHC impairment due to ototoxic drug kanamycin. Figure taken from Kiang et al. (1986).

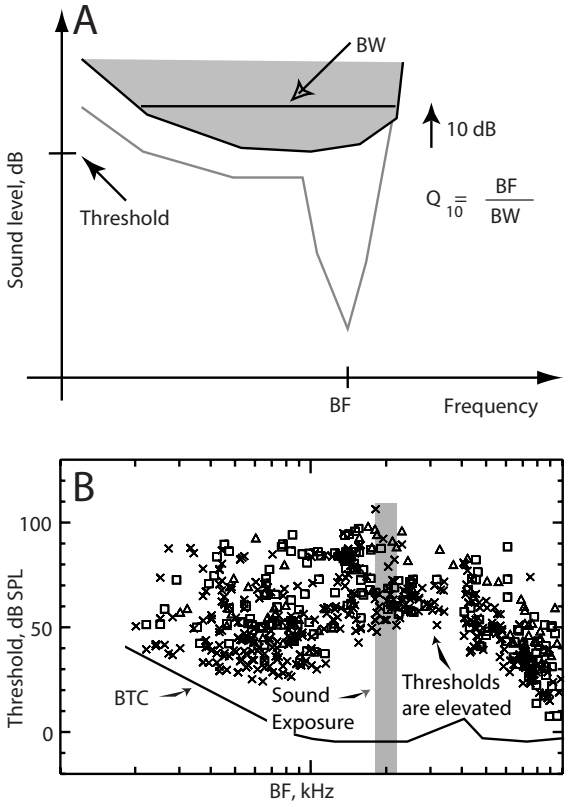


Figure 2.13: In A: example impaired tuning curve at a specific CF in a cat auditory periphery. In B: changes in best frequency thresholds as a result of sound induced hearing. In this case, 110 dB noise exposure centered at 2 kHz. Figure taken from Sachs et al. (2002).

frequencies in sounds. One important feature associated with this mechanism is a high degree of nonlinearity called compressive non-linearity. It allows the membrane to respond quite rapidly to sound pressures starting from very low sound pressure levels, however at a certain compression point, the basilar membrane decreases its response with increases in sound pressure level. By damaging the OHCs completely and thereby removing the cochlear amplifier, compressive non-linearity is removed and hearing threshold is elevated.

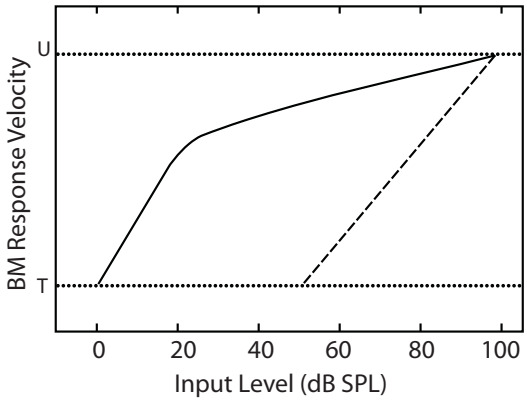


Figure 2.14: Basilar membrane input-output curve at an arbitrary CF. Basilar membrane response velocity (arbitrary units) is shown on the ordinate, and input SPL on the abscissa. Figure modified from Oxenham and Bacon (2003).

Figure 2.14 shows the response of one point along the basilar membrane to a tone at the characteristic frequency. The dotted lines along points T and U show the level of threshold where sounds are barely audible and where sounds becomes uncomfortable. The solid and dashed lines represent the basilar membrane input-output curves, with the solid line representing the normal cochlea and the dashed line an impaired cochlea. The difference between the two curves is purely the cochlear amplifier contribution by OHCs. A normal basilar membrane shows no compression at low SPLs, that is, a 1:1 ratio between change in SPL and change in response velocity, and high compression within mid-range sounds pressures, normally about 5:1. This differs from impaired cochlea where lack of the cochlear amplifier may result in a direct 1:1 relationship.

Examining the plots more closely provides a lot of information about hearing impairment and speech perception. From the normal curve, sounds are first perceived at 0 dB SPL and rise compressively to become uncomfortable at 100 dB SPL. The normal cochlea is able to provide the full 100 dB SPL dynamic range by becoming more compressive, or less responsive, as sound pressure levels increase. From the impaired curve, however, sounds become uncomfortable at the same 100 dB SPL as in the normal cochlear, but are first perceived at 50 dB SPL. The rise in membrane response is quite quick compared to the normal curve, and covers a much smaller

dynamic range of about 50 dB.

Hearing aids are designed to overcome a reduced dynamic range and loss of compressive non-linearity by amplifying soft sounds to threshold, avoiding amplification of loud sounds to painful or damaging levels, and restoring normal compressive growth of sound pressures.

2.4.3 Loudness Recruitment

Inner and outer hair cell impairment are two factors that regulate the function of the auditory nerve (AN) discharge rate function (see Fig. 2.15). The auditory nerve discharge rate function is the rate at which action potentials emanate in a nerve at specific CF location as a function of tone SPL. Since damage to outer hair cells removes compression and the cochlear amplifier effect in the basilar membrane, this acts to both increase hearing threshold and steepen the AN fiber discharge rate function. Threshold here is the SPL at which spike rate increases above the base-line spontaneous rate. Second, with damage to the inner hair cells, however, there is a precipitous decline in both the the slope of the AN discharge rate function and the maximal discharge rate response. The combination and degree of these two impairments produce a wide range of slopes and threshold shifts in the AN rate function, increasing the difficulty of creating hearing prescriptions to counter this effect.

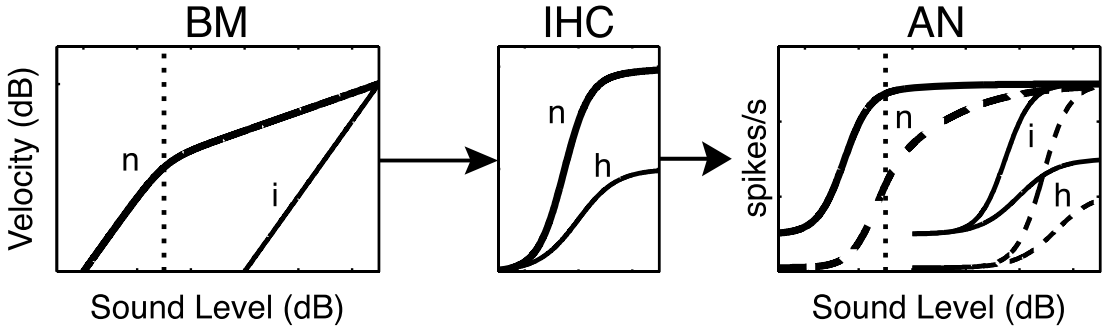


Figure 2.15: Left: Basilar membrane velocity in the normal (n) and impaired (i) ear as a function of sound input. Loss of OHC cochlear amplifier primarily responsible for differences seen between the two curves. Centre: Discharge rate of the inner hair cell as a function of sound pressure is shown in the normal(n) and impaired (h) ear. Right: High spontaneous rate fibers shown in dashed lines and low spontaneous rate fibers in solid. Three cases are presented; No impairment (n), OHC impairment only (i), and both IHC and OHC impairment (h). Figure taken from Heinz et al. (2005).

One effect of steepening of the AN rate level function is loudness recruitment. The term loudness recruitment means a faster than normal perceived growth in sound amplitude between elevated threshold levels and high sound pressure levels. Studies

trying to identify the cause of loudness recruitment have shown that growth of loudness corresponds well to the abnormally fast growth in basilar membrane response when outer hair cells are damaged (refer to Fig. 2.14). This suggests that loudness recruitment may be a result of a steepened basilar membrane response and that this response must be conveyed to higher processing centers via steepened rate level functions. Heinz et al. (2005), however, has concluded that for mixed auditory hearing loss, AN rate level functions were not steeper. OHC damage does in fact steepen the slopes of the AN rate level functions but not enough to overcome the shallowing of the rate level function by IHC damage, except in the case of extreme hearing loss. Hearing loss typically is a mixture of both inner and outer hair cell damage rather than just one of the two. Since mixed hearing loss produces shallower rate level functions, AN rate level functions themselves would not be sufficient to explain growth of loudness. Heinz et al. (2005) has offered an alternate explanation for loudness recruitment: that it could ultimately depend on neural mechanisms of the central auditory system that are sensitive to AN discharge spike trains which have synchronized to frequencies other than at the BF.

2.4.4 Adaptation

When presented with a sudden, sustained stimulus, the auditory nerve initially responds with a high discharge rate that decreases monotonically to a steady-state rate. This decline of discharge rate is called adaptation. Adaptation in the auditory nerve has been described by a number of researchers in response to acoustic (Eggermont and Spoor, 1973; Westerman and Smith, 1984; Javel, 1996) and electric stimuli (Killian et al., 1994; Haenggeli et al., 1998). The mechanisms behind adaptation are many and include release of neurotransmitter from IHCs, activation of neurotransmitter receptors in auditory nerve, and transmission of action potentials along the auditory nerve (Nourski et al., 2006). Overall, the general function of adaptation is to keep the auditory system operating in a region of maximum sensitivity (Fettiplace, 2006).

Figure 2.16 below shows the histogram from a physiological experiment conducted by Smith (1977) in which an adaptation response was taken from a 6 to 12 month old Mongolian gerbil auditory nerve. A 204 ms adaptation test tone at 25 dB SPL with a frequency equal to the auditory nerve CF was applied and the response was taken directly from the an auditory nerve using a glass pipette. On application of the test tone, we see an initial high rate of response that is followed immediately by adaptation which reduces the response to steady state and sets it into a new region of maximum sensitivity.

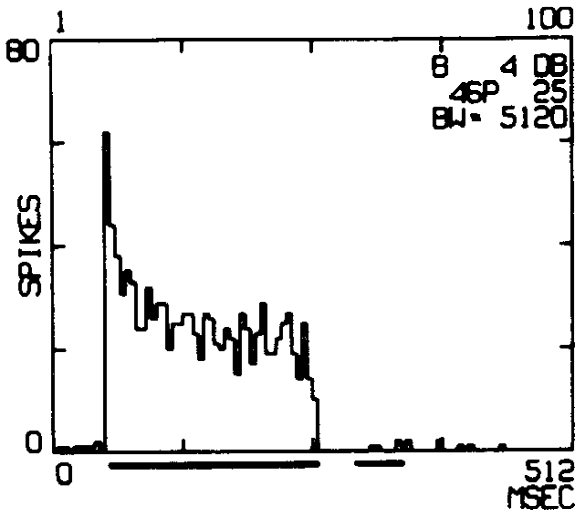


Figure 2.16: Adaptation in a Mongolian gerbil auditory nerve in response to a 204ms test tone with frequency equal to nerve CF, 7.4 kHz in this case. Spike count is shown on the ordinate and histogram bins of 5120 µs in width to 512ms on the abscissa. Figure taken from Smith (1977).

2.5 Neural Coding

Communication between the cochlear and higher auditory processing centers in the brain consists of nothing more than spike trains. Since spike trains can vary only in the number of spikes in a given amount of time and the length of the inter-spike interval, all relevant information of speech sensory perception has to be encoded by manipulating only these two parameters (Yarmush et al., 2003). Some properties of speech can be encoded with longer intervals between spikes and some with shorter. Some properties can be encoded linearly or perhaps non-linearly (Yarmush et al., 2003). When analyzing the neural code, we can represent it as: a) average discharge rate and b) spike-timing information. We assume here that, common to both representations, tonotopic mapping preserves the resonance information along the basilar membrane in the auditory nerve (Yarmush et al., 2003).

If a speech stimulus is applied, and neural spike trains are elicited, certain temporal patterns, or properties such as strength of neural response, can be observed by obtaining the rate representation of the neural output. A rate representation is calculated by averaging the neural response over large time bins on the order of milliseconds or more (refer to Fig. 2.17). The result of this procedure is called the average discharge rate response. Because of the large time bins, average discharge rate representations usually have low time resolutions but show the slow moving temporal patterns of the speech signal, referred to as the “envelope”.

Naturally we can also observe the fine spike-timing of the neural discharge events to view the fine temporal response of a speech stimulus (refer to Fig. 2.17). Information gathered here can give clues as to how well neurons capture and phase lock to certain frequencies in speech. This is because the basilar membrane resonates at the same frequency of the stimulus at the CF region on the membrane and as such the vibration patterns of the stimulus frequency is transferred into the neural impulses travelling to the auditory periphery.

Neurons can completely capture frequencies up to about 1000Hz from which point the absolute refractory period becomes too long. At frequencies greater than 1000Hz and less than 5000 Hz, synchrony falls off and fiber discharge events become sparse. In this case, it is assumed that the summed response of many nerve fibers synapsed to an IHC produces a complete response in the central auditory system (Garner, 1950). Above 5Khz, IHCs cannot synchronize well to the speech stimuli.

In addition to the mean-rate response and the fine-timing information, we assume that the auditory periphery preserves tonotopic mapping. If we recall from Fig. 2.7, the basilar membrane resonates at specific points along its length to specific frequency components in speech. Nerve cells that synapse to hair cells along the basilar membrane fire in response to the basilar membrane at that point of connection. The auditory nerve thus preserves the spatial orientation of CF points along the basilar membrane, an arrangement otherwise known as tonotopic mapping (Roberts, 2002). Thus, different frequency components of speech, such as formants, are represented in separate populations of AN fibers and sent in this manner to the central auditory system (Gelfand, 1998; Reale and Imig, 2004).

Presumably, it is likely that speech coding and perception can be characterized by a combination of these three representations, or possibly yet unproposed theories. For the purposes of computer simulations in Chapter 4, and experiments to follow in Chapter 5, we will examine the neural representation of speech as either spike-timing information or as mean spike rates.

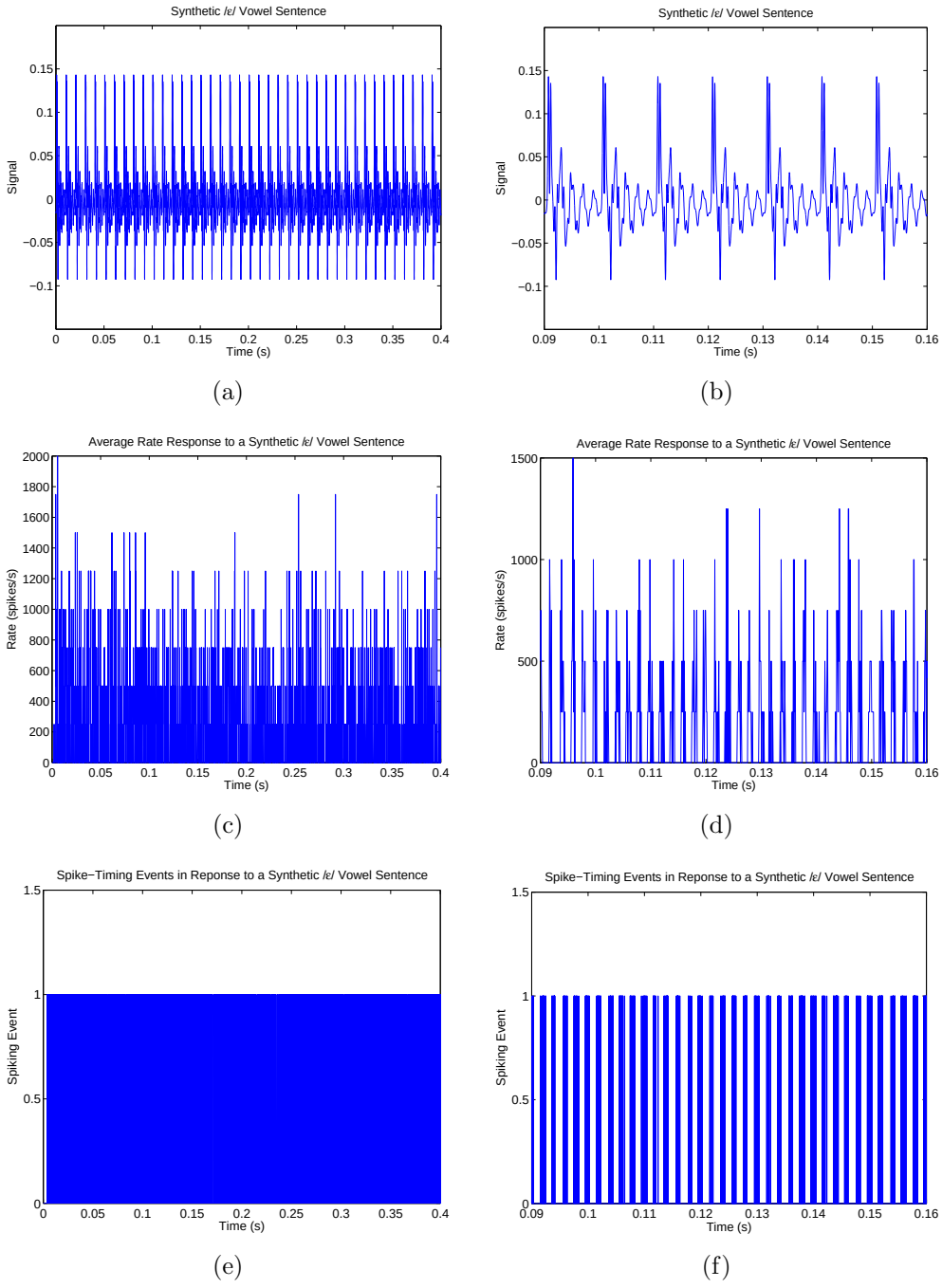


Figure 2.17: (a) shows a sample speech sentence of a repetitive /ε/ vowel with a closer portion shown in (b). Subfigures (c) and (d) show the mean-rate representation of the speech signal, having used a bin-width of 0.1 ms. Subfigures (e) and (f) show the spike-timing neural discharge events in response to the speech signal. Figures were generated using the Zilany and Bruce (2006) auditory periphery model described in Chapter 4.

Chapter 3

Hearing Aid Prescription Algorithms

3.1 Hearing Impairment

While we know much about the anatomical pathologies associated with sensory neural hearing loss, the resulting physiological and psychoacoustic effects are still under much ongoing research. As alluded to before, the combination and degree of damage to the inner and outer hair cells in sensory neural hearing loss produce a wide range of effects to the AN rate function. Hearing aids unfortunately have no option but to target a hearing aid prescription to the cumulative rather than individual effects of IHC and OHC impairment, making it very difficult to create adequate hearing aid prescriptions.

For much of the history of hearing aids, amplification prescriptions has been an endeavor of trial and error. Early attempts at creating hearing aid amplification prescriptions showed that the most comfortable gain at a particular frequency equals approximately half the hearing threshold at the same frequency. This is referred to as the "half-gain" rule. That is, for every 1 dB increase in hearing threshold, a 0.5 dB gain adjustment is needed for tones to be deemed most comfortable (Lybarger, 1978). This seems quite contrary to intuition which suggests that in order to compensate for hearing loss at a particular frequency, one need only apply an amount of gain equal to the amount of hearing loss. However this is not the case, and attempting to restore the dB loss at each frequency does not result in ideal restoration on the whole.

As we will describe shortly, popular linear hearing aid prescriptions, including the National Acoustic Laboratories Revision 1 (NAL-R) and Desired Sensation Level (DSL) amplification prescriptions are based on variations of the half-gain rule as well as perceptual criteria such as judgments of speech intelligibility, sound comfort, and loudness equalization (Dillon, 2001).

3.2 Hearing Aid Prescriptions

3.2.1 Real Ear Aided, Real Ear Unaided, and Insertion Gains

Hearing aid prescription formulae specify the desired amount of amplification in terms of either the real ear aided gains (REAG) or the real ear insertion gains (REIG). Real ear aided gains and insertion gains are related to a third type of gain called real ear unaided gains (REUG) by the formula:

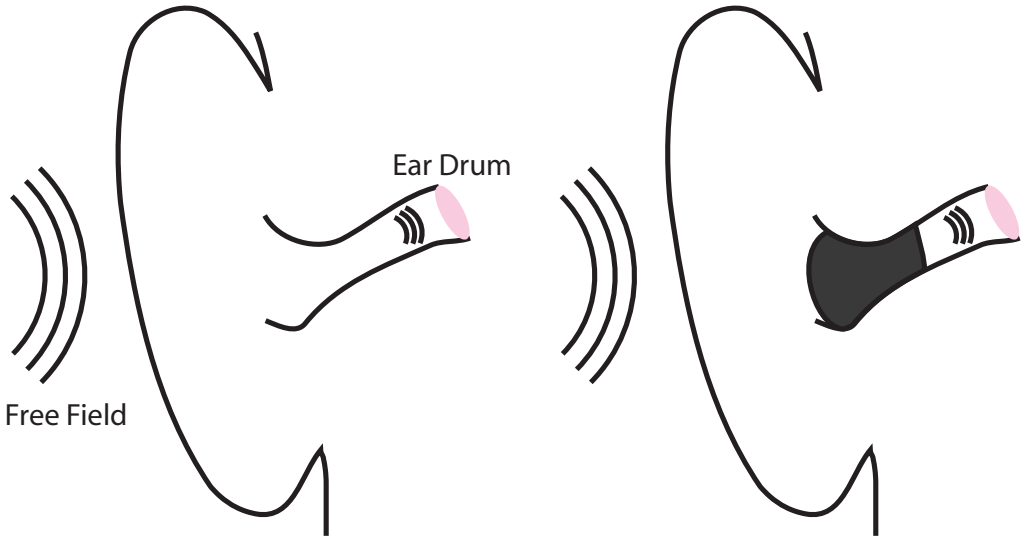
$$\text{REAG} = \text{REIG} + \text{REUG} \tag{3.1}$$

Real ear unaided gains are amplification gains provided by the transfer function of the ear canal. The gain profile is measured as the difference in sound pressure level between the point beside the ear drum relative to the point just outside the ear at the pina, namely, at free field (Fig. 3.1(a)). We have already seen the REUG profile for a typical adult ear as seen in the outer ear transfer function in Fig 2.5. The transfer function represents the gain-frequency response profile by the auditory canal itself. In contrast, real ear aided gain (REAG) is the gain-frequency response profile of an ear canal containing a working hearing aid. It is measured from the point of the ear drum relative to free field as shown in Fig. 3.1(b). The difference between REAG and REUG is called real ear insertion gains (REIG) and are the gains provided solely by the hearing aid device itself, beyond restoration of the REUG (Fig. 3.2). Audiologists and auditory researchers use these different gain measurements as ways to accurately fit amplification schemes to a broad range of individuals, each having their own unique differences in canal shapes and auditory loss profiles.

3.2.2 National Acoustics Laboratory Prescriptions

The Australia National Acoustics Laboratories' Revision 1 (NAL-R) hearing aid amplification scheme (Byrne and Dillon, 1986) has proved to be one of the most well-established amplification scheme for over a decade in hearing aid centers throughout Australia. The NAL-R originated from improvements over the amplification scheme developed by Byrne and Tonisson (1976) and has held true to the original aims: to amplify all frequency bands of speech to equal loudness.

The procedure proposed by Byrne and Tonisson (1976) describes insertion gains at specific frequencies for a given increase in hearing threshold. The general idea is to elevate the different frequency components in speech so that much of it would be audible when the user has selected the most comfortable hearing levels. Research from the National Acoustics Laboratories has shown that hearing aid users have notable improvements in speech intelligibility when all bands were equal in perceived loudness. It was reasoned that if hearing aid gains were to be based on all frequency bands, that this would maximize the amount of speech signal brought into a level of



(a) (b)

Figure 3.1: In (a), REUGs are measured over the unoccluded ear canal whereas in (b), REAGs are measured over an active hearing aid device in the auditory canal.

equal loudness.

The initial Byrne and Tonisson (1976) amplification formulation was based on the half-gain rule, a rule proposed by Lybarger (1944) and which forms the basis of many hearing aid prescriptions on the market. Their procedure prescribes frequency gains as 0.46 times the increase in hearing thresholds plus or minus a frequency dependent constant. The revised NAL-R formula, however, extends the principles from Byrne and Tonisson (1976) to include a constant, H_{3FA}, which is the hearing threshold three frequency average at 500Hz, 1000Hz and 2000Hz. The NAL-R formulation for calculating insertion gains is given in Table 3.1

Gains given by the NAL-R prescription are in terms of insertion gains (IG), that is, the gains provided by the hearing aid above the gains normally supplied by the outer ear's natural amplification.

3.2.3 Desired Sensation Level Prescriptions

The Desired Sensation Level prescription differs from the NAL-R procedure in that it does not try to make speech equally loud, but rather comfortably loud. DSL, developed by researchers at the University of Western Ontario in the 1980's, originated as a tool for pediatric audiologists to tailor amplification schemes specifically for the unique needs of children and infants. Motivation came after a 1974 rubella outbreak

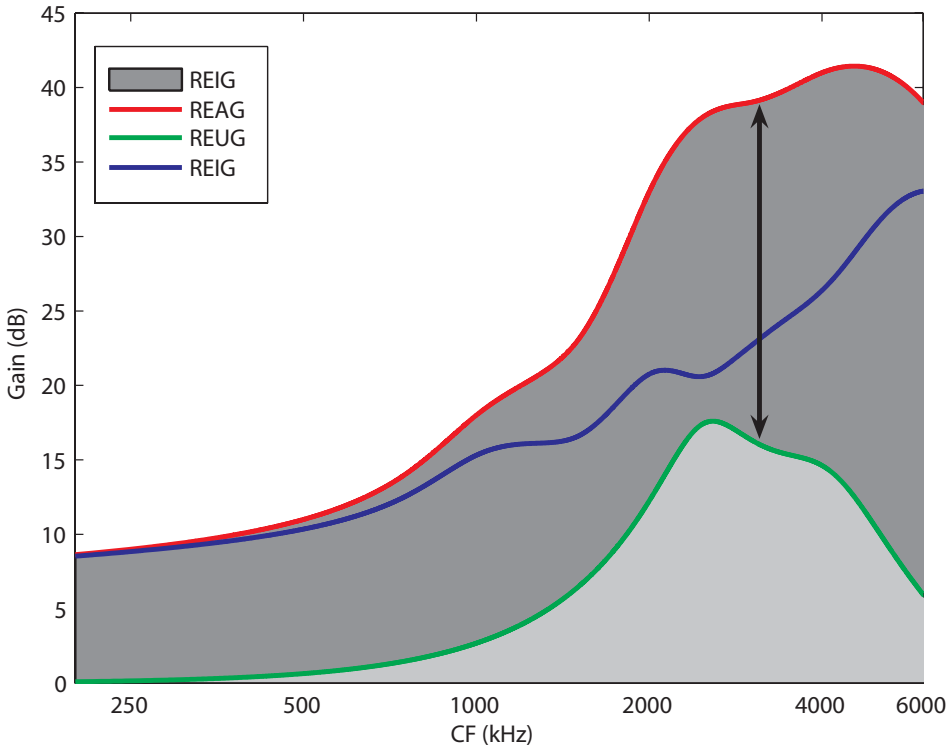


Figure 3.2: Real ear aided gains from the DSL algorithm is shown as the red curve. Insertion gains, shown in dark gray, is the difference between the real ear aided (red) and unaided (green) gains, and is also represented as the blue curve.

Table 3.1: NAL-R calculation. Table taken from (Byrne and Dillon, 1986).

1.	H_{3FA}	=	$(H_{500} + H_{1000} + H_{2000})/3$
2.	X	=	$0.15 \cdot H_{3FA}$
3.	G_{250}	=	$X + 0.31H_{250} - 17$
	G_{500}	=	$X + 0.31H_{500} - 8$
	G_{750}	=	$X + 0.31H_{750} - 3$
	G_{1k}	=	$X + 0.31H_{1k} + 1$
	$G_{1.5k}$	=	$X + 0.31H_{1.5k} + 0$
	G_{2k}	=	$X + 0.31H_{2k} - 1$
	G_{3k}	=	$X + 0.31H_{3k} - 2$
	G_{4k}	=	$X + 0.31H_{4k} - 2$
	G_{6k}	=	$X + 0.31H_{6k} - 2$

in the Children's Hospital in Halifax, Nova Scotia that left two clinical audiologists Richard Seewald and Patricia Stelmachowicz with an unexpected case of more than 30 children and infants with hearing loss. At the time there were no tools available to fit amplification schemes to young children, and most adult amplification schemes were not useful for children. In 1976, Seewald, in collaboration with Dr. Mark Ross at the University of Connecticut, began development of a simple linear amplification scheme, the Desired Sensation level (DSL) method, publishing their first original results in 1985 (Seewald et al., 1985).

The original version of the DSL algorithm was aimed for use in children, but continued development and research has expanded its role for use with adults. The original goal of DSL remains the same: to make speech sufficiently audible without discomfort. The desired sensation levels, or gains, are based on data describing the most comfortable hearing level associated with speech presentation level for different hearing impairments (Kamm et al., 1978). DSL prescribed hearing gains are given in the form of look-up tables (see Table 3.2) and expressed in terms of the real ear aided gain (REAG).

3.3 Audiograms

An individuals' hearing loss profile is represented as an audiogram. An audiogram is a semi-logarithmic plot, with frequency presented along a logarithmic axis and corresponding hearing loss values along the linear axis. Amplification frequencies are commonly selected as points between the range of 250 and 8000 Hz as this is where almost all speech hearing and intelligibility resides (Gelfand, 1998). Audiologists measure an individuals' audiogram by presenting pure tones at varying amplitudes

Table 3.2: Desired Sensation Level look up table. Table taken from Dillon (2001).

dB HL\Frequency	250	500	750	1000	1500	2000	3000	4000	6000	8000
0	0	2	3	3	5	12	16	14	8	6
5	3	4	5	5	8	15	18	17	11	10
10	5	6	7	8	10	17	20	19	14	13
15	7	8	10	10	13	19	23	21	17	16
20	9	11	12	13	15	22	25	24	20	19
25	12	13	14	15	18	24	28	27	23	22
30	14	15	17	18	20	27	30	29	26	25
35	17	18	19	21	23	30	33	32	29	27
40	20	20	22	24	26	33	36	35	32	30
45	22	23	25	27	29	36	39	38	36	35
50	25	26	28	30	32	39	42	41	39	38
55	29	29	31	33	35	42	45	45	43	42
60	32	32	34	36	38	46	48	48	46	45
65	36	35	37	40	42	49	52	51	50	49
70	39	38	40	43	45	52	55	55	54	53
75	43	42	43	46	48	56	59	58	58	58
80	47	45	47	50	52	59	62	62	61	60
85	51	48	50	53	55	63	66	65	65	64
90	55	52	54	57	59	66	69	69	69	68
95	59	55	57	60	62	70	73	73	73	72
100	62	59	61	64	66	73	76	76	76	75
105	66	62	64	68	70	77	80	80	81	80
110	68	66	68	71	73	80	83	84	85	85

and pitches. When the tone is barely audible to the individual, the sound pressure level and pitch are recorded on the audiogram.

A typical person with normal hearing would have curves that lie around the 0 db hearing loss mark on the audiogram, however, individuals with some hearing loss have profiles that fall into one of 4 categories: Mild, Moderate, Severe, or Profound hearing loss. To calculate which category profiles belong to, we use the formula:

$$X = \frac{H_{250} + H_{500} + H_{1000} + H_{2000} + H_{4000}}{5} \tag{3.2}$$

If values of X lie between 0 and 20, then it is considered that there is no hearing loss, and that the individual's auditory periphery is functioning normally. For values of X greater than 20, hearing loss is present and considered mild if the value of X lie between 21 and 40, moderate between 41 and 70, sever between 71 and 95, and

profound for values greater than 95 (British Society of Audiology, 1988).

The NAL-R and DSL amplification prescriptions use different formulas to restore hearing and in doing so, it is not surprising that each prescriptions provides different gains for each type of hearing loss profile. Figures 3.3 and 3.4 demonstrate differences between the NAL-R and DSl algorithms in restoring examples of each of the 4 types of hearing loss profiles.

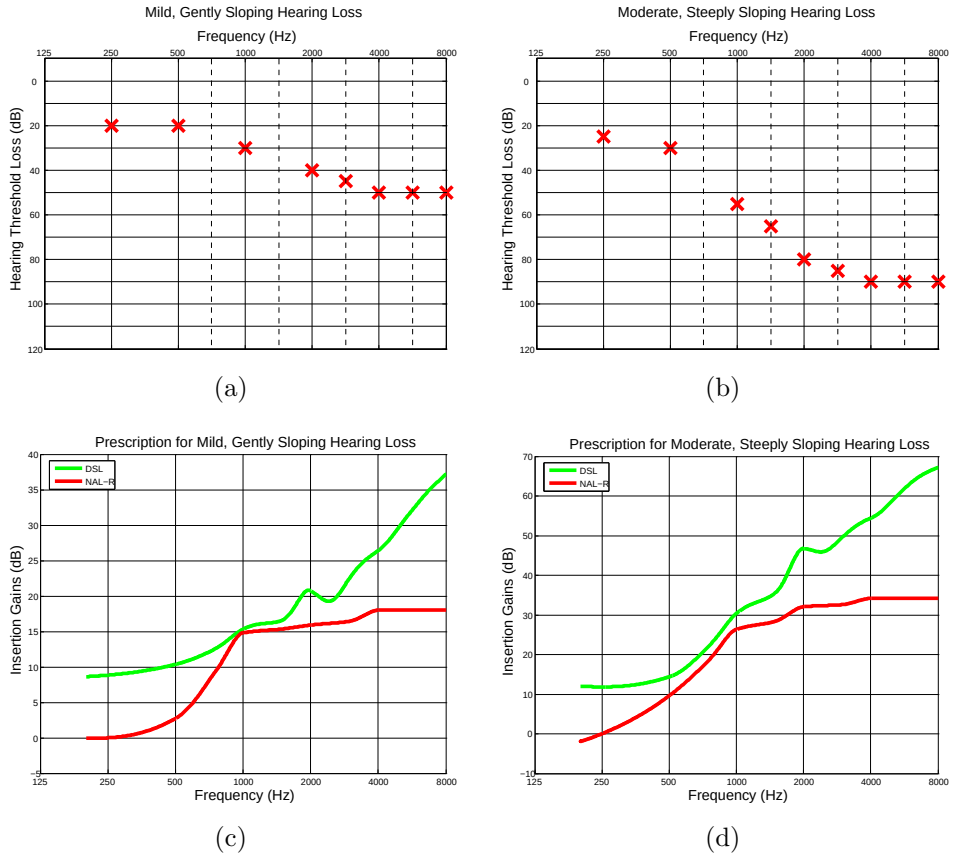


Figure 3.3: Example of mild and moderate hearing loss profiles (top row) with linear hearing aid amplification gains shown below (bottom row)

3.4 Compression Schemes

People with hearing impairment often suffer from a reduced dynamic range resulting in an abnormally rapid growth of perceived loudness. When sounds are too soft, they are inaudible, but with increases in sound pressure level, sounds quickly become too loud for comfort. Compensating for a reduced dynamic range is problematic for

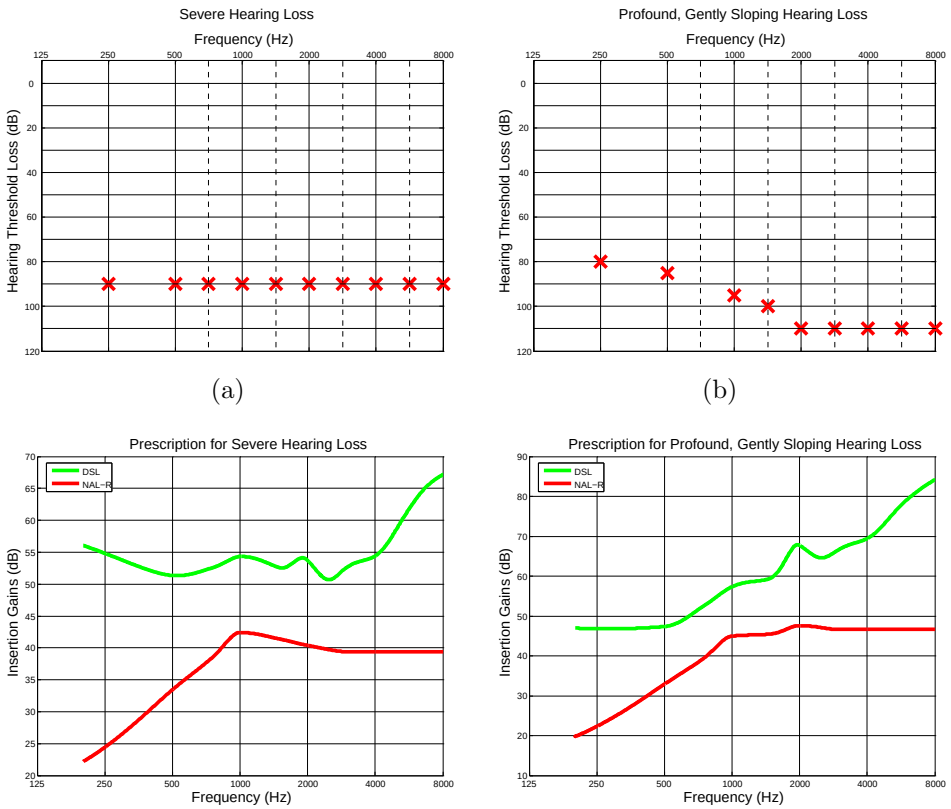


Figure 3.4: Example of severe and profound hearing loss profiles (top row) with linear hearing aid amplification gains shown below (bottom row)

linear hearing aids which work independently of input sounds pressure level by applying a certain amount of gain for a given threshold shift at a specific frequency. Thus, hearing aids often utilize compressions, the role of which is to reduce the dynamic range of incoming sounds. Compression gains act to adjust sounds pressure level at the output of the hearing aid device so that weak sounds are amplified more than loud sounds, and that very loud sounds are limited in SPL. Compression gains can be single band where gains are applied identically over all frequencies, or multi-band, where compressions gains vary independently over different frequency bands. Multi-band compression is necessary if dynamic ranges vary significantly in different frequency bands.

When compression gains are applied over a large range of input sounds pressure levels it is called wide dynamic range compression (WDRC). An example of WDRC is shown in Fig. 3.5. In the linear region, an overall gain of 30 dB is applied to sounds with SPLs below 40 dB. After this SPL, compression reduces gains by 10 dB for every 20 dB increase in SPL, giving an effective 2:1 compression ratio. Beyond 100 dB SPL, negative compressions gains are applied to prevent sounds from reaching dangerous levels.

Compression gains are applied to sounds dynamically according to changes in the instantaneous sound pressure level. The dynamic behaviour of compressions gains is characterized by attack and release times. When input SPL increases, the time it takes for the compressor to react to this increase is called the attack time, and similarly, when input SPL falls, the time it takes for compressor to react is referred to as the release time. For the purposes of this thesis, we assume that compression gains are applied instantaneously and constantly over a length of a given speech segment.

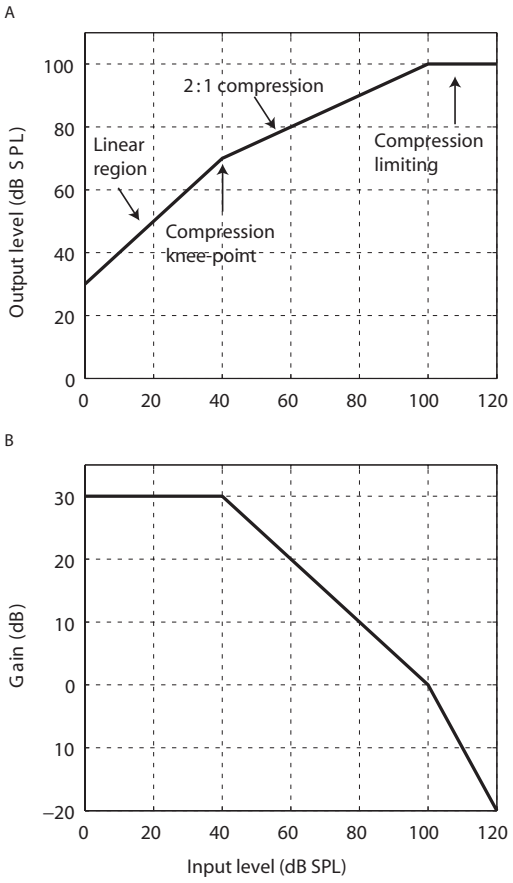


Figure 3.5: A: input vs output sound pressure levels. A 2:1 compression ratio is applied after 40 dB input SPL and output is limited after 100 dB SPL. B: input sound pressure level vs applied gains. Figures taken from Bruce (2007)

Chapter 4

Modelling the Auditory Periphery

In this chapter we will detail ways in which we can examine neural response of the auditory nerve using the Zilany and Bruce (2006) auditory periphery model, and show in what ways common amplification prescriptions, such as the NAL-R and DSL, affect the neural representation of speech.

4.1 Zilany-Bruce Cat Auditory Periphery Model

Experimental studies on the effects of hearing impairment in the auditory periphery have been performed on, for the most part, cats owing to physiological similarity between the human and cat auditory peripheries. However, the need for computer models becomes apparent given the costly and time consuming nature of creating appropriate hearing loss types and measuring the neural correlates of those impairments in animals. Computer models that simulate the auditory nerve provide a fast and easy way to determine the physiological effects of hearing impairment. Efforts by auditory researchers over the past 40 years have produced a number of physiological data sets, theories, and models describing the mechanisms of the auditory periphery. Early development of auditory periphery computational models in the 1980's utilized linear basilar membrane filters to model AN responses, however it was soon realized that non-linear basilar membrane properties account for important characteristics in AN responses. One of the first accurate models of the auditory periphery to include a non-linear basilar membrane was developed by Carney (1993). Over the past 15 years, continual improvements have brought about more complete and accurate models of the auditory periphery. Notable improvements made by Bruce et al. (2003) provide greater accuracy of normal and impaired auditory responses by allowing for IHC and OHC impairments. The latest model by Zilany and Bruce (2006) (see the schematic in Fig. 4.1) matches a much broader set of physiological data, is able to simulate the full dynamic range of hearing, and produces AN responses over a wider range of center frequencies than models before (up to 40 kHz rather than 4 kHz).

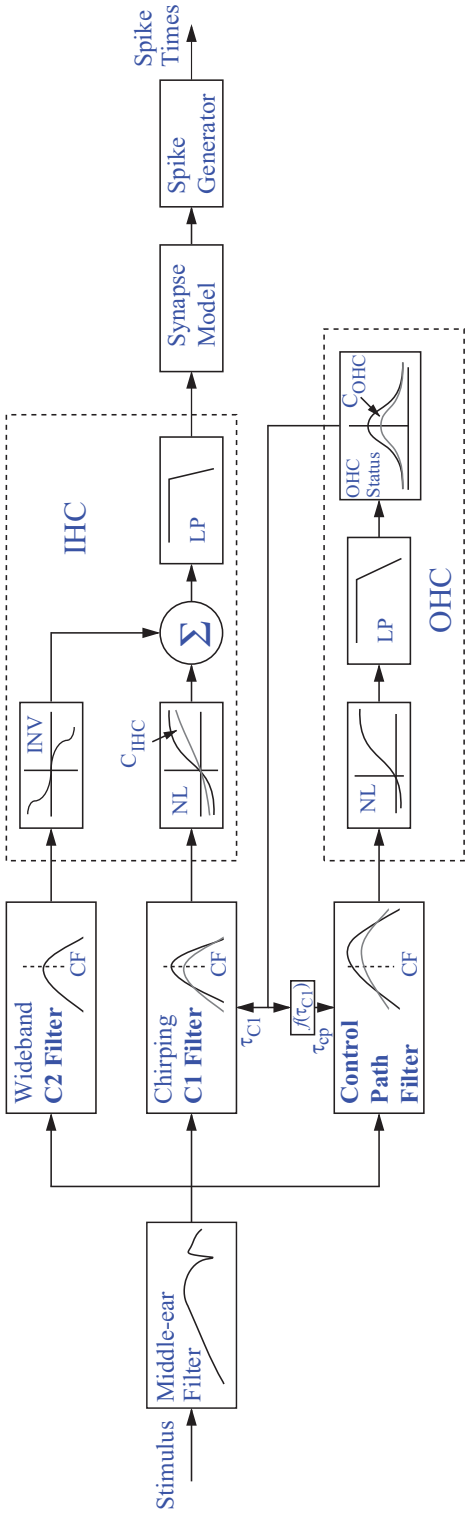


Figure 4.1: Zilany and Bruce cat auditory nerve model. Figure taken from Zilany and Bruce (2006).

The Zilany and Bruce (2006) model describes the auditory pathway from the middle ear through to the auditory nerve. The outer ear is not included in the model, but follows from the head-related transfer function (Wiener and Ross (1946)) described in section 2.3.1. Input to the middle ear consist of speech waveforms with instantaneous pressures in units of Pascal, sampled at a rate of 100 kHz. A high sampling rate is necessary to ensure stability in the middle ear filter and proper statistics in the neural spike generator at the output. After middle ear filtering, the signal is passed through three parallel filter paths: C1, C2, and the control path filter. C1 is a narrowly tuned and adjustable filter which dominates at low to moderate input pressures. The feed-forward input to C1 from the control pathway allows for modelling of the OHC's cochlear amplifier effect by regulating the gain and bandwidth of the C1 filter. The C2 filter is broadly tuned, highly resistant to trauma, and designed to dominate the response at high input pressures. After transduction from BM vibrations to IHC potentials, the C1 and C2 signals are summed and low-pass filtered to produce an IHC receptor potential. This potential drives an IHC auditory nerve receptor synapse model, leading to a spike generator producing a Poisson distribution of neuronal spiking events at the output for a specific characteristic frequency (CF). The reader is referred to Zilany and Bruce (2006) for a more detailed description of the cat auditory model used in this research.

Model parameters C_{IHC} and C_{OHC}, which control the level of inner and outer hair cell impairment, respectively, can be adjusted to provide a desired hearing threshold shift at a specific CF. A C_{IHC} or C_{OHC} of 0 produces full impairment whereas 1 produces normal function.

In Fig. 4.2 we see an example output from the Zilany and Bruce (2006) model. The synaptic output is presented to the neural spike generator at the output which produces spiking events based on a Poisson distribution. In 4.2(b) we see a marked increase in rate upon stimulus onset which diminishes as adaptation in the model kicks in. For a review of adaptation, please refer to Section 2.4.4.

4.1.1 Neurogram

Shown in Fig. 4.3A is an example speech sentence with a sampling rate of 16 kHz, taken from the TIMIT speech database. The speech sentence reads, in an American dialect, "They will take a wedding trip later". We can view the spectral components of this speech waveform over time by periodically calculating the Fourier power coefficients of a windowed segment of the speech sentence. In Fig. 4.3B, this was done using a 256 point Hamming window moving with 128 point overlap. Viewing a signal this way, as a function of time and frequency, is more commonly referred to as the spectrogram. In a similar manner, we can visualize the neural representation of speech in the auditory nerve by displaying the neural response as a function of CF and time as a plot called the "neurogram". A neurogram can include the spike timing

information of the neural responses by maintaining a small time bin size (10 µs), as seen in Fig. 4.3D. Alternatively, we can view the average discharge rate representation by computing the moving average with a window of up to several milliseconds. In Fig. 4.3C, the average discharge rate representation was calculated by binning the fine timing neurogram every 80 µs.

In later sections and in the next chapter, we calculate the difference between the normal and impaired neurogram. To avoid inaccuracies in taking the difference between high frequency components, we low-pass filtered each neural representation first with a 12th order, zero phase, Butterworth filter. Spike-timing neurograms were filtered with a 3dB cut-off at 6000 Hz. We filter the spike timing information to smooth the response, but we are careful to keep the cut-off high enough to avoid cutting out high frequency speech components that may be contained in the neural output. Average discharge rate representations, on the other hand, were filtered with a 3dB cut-off at 250 Hz. We considered such a low cut-off here because we are looking for slow varying temporal patterns in the rate response. Lastly, throughout this thesis, calculation of the spike-timing and average-discharge rate neurograms will remain the same unless specified otherwise.

Neurograms on their own are somewhat complex ways of visualizing the speech behaviour at the level of the auditory nerve. In the following section we will describe equations that will allow us to better analyze and visualize the information contained in a neurogram.

4.2 Representing Auditory Nerve Responses

4.2.1 Synchronized Rates and Box Plots

A synchronized rate calculation (Eqn. 4.1) is quite similar to a Fourier transform. Its purpose is to calculate the frequency components (as rates) that exist in the neural output from a neuron, or bundle of neurons, at a particular CF in response to speech. Synchronized rates display the effects of acoustic trauma at specific CF locations. This is because hearing damage generally exhibits broadened tuning curves at specific CF locations, making neurons highly susceptible to capturing frequency components other than their their own best tuned frequency.

Synchronized rates are calculated on steady portions of a neural response, where adaptation effects have stabilized and stationary characteristics remain (Miller et al., 1997). For the purposes of this thesis, we use an 80 ms Hamming window to isolate a stationary portion of the spike timing neural response. The spike timing neural response is first normalized to have units of spikes/sec by multiplying each spiking event by the sampling frequency. We use the Fourier transform in our calculation of synchronized rates as follows:

$$\mathcal{R}(kf_t) = \frac{\left| \sum_{n=0}^{N-1} w(n)p(n)e^{-j2\pi kn/N} \right|}{\sqrt{N \sum_{n=0}^{N-1} w(n)^2}} \tag{4.1}$$

where $\mathcal{R}(kf_t)$ are the synchronized rates at k frequencies and where f_t is the fundamental frequency of speech. The term in the denominator is a correction for the attenuation of the signal by the Hamming window. In this thesis, we avoid analysis of frequencies greater than 4 kHz because of loss of synchrony that exists beyond this range.

Synchronized rates represent a one-dimensional view of frequency components at a single CF. By extending the synchronized rate calculation to many CF locations, such as if we would like to apply the calculations to a neurogram, we develop a 2 dimensional structure called a "box-plot", allowing us to view the response of many CFs to a particular speech stimulus simultaneously. With a box-plot, we can gather information on how well rate information of a particular fiber compares relative to those of other center frequency fibers. This may be, for example, beneficial when identifying spread of synchrony.

In Fig. 4.4 we see a calculation of the box-plot using a 65dB SPL /ɛ/ vowel segment with a 100Hz fundamental frequency. Depending on the magnitude of the rate of response, synchronized rates are categorized into bins and are displayed as a 'box' marker of one of 8 sizes (see figure legend) at kf_t frequencies.

4.2.2 Power Ratio as a Measure of Synchrony

The degree of synchrony of an AN fiber to a particular frequency component in periodic speech is measured using the power ratio (Eqn. 4.2). The power ratio is defined as the sum of power in the AN Fourier response R at the frequency f_x and its harmonics, divided by the total power in the response:

$$\mathcal{PR}(f_x) = \frac{\sum_{m=1}^{u} R^2(m \cdot f_x)}{\sum_{n=1}^{v} R^2(n \cdot f_0)} \tag{4.2}$$

with $u < 4$, $m \cdot f_x \leq 5$kHz, $v = 50$, and $n \cdot f_0 \leq 5$kHz, where f_0 is the fundamental frequency (Miller et al., 1997). Because phase locking in cats is not observed above 5kHz, the summations are limited to frequency components below 5kHz. One important point to consider is that when using the power ratio, the frequency of interest, f_x, needs to be a multiple of f_0. This is to ensure that the total power in the denominator includes the power in the numerator as well. As we will see in experiments in the next chapter, this becomes a particular point of trouble when the fundamental

frequency of speech is not a multiple of the formant frequency. In this case, we opt to avoid using the power ratio.

In Fig. 4.5 we see the power ratio response to over 160 center frequencies, spaced logarithmically along the basilar membrane in a fully functioning (green) and impaired (red) cochlear. Power ratios are given for formants F1, F2, and F3, respectively in response to a 65 db SPL synthetic /ε/ vowel presentation. Comparing the subplots, we see the first formant typically contains the most energy in speech, followed by the second and third formants.

Around low to moderate SPLs in a normal cochlea, the neural response to speech at a particular CF fiber shows great specificity, or synchrony, around similar frequencies in speech. Toward high SPLs, synchronized responses to formants tend to dominate the neural response over a broad region of higher CF fibers, and not just the CF at the formant location. This phenomenon is known as the upward spread of synchrony and was first observed by Young and Sachs (1979) in studies of auditory nerve responses. Spread of synchrony is easily explained by the tuning curve seen in Fig. 2.11A. As the sound pressure level increases, two things happen. First, with higher sound pressure levels, the AN tuning curves tend to broaden, and second, the formants which already contain the majority of speech energy end up surpassing the threshold limit at other CF locations, thereby causing CF to respond to frequencies other than the fiber's center frequency. The effect is seen as a widening in the power ratio plot.

Hearing damage too tends to broaden the power ratio plots, primarily because hearing damage causes a shallowing and broadening of CF tuning curves (refer to Fig. 2.13A). Broadening of the tuning curves means that CF fibers are less selective to their most sensitive (best) frequency component. Figure 4.5 demonstrates this effect in the case of mild hearing loss. Upward spread of synchrony becomes quite apparent particularly at the third formant frequency.

Once again, broadening of the power ratio is a direct result of a broadening and shallowing of tuning curves along the basilar membrane. Trying to restore normal response in auditory nerve fibers at a particular CF so that the same frequency components of speech dominate the neural response could cause other fibers with broadened tuning curves to fire in response to that amplified component rather than at their own best frequency. Both synchrony capture and the spread of synchrony reflect non-linear processes of the cochlea that complicates the matter of hearing restoration.

4.2.3 Phase Response

Temporal coding theory proposed by Carney (1994) hypothesizes that auditory nuclei in the brain uses phase information in the AN neural output to determine speech perception and loudness. The degree of loudness is thought to correspond to

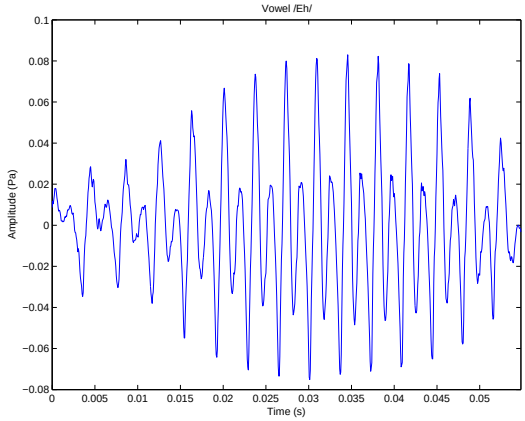
how much the phase at a certain frequency in the neural output changes across CFs. Studies by Carney (1994) have shown that as speech SPLs increase, and, similarly, as hearing impairment worsens, phase to a particular frequency across CFs tends to become more synchronized (see Fig. 4.6).

Phase response is calculated relative to the phase of the fiber with CF equal to the frequency of interest and is only measured for adjacent fibers with power ratios greater than 0.1. For the purposes of this thesis, power ratios less than 0.1 are deemed insignificant and are ignored. With the appropriate selection of nerve fibers around CF, the Fourier transform of neural output is taken and the phase at the particular frequency of interest is measured. The phase response is plotted on a phase-CF axis and, generally, tends to be quite linear crossing the relative zero-phase mark at the CF location.

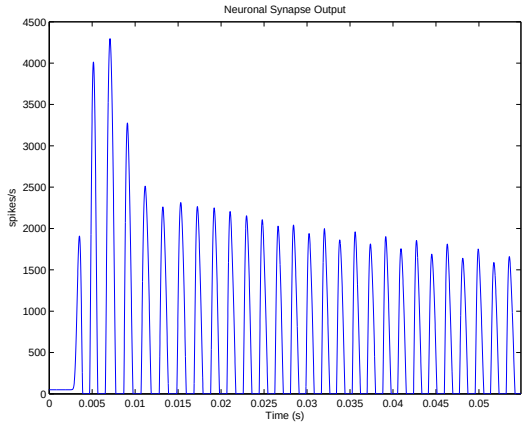
4.2.4 Histogram of Average Discharge Rates

So far we have discussed two important plots for analyzing rate patterns. Those are box-plots, which look at the rates on a CF vs frequency axis and the average discharge rate neurogram, which looks at rates distributed temporally over the course of the speech stimuli. A third method of analyzing rate patterns is by viewing the histogram of the average discharge rate neurogram. In doing so, we exclude the time domain and show rates as they are distributed by relative number alone. The idea here is to collapse rate response into distribution patterns that are easy to compare.

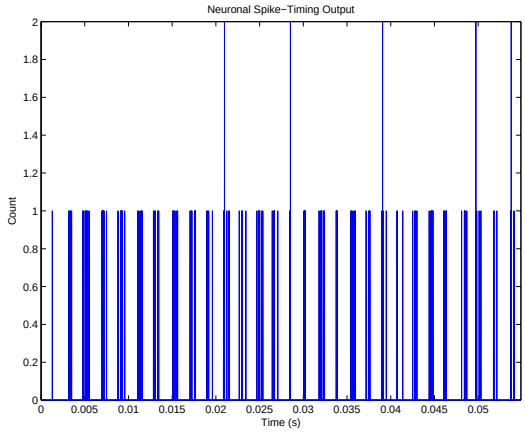
In Fig. 4.7, we bin the rates in the discharge rate responses, up to and including 400 spikes/sec, into bins varying in steps of 5. We visualize the result on a 2 dimensional plot of rate count vs CF. Figure 4.7(a) shows the histogram response to an /ɛ/ phone presented at low to moderate SPL. With greater impairment of the auditory periphery, the distinctive features in the count distribution tend to fall, producing a uniform rate distribution across CFs in the impaired region (Ref. Fig. 4.7(b)). This abnormal distribution is brought closer to normal with amplification prescriptions (Ref. Fig. 4.7(c) and 4.7(d)).



(a)



(b)



(c)

Figure 4.2: Sample /ε/ phone in (a) is presented to the Zilany and Bruce (2006) model and the synaptic output from a nerve fiber of CF = 542 Hz is seen in (b). The action potential events from a bundle of 50 identical neurons is shown in (c).

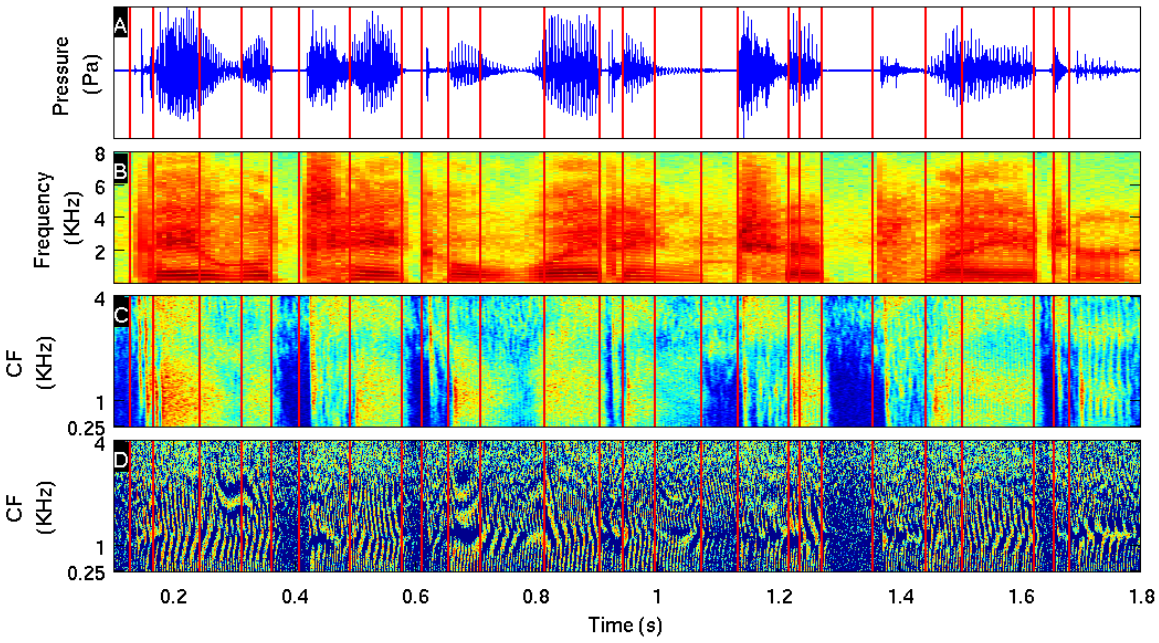


Figure 4.3: An example sentence from the TIMIT database and the corresponding spectrogram and neurograms. (A) Time-domain pressure waveform; (B) Spectrogram; (C) Neurogram based on the average discharge rate; (D) Neurogram based on the spiking timing information. Phoneme boundaries are indicated by the vertical red lines. Sub-figures B and D are scaled by log10 for better contrast.

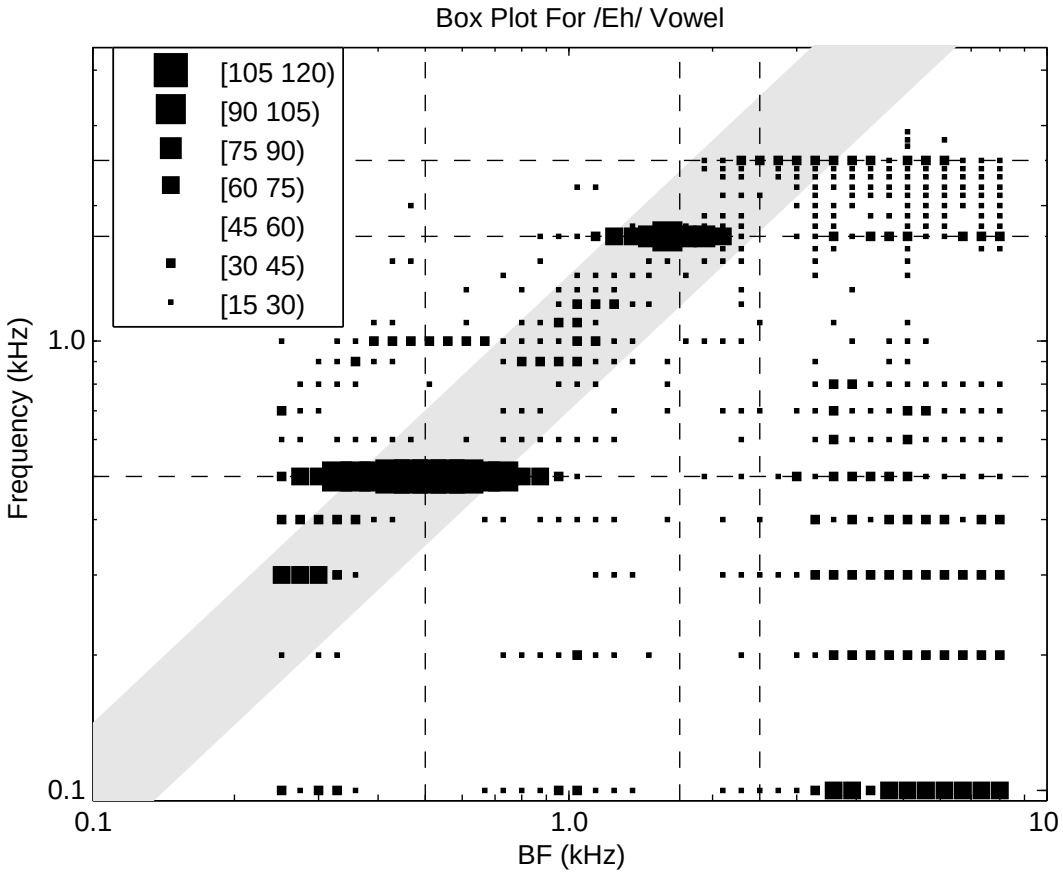
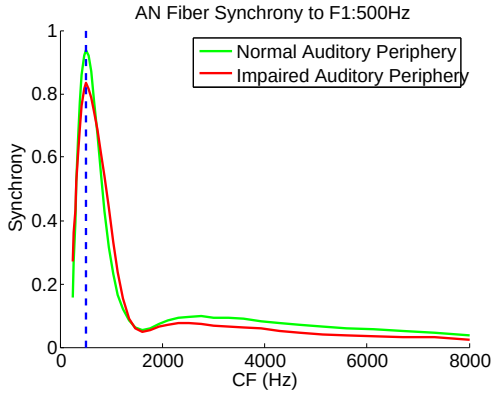
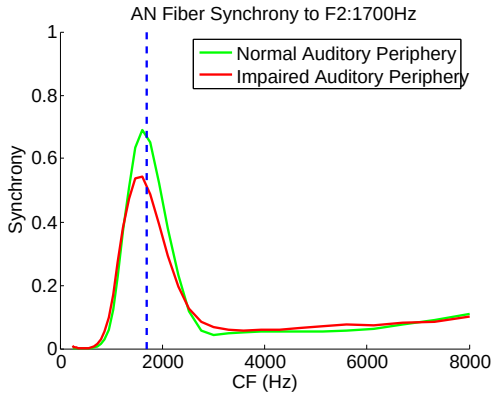


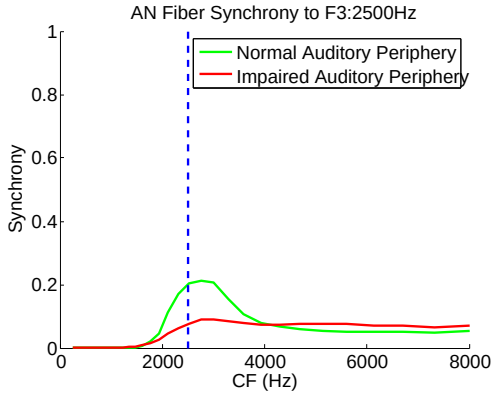
Figure 4.4: Box plot of model AN response to a synthetic /ɛ/ vowel. The grey diagonal region highlights the fiber CFs which approximately corresponds to the same fourier frequency and the dashed lines indicate formant frequency locations of this vowel.



(a)



(b)



(c)

Figure 4.5: Synthetic /ɛ/ vowel: AN synchrony to the first formant of 500 Hz is shown in (a), second formant of 1700 Hz in (b), and the third formant of 2500 Hz in (c). Results were simulated using the Zilany and Bruce (2006) AP model. Mild hearing loss was used in simulating the impaired auditory periphery.

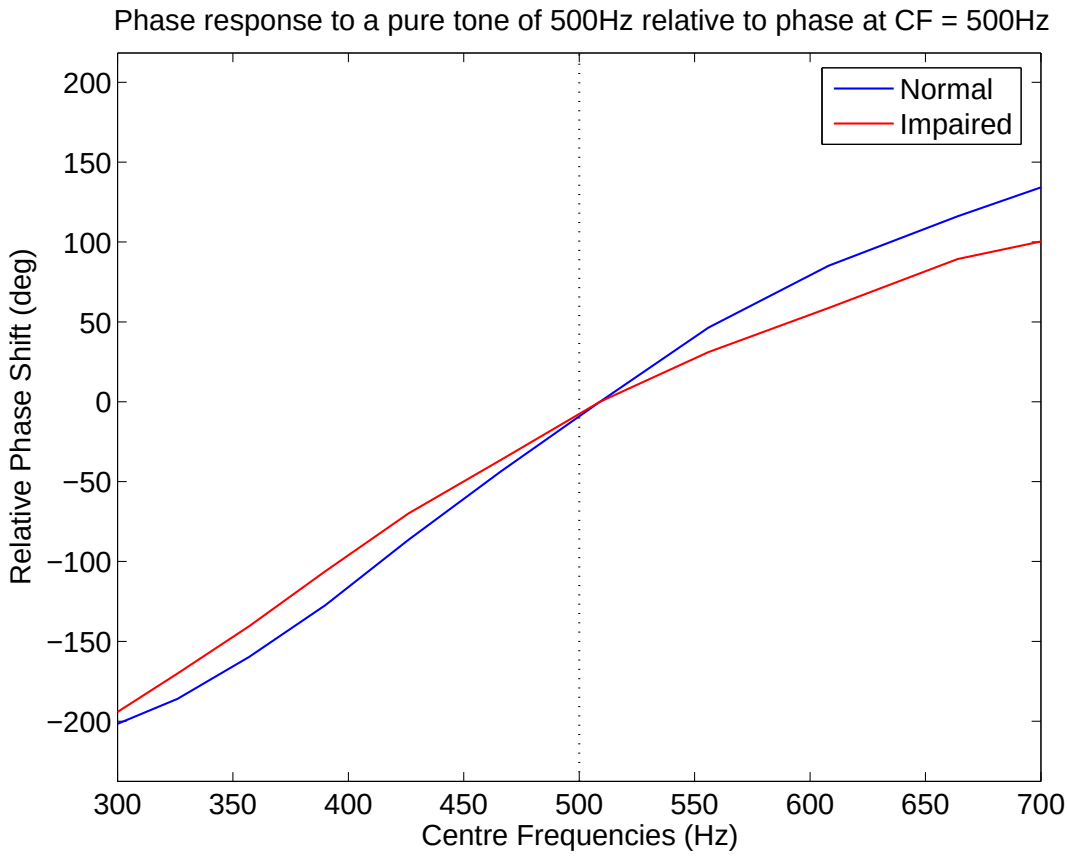


Figure 4.6: Phase Response to a 500Hz tone, relative to CF = 500Hz. Normal auditory periphery response shown in blue and mildly impaired response in red.

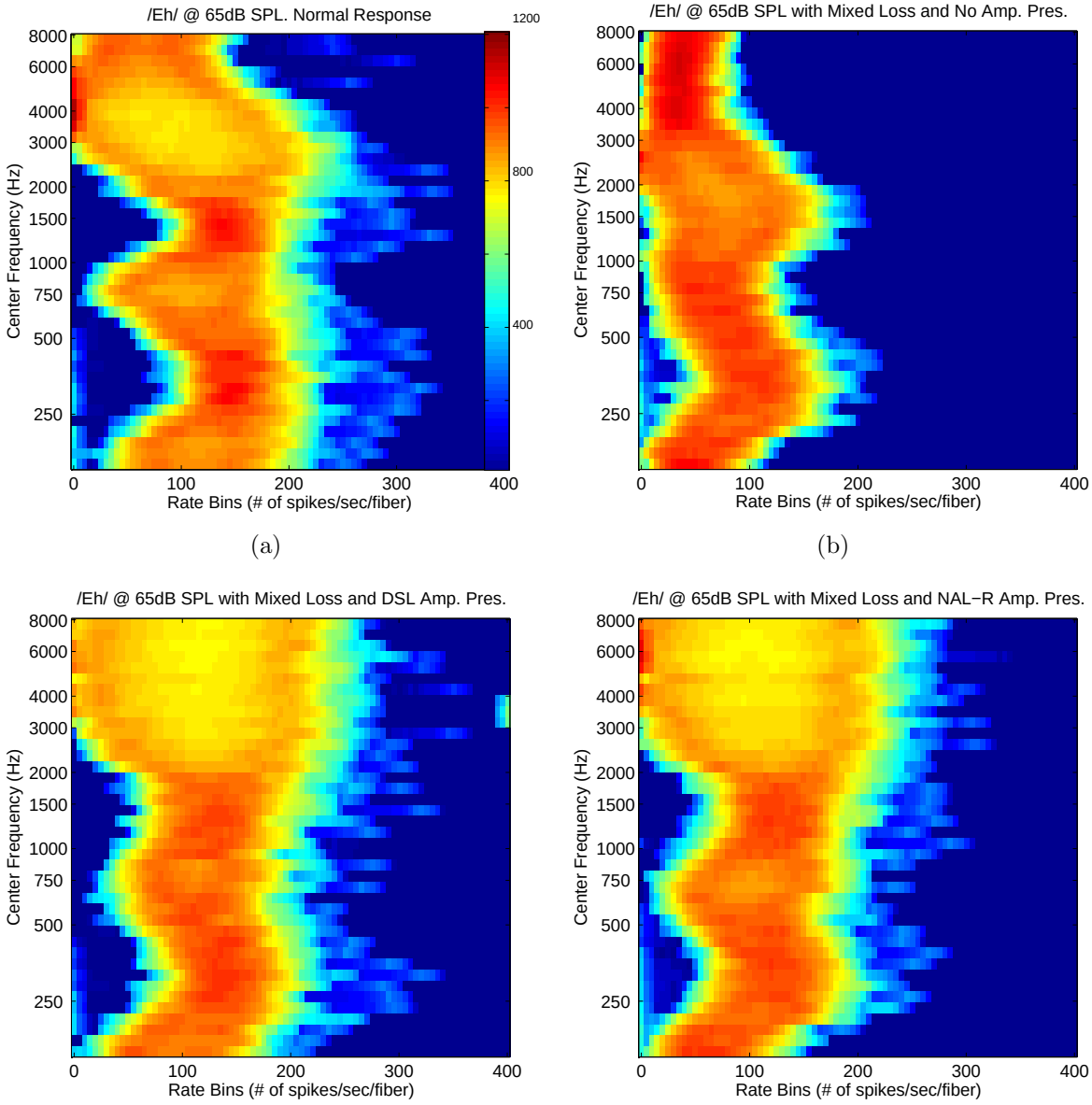


Figure 4.7: The figures above were generated in response to a 65dB SPL /ε/ vowel segment with a 100Hz fundamental frequency. The histograms are composed of 80 equally spaced bins of width 5 spike/s. Each histogram represents the distribution of bins with the colourbar in (a) indicating the magnitude of each bin. The data in each plot is scaled by log base 10 for enhanced contrast, but the colour bar is shown linearly.

Chapter 5

Neurophysiological insights into hearing aid amplification schemes

5.1 Motivation

In Section 2.5 we touched on two ways in which we may represent the neural representation of speech; either as a mean-rate representation or as a spike-timing representation. Here we use the tools described in the previous chapters to find optimal single-band compression gain adjustments around the NAL-R and DSL prescribed gains using the mean-rate and spike-timing representation of speech. Our objective is to determine whether and how hearing aid gain prescriptions optimize the restoration of mean-rate and spike-timing representation of speech.

Our objective is carried out through a series of experiments. In the first, we use spoken speech sentences from which we try to optimize compression gains for each phone in order to ideally restore the impaired speech neurogram to normal. Next we use either synthetic or spoken vowels. Synthetic vowels are ideal with known characteristics such as the vowel's fundamental frequency and formant frequencies, whereas when using spoken vowels, these characteristics had to be estimated. Similar to the first part of the study, we adjust compression gains for each phone in order to ideally restore the impaired neurogram to normal. The auditory periphery model in both speech sentence and vowel studies is assumed to have mixed hearing loss. Our last study examines optimal compression gains using a synthetic vowel in which hearing loss, rather than being mixed, is due solely to inner or outer hair cell impairment. Before moving on to each of the said studies, we will first briefly describe the set-up of the auditory model used throughout each.

5.2 Model Set-up

The auditory periphery model used throughout studies in this chapter consisted of 40 characteristic frequencies (CFs) spaced logarithmically between 250 and 8000 Hz. We consider the region between 250 and 8000 Hz because it is assumed that most of the frequency components for speech intelligibility reside here (Gelfand, 1998). At each CF fiber, the simulated neural response was composed of 50 AN fiber responses. We chose to use 50 fibers as that gives an appropriate sampling of auditory nerve fibers at and around the CF region. As well, in accordance to Liberman and Kiang (1978) on the makeup of the cat auditory periphery, 60% of fibers were chosen to be high spontaneous rate (>18 spikes/s), 20% medium (0.5 to 18 spikes/s), and 20% low (<0.5 spikes/s).

The impaired auditory periphery model used in this thesis was characterized as having either mild or severe hearing loss audiogram patterns (see Fig. 5.1). Hearing loss patterns can be caused by a combination of IHC and OHC impairment to produce the desired threshold shift. Until we begin the last study, it is assumed that roughly 1/3 of the hearing threshold is a result of inner hair cell impairment and 2/3 a result of outer hair cell impairment. For more information on how IHC and OHC impairments contribute to a desired threshold shift, the reader is referred to Zilany and Bruce (2007).

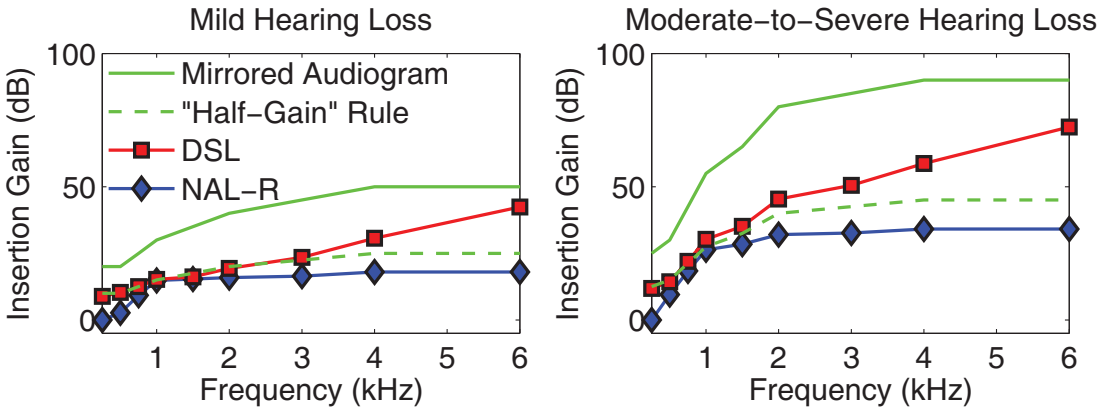


Figure 5.1: Mild and Severe hearing loss patterns. NAL-R and DSL prescribed gains are presented here as insertion gains.

5.3 Stimuli

Input to the auditory periphery model consisted of either speech recordings, voiced vowels, or synthetic vowels. Speech recordings were taken from the TIMIT corpus of

prompted utterances. The TIMIT corpus consists of 450 phonetically-compact and 1890 phonetically-diverse English-read speech sentences in a wide variety of American dialects.

In place of whole sentences, some experiments made use of either voiced or synthetic periodic phones. Synthetic phones were ideal /ε/ phones of known characteristics: having a fundamental frequency of 100 Hz and formants at exactly 500, 1700 and 2500 Hz. On the other hand, characteristics of voiced phones were slightly more difficulty to determine. Voiced phones were stable long periodic phones segments selected from TIMIT sentences. Fundamental frequencies were determined using cross correlation methods and formants were determined using linear predictive coding (refer to code in Appendix B-8 and B-9) With long repetitive phones, these methods are quite accurate in estimating fundamental and formant characteristics.

To understand how linear predictive coding (LPC) determines formant frequencies, we refer to speech production in Section 2.2. We stated that the voice box acts as an impulse generator that inputs a series of puffs into the vocal tract. Furthermore, the vocal tract itself acts as a filter with an impulse response that is quite dynamic, that is, by moving our tongue and constricting or relaxing the muscles of our vocal tract, we can change the impulse response of our vocal tract to shape the sounds we wish to produce. In periodic speech, periodic input from the vocal cords into the vocal tract 'samples' the Fourier spectrum of the vocal tract impulse response. Using LPC, we can determine from a speech recording the approximate vocal tract filter coefficients and, therefore, the formant frequencies. LPC essentially states that the current speech sample, $x[n]$, can be approximated as a linear combination of 'P' previous samples according to:

$$x[n] = \sum_{k=1}^{P} a_k \cdot x[n-k] + e[n] \tag{5.1}$$

where P is the prediction order, a_k is the prediction coefficient, and $e[n]$ is the prediction error for the current sample $x[n]$. We find the most suitable prediction coefficients by minimizing the prediction error:

$$\min(e[n])^2 = \min(x[n] - \sum_{k=1}^{P} a_k \cdot x[n-k])^2 \tag{5.2}$$

Ideally, LPC should be able to decompose the vowel into the voice box impulses, as the prediction error, and the vocal tract filter coefficients, a_k, however errors exist in approximating the order of coefficients as well as inherent errors associated with real-world recordings.

Another problem pertaining to vowel formation, however, is that resolving the formants of a particular voiced vowel segment may not necessarily be a multiple of

the fundamental frequency. Thus, when the vocal tract spectrum is being 'sampled' at the fundamental frequency, the main peak of the formant may appear between 'samples' and therefore may not be entirely appropriate when working with the power ratio response. For this reason we make extensive use of the synthetic /ɛ/ vowel as it is designed to provide formants at exact multiples of the fundamental frequency.

5.4 Optimal Gains Using Sentences From the TIMIT Database

5.4.1 Gain Optimization Strategy

We determine optimal single-band gain adjustments around the hearing aid prescription gains using the gain optimization strategy shown in Fig. 5.2. The gain optimization strategy compares neural responses of spoken sentences on a phoneme-by-phoneme basis for the impaired and normal models. In order to avoid the confounding and complicating effects of compression attack and release times, a constant gain adjustment was applied for the duration of each phone, using the known phone boundaries from the TIMIT transcriptions.

Before applying the gain optimization strategy, each sentence was normalized to either 45, 65, or 85 dB SPL before being presented to the model. We thereby ensure good consistency and sound pressure level coverage for the gain optimization strategy detailed next.

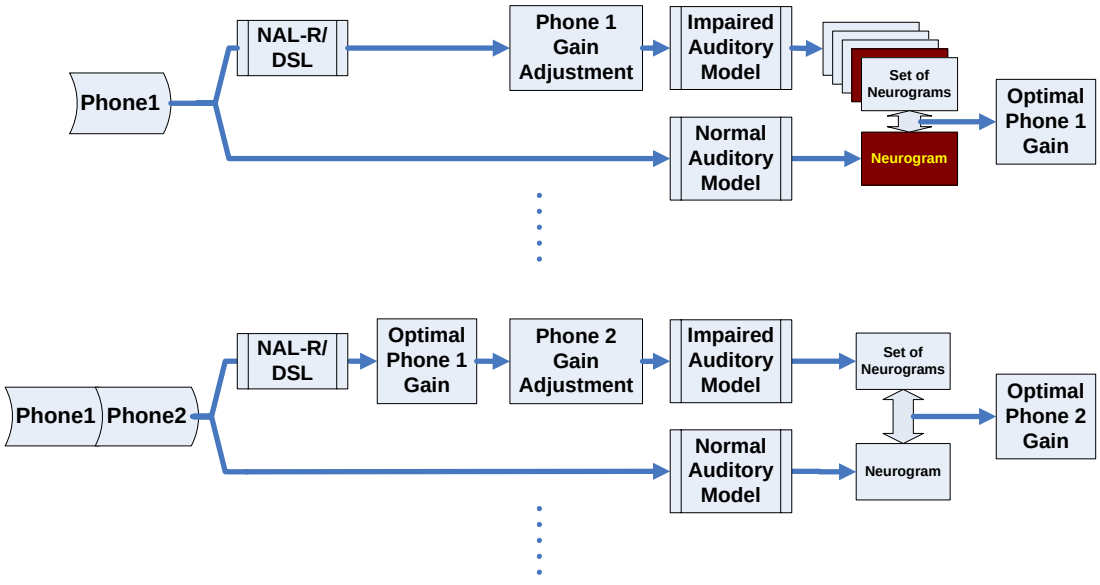


Figure 5.2: Flow diagram of gain adjustment strategy.

The strategy begins by passing the first phone through the normal model to derive the normal neurogram. In the impaired pathway, the phone is passed though either the NAL-R or DSL amplification prescription before a single-band gain adjustment is applied. Gain adjustments range from -40 to $+40$ dB in -5 dB increments resulting in 17 uniquely amplified phones. The phones are passed through the impaired model, producing a set of 17 neurograms. The gain adjustment that minimizes the mean absolute error between the normal and impaired neurograms is deemed the optimal gain adjustment for that phone. For each amplification prescription, optimal gain adjustments were found by comparing either the neurograms with spike timing information or the average discharge rate neurograms.

The second and all subsequent phones are analyzed in the same manner as the first, however, due to adaptation in the auditory-periphery model, all prior phones are prepended. The range of gain adjustments is applied only to the current phone and all previous phones are amplified with their optimal gain adjustments.

5.4.2 Mean Absolute Error Metric

We assume that the brain interprets components in the neurogram with equal weighting, such that large values in the neurogram contribute as equal significance as that of smaller values. When determining an error metric, we settle for the mean absolute error metric which determines an error value by weighting all differences between normal and impaired auditory responses equally. The mean absolute error metric (Eqn. 5.3) works by taking the absolute difference between two data sets and calculating the mean value. When two data-sets are in close relationship the error produced by the mean absolute error metric approaches 0. Thus when determining the optimal compression gains, we seek to minimize the mean absolute error between the normal and impaired neural responses.

$$\text{MAE} = \frac{1}{n} \cdot \sum^{n} |x - y| \tag{5.3}$$

The minimum mean absolute error is versatile and can be used to compare neurograms, power ratios and phase responses (at each formant frequency), box plot, and histogram plots between the normal and impaired cases. An example error curve using the mean absolute error metric shown in Fig. 5.3. The error curve was produced by comparing the normal vs impaired neurogram from an /ɛ/ vowel at 65 dB SPL as input to a mildly impaired auditory periphery. DSL prescription and compression gain adjustments were subsequently applied in the impaired case. Note the appearance of a local minimum at approximately -20 dB

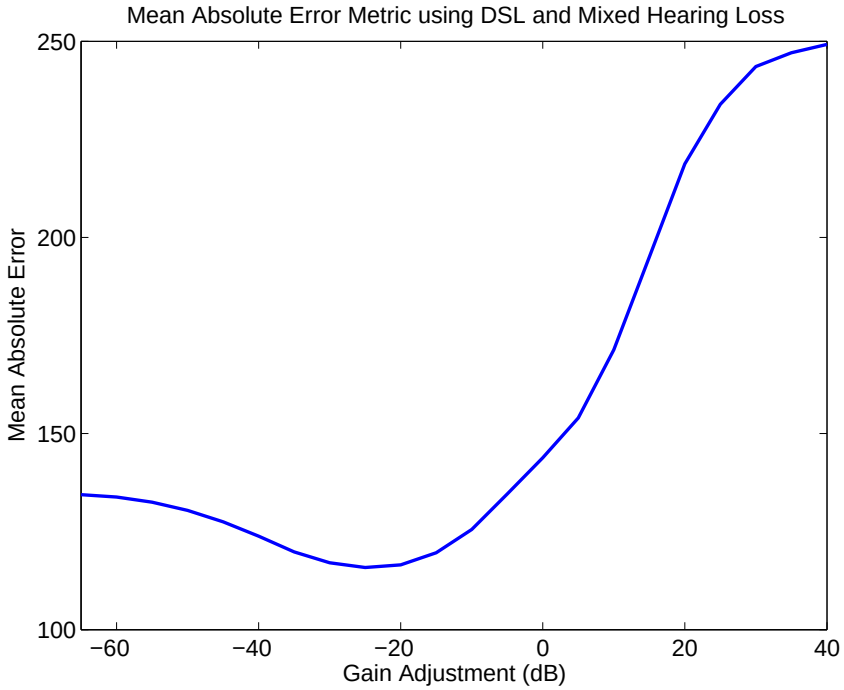


Figure 5.3: Mean absolute error metric showing errors between the normal neurogram produced in response to an /ɛ/ vowel at 65 dB SPl and the impaired neurogram to the same vowel with additional compression gain adjustments.

5.4.3 Optimal Gain Adjustments For TIMIT Sentences

The results from the gain optimization strategy[1] are shown in Fig. 5.4 and were generated using test sentences from the TIMIT database. Two sentences were presented to the auditory model with mild hearing impairment and another two were presented to a model with moderate-to-severe hearing impairment. Again, each sentence was delivered to the gain optimization strategy at 3 different mean sound pressure levels, thereby providing a number of diverse phone types and sound pressure levels for examination. The curves in Fig. 5.4 show the data points collected using the mean absolute error metric between normal and impaired neurograms.

The results from the gain optimization strategy suggest that positive gain adjustments above the prescribed gains better restore the mean discharge rate representation of speech. However for the fine timing neurogram, the gain optimization strategy suggests negative gain adjustments. Interestingly, NAL-R seems to be more suited for the spike timing neurogram, having optimal gains closer to zero than those for DSL, whereas DSL seems to be more suited for the average discharge rate neurogram. Presumably because of higher amplification gains in DSL, results here suggest that DSL may be weighted more towards optimizing the average discharge rate neurogram and NAL-R more towards the spike timing neurogram. We will analyze this conclusion in more depth in sections to come. For now, the discrepancy between positive and negative optimal gain adjustments between the two neural representations is not clear and warrants further investigation. In particular, it is difficult from the neurogram output alone to determine in exactly what ways gain adjustments affect the neurogram.

5.5 Optimal Gain Adjustments Using the Synthetic /ɛ/ Vowel

Using TIMIT speech sentences, we deduced that negative gain adjustments bring neurograms back closer to normal for neurograms containing fine timing information, and positive gain adjustments for neurograms containing the average discharge rate. In the gain adjustment strategy in this section, we use a synthetic vowel to see if the same optimal gain adjustments hold when comparing the power ratio, phase response, and box plots at formant frequencies. We use a similar strategy as the gain adjustment strategy described in the previous section, however not on a phoneme by phoneme

[1]For this section, calculation of neurograms were slightly different from what was described in Section 4.1.1. Fine timing neurogram was calculated using a 2 µs binwidth followed low-passing filtering via a 128 point Hamming window. The 3dB cut-off was approximately 5 kHz. The Average-discharge rate neurogram was calculated using a 62.5 µs binwidth followed by low-passing filtering via a 128 point Hamming window. The 3dB cut-off in this case was approximately 150 Hz.

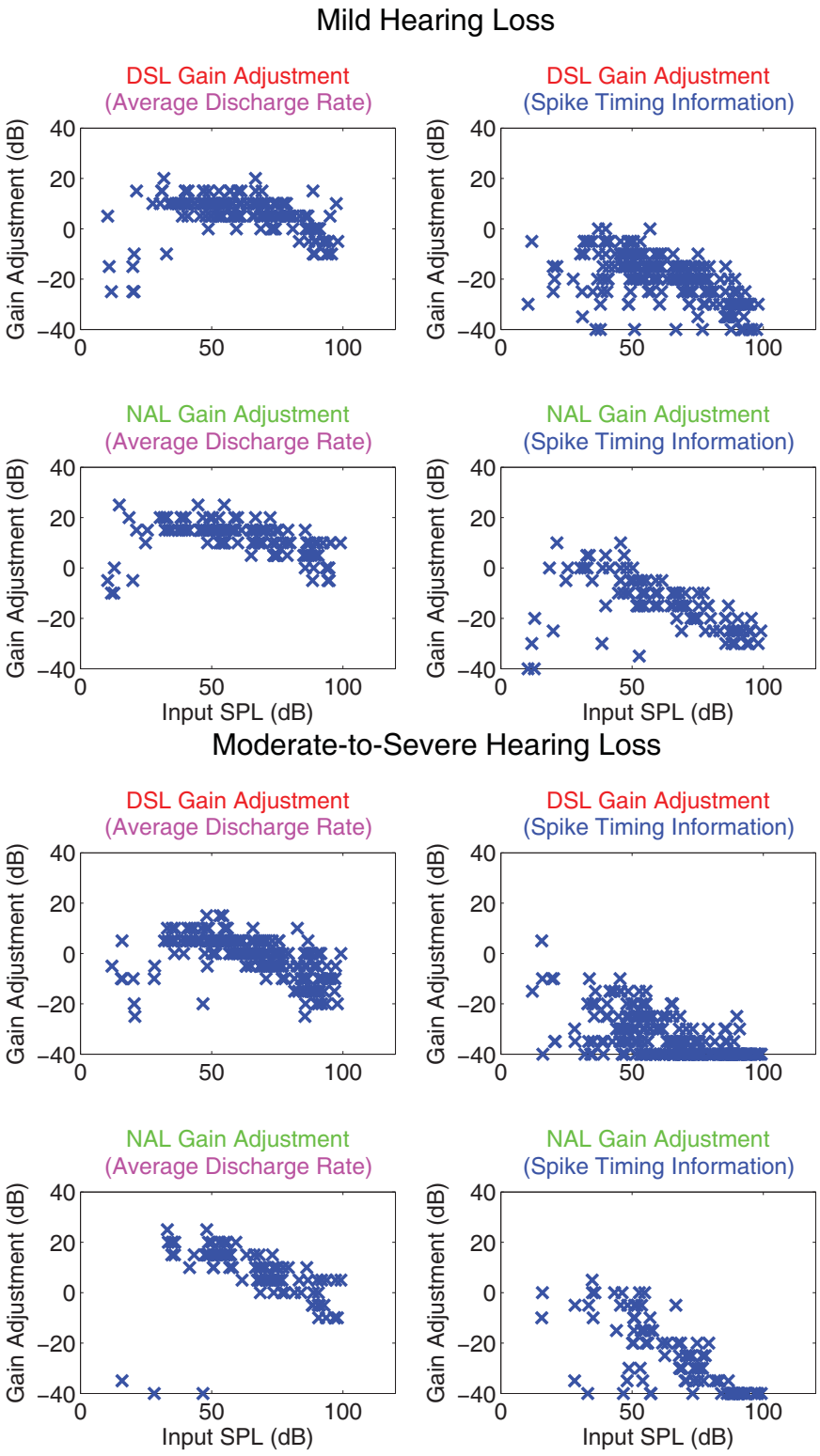


Figure 5.4: Raw data points from the gain optimization strategy

basis. The strategy here uses a test sentence containing 40 identical synthetic /ɛ/ phones over a duration of 400 ms with an SPL between 40 and 100 dB. This stimulus was passed through either a normal auditory periphery model or a mildly or severely impaired auditory periphery model. In the mildly impaired case, the sentence was processed using either the NAL-R or the DSL amplification schemes with overall compression gain adjustments from 40 dB below to 40 dB above the prescribed gains in 5 dB steps. In the severe case, gain adjustments ranged from −60 dB to 20 dB. Neural responses were obtained using the cat auditory-periphery model of Zilany and Bruce (2006, 2007), and at each gain adjustment, mean absolute errors between the normal and impaired neurograms, power ratios and phase responses (at each formant frequency), box plots, and histogram analysis were recorded. The algorithm repeats these steps over a 40 to 100 dB SPL range of input presentation levels, increasing in steps of 5dB SPL.

5.5.1 Synthetic Vowel Optimization Results

Optimal gain adjustments for neurograms containing average discharge rate information is shown in Fig. 5.5 and were found by minimizing the mean absolute error between the normal and impaired histogram interpretation. In compensating for mild and severe hearing impairment, the DSL and NAL-R algorithms are fairly effective at restoring the impaired rate distribution count to normal. The results are consistent with those for the TIMIT sentences; there seems to be a need for a slightly positive gain adjustment, particularly when dealing with the NAL-R algorithm since it gives less overall gain than DSL.

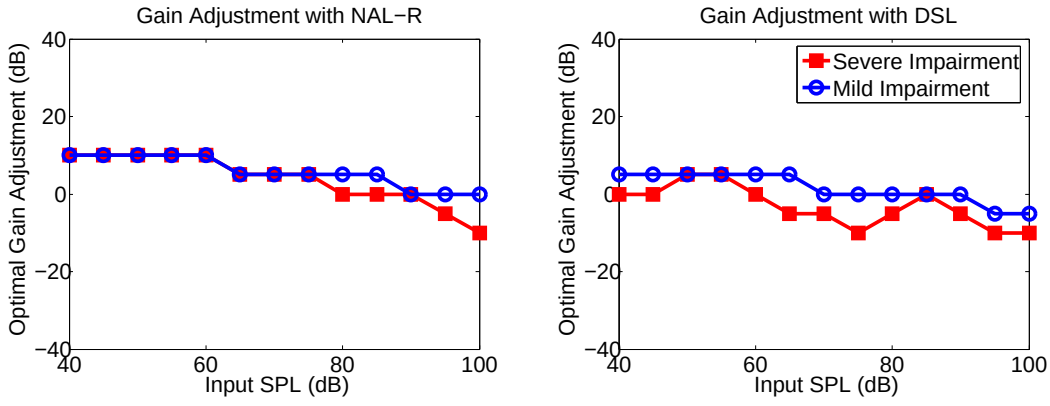


Figure 5.5: Optimal gains using the average discharge rate neurogram and a synthetic /ɛ/ vowel. NAL-R amplification scheme is shown in (a) and DSL shown in (b).

Optimal gain adjustments to restore neurograms containing fine-timing information show, for both mild and severe hearing impairment, that negative compression gain adjustments are required for linear hearing aid amplification schemes (Fig. 5.6). This is precisely what the gain optimization strategy for the TIMIT database suggests (Fig. 5.4). The reason for this requirement is that individual power ratios and phase responses (see Appendix C) of the impaired auditory periphery broaden and flatten with respect to the normal auditory periphery with increasing gain adjustments, a phenomenon known as spread of synchrony Miller et al. (1997). With higher gains, phase deviations, in addition to spread of synchrony, become a driving factor in the error metric, leading to the requirement of negative gain adjustments.

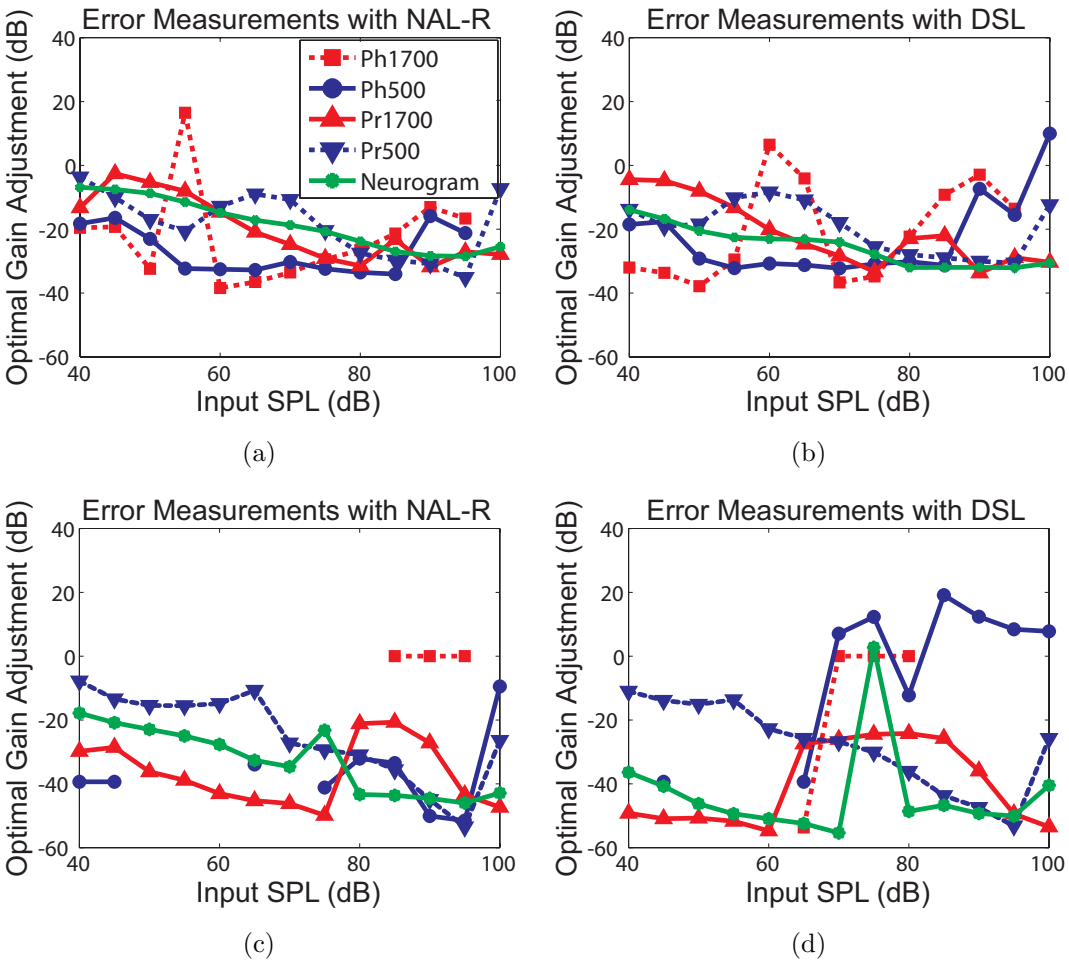


Figure 5.6: Fine timing neurogram: optimal gains using a synthetic vowel /ɛ/ in the case of mild, (a) and (b), and severe, (c) and (d), hearing loss. Ph500 and Ph1700 show the optimal gain adjustment for the impaired phase response at $f = 500$ and $f = 1700$Hz (note that some data points are missing since the power ratio at these points is < 0.1 through the range of gain adjustments). Likewise, Pr500 and Pr1700 represent the optimal gain adjustments for power ratios measured at $f = 500$ and $f = 1700$Hz.

5.6 Optimal Gain Adjustments Using Spoken Vowels

The gain analysis using the synthetic /ɛ/ vowel confirmed the results from the gain optimization strategy in section 5.4 which used sentences from the TIMIT speech corpus. Since the synthetic /ɛ/ vowel was manufactured with ideal fundamental and formant frequencies, for completeness, we would like to verify that we obtain the same results with spoken vowel segments. We apply the same gain adjustment strategy as we did to the synthetic /ɛ/ vowel, however using spoken vowel segments from the TIMIT speech database. Vowels were chosen to be /ɛ/ vowels and of good length and periodicity. Since we are dealing with spoken speech, we need to approximate formants based on the fundamental frequency were we to use either the power ratio or phase analysis. As a consequence, we have opted not to use the power ratio and phase analysis in this section, however this information should be encoded in the neurogram and box plots. Once again, optimal gains were determined by minimizing the mean absolute error between computed spike-timing and mean discharge rate neural information.

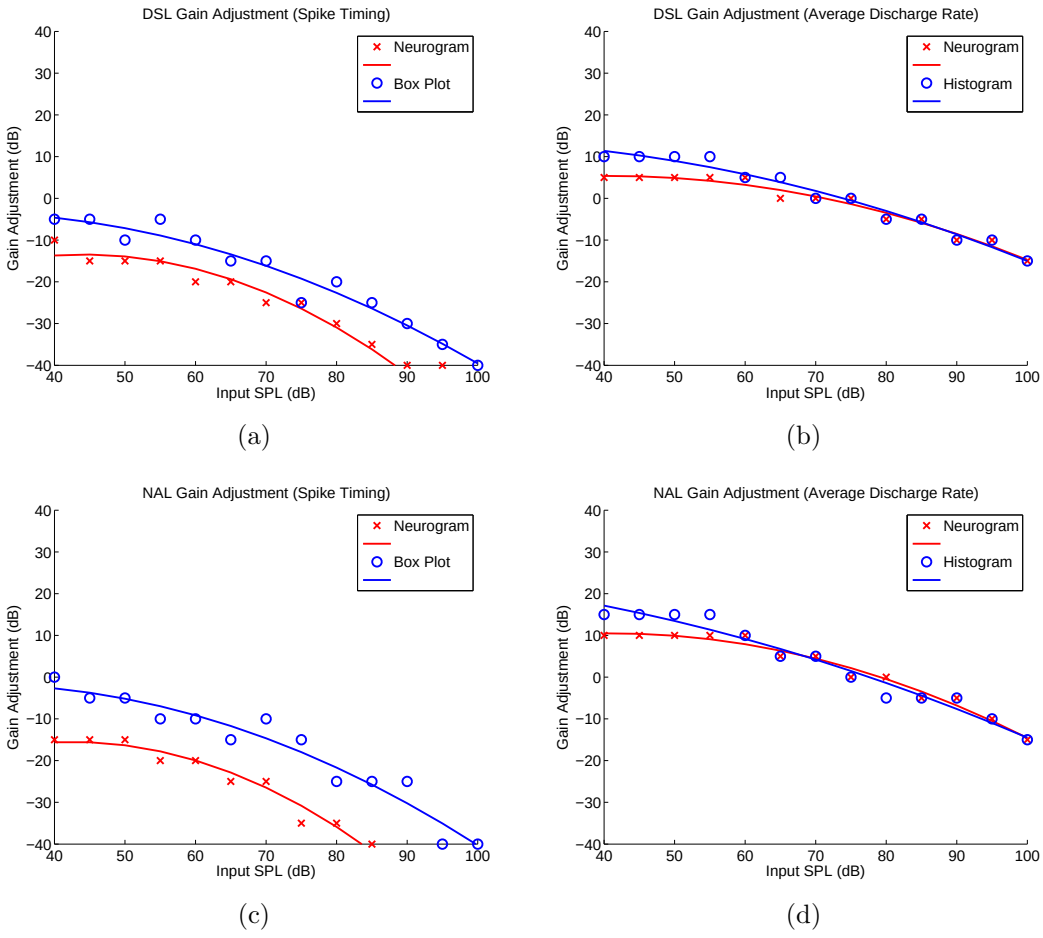


Figure 5.7: Optimal gains for a spoken /ɛ/ phone from the TIMIT database. Figures for optimal gains based on the fine-timing neurogram are shown on the left and those based on the average discharge rate are shown on the right. The DSL amplification prescription is applied in subfigures (a) and (b) and NAL-R in subfigures (c) and (d).

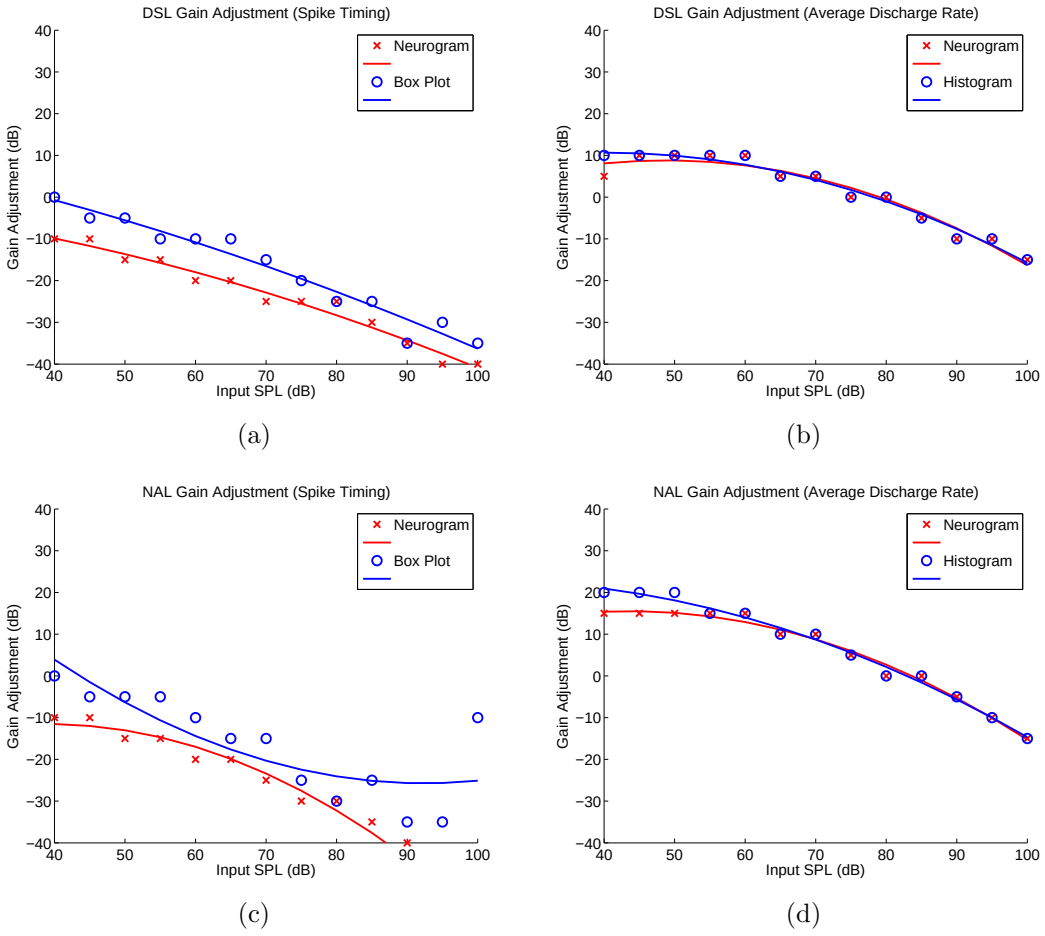


Figure 5.8: Optimal gains for a spoken /ε/ phone from the TIMIT database.

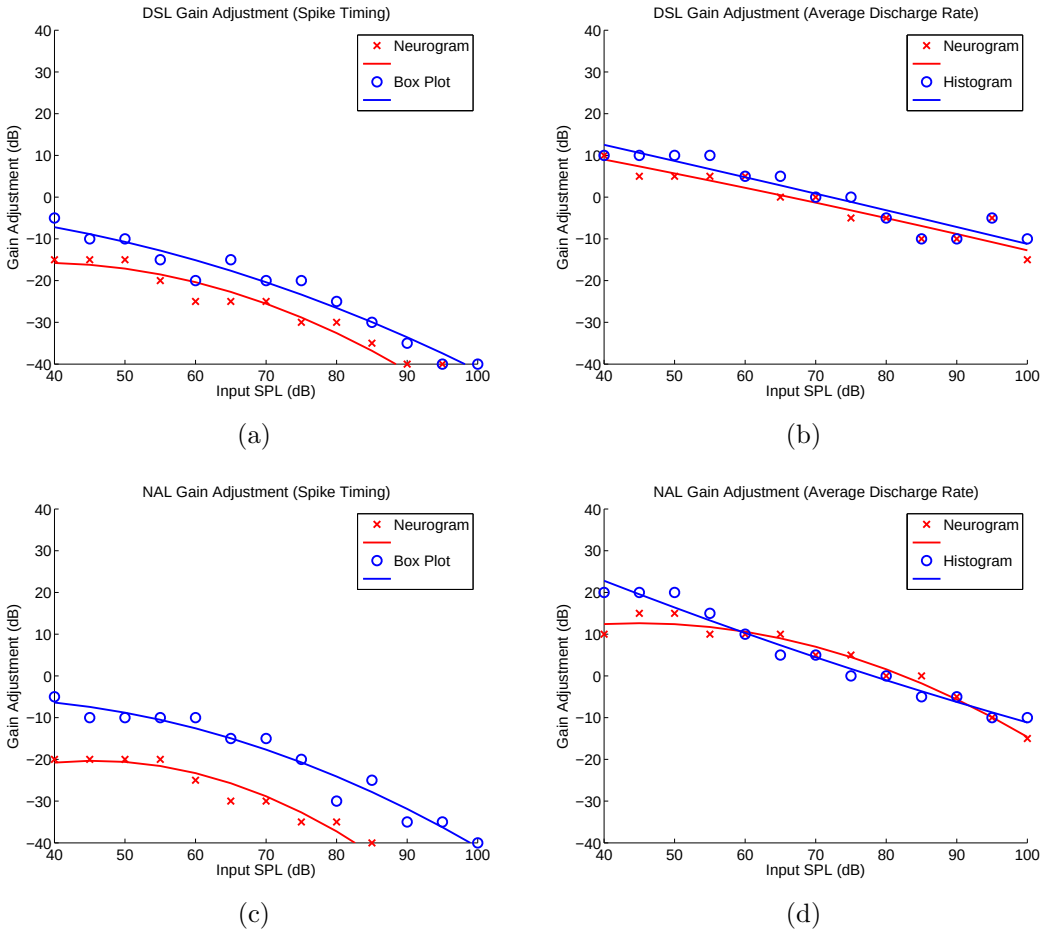


Figure 5.9: Optimal gains for a spoken /ɛ/ phone from the TIMIT database.

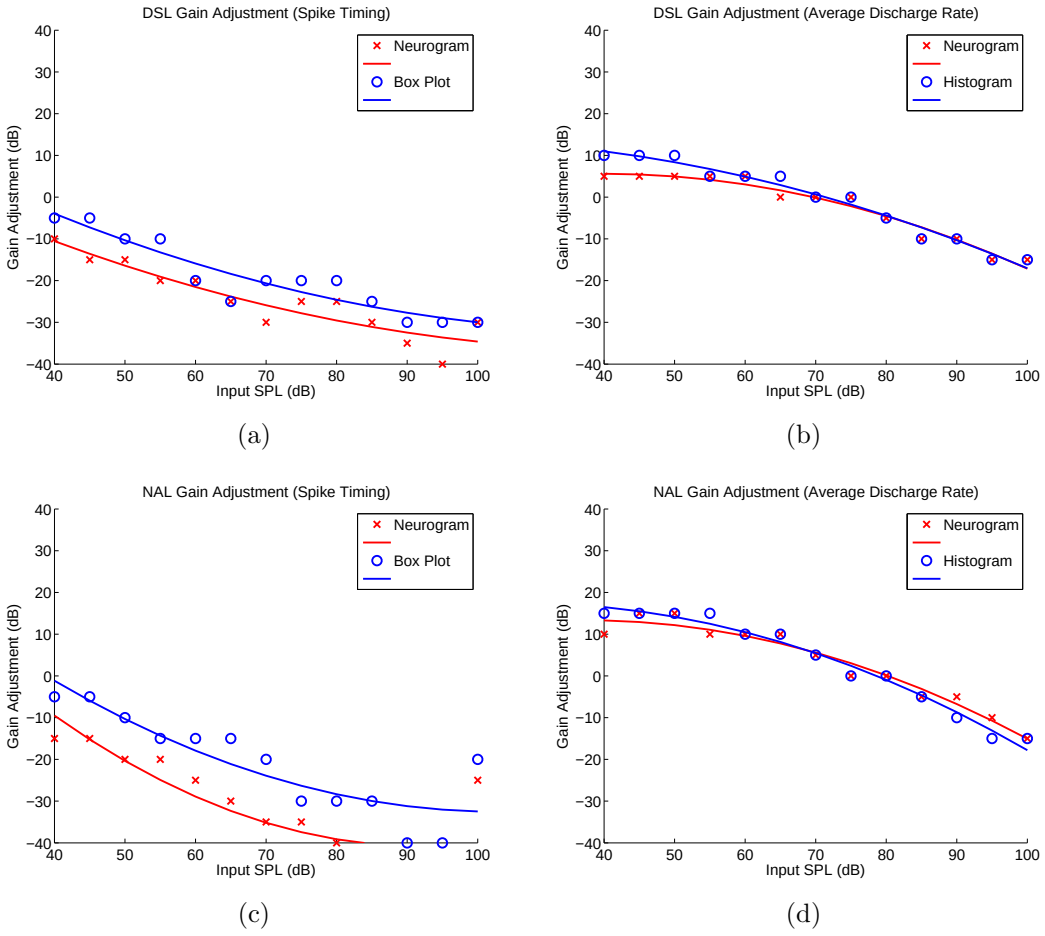


Figure 5.10: Optimal gains for a spoken /ɛ/ phone from the TIMIT database.

Figures 5.7, 5.8, 5.9, and 5.10 confirm the results from the gain optimization strategies using TIMIT speech sentences and synthetic /ε/ vowel; for neurograms containing spike timing information, the box plots and the neurograms themselves indicate that, on the whole, lower gain adjustments are necessary to optimally restore the neural representation of speech. This is in contrast to the average discharge rate responses that indicate greater gains are necessary to restore the neurograms. Negative slopes in the gain adjustments indicate increasing compression is needed, possibly due to properties of IHC and OHC impairment as we approach higher SPL. In the next section we investigate optimal compression gains when hearing thresholds increase as a result of purely inner and outer hair cell impairment.

5.7 Gain Adjustments Using Solely IHC or OHC Impairment.

In experiments untill now, hearing loss was predefined as a mixture of inner and outer hair cell impairment, in that 1/3 of the hearing threshold shift resulted from inner hair cell impairment and 2/3 from outer hair cell impairment (Zilany and Bruce, 2007). In this section we investigate the neural response characteristics given inner or outer hair cell impairment alone. We would like to explore why and how gains differ when we remove the cochlear amplifier by impairing the outer hair cells, or impede the neuronal response by impairing the inner hair cells.

5.7.1 Responses With and Without Prescribed Amplification

Our initial exploration examines the behaviour of the box plots, power ratios, and histogram plots when sound is introduced to an auditory periphery with hearing thresholds that are a result of either solely inner or outer hair cell impairment. We compute both results before and after prescribed hearing amplification. Our simulations are run using an /ε/ vowel to help ensure proper measurements of the box plots and power ratios. A mildly impaired auditory periphery is utilized so that hearing loss can be fully attributed to either inner hair cell or outer hair cell impairment. Figures 5.11, 5.12, 5.13 highlight the results from this experiment.

We notice from the box-plots that the neural response is markedly diminished with inner hair cell impairment. The reason for this is that greater damage to hair cell stereocilia prevents these cells from being readily depolarized. This in turn directly affects the rate of discharge of auditory nerve neurons that are synapsed to the damaged hair cell. Interestingly enough, when looking at the power ratios, significant spread of synchrony is observed when amplification prescriptions are applied, however looking at the box plots, rates are more pronounced, reaching levels closer to normal,

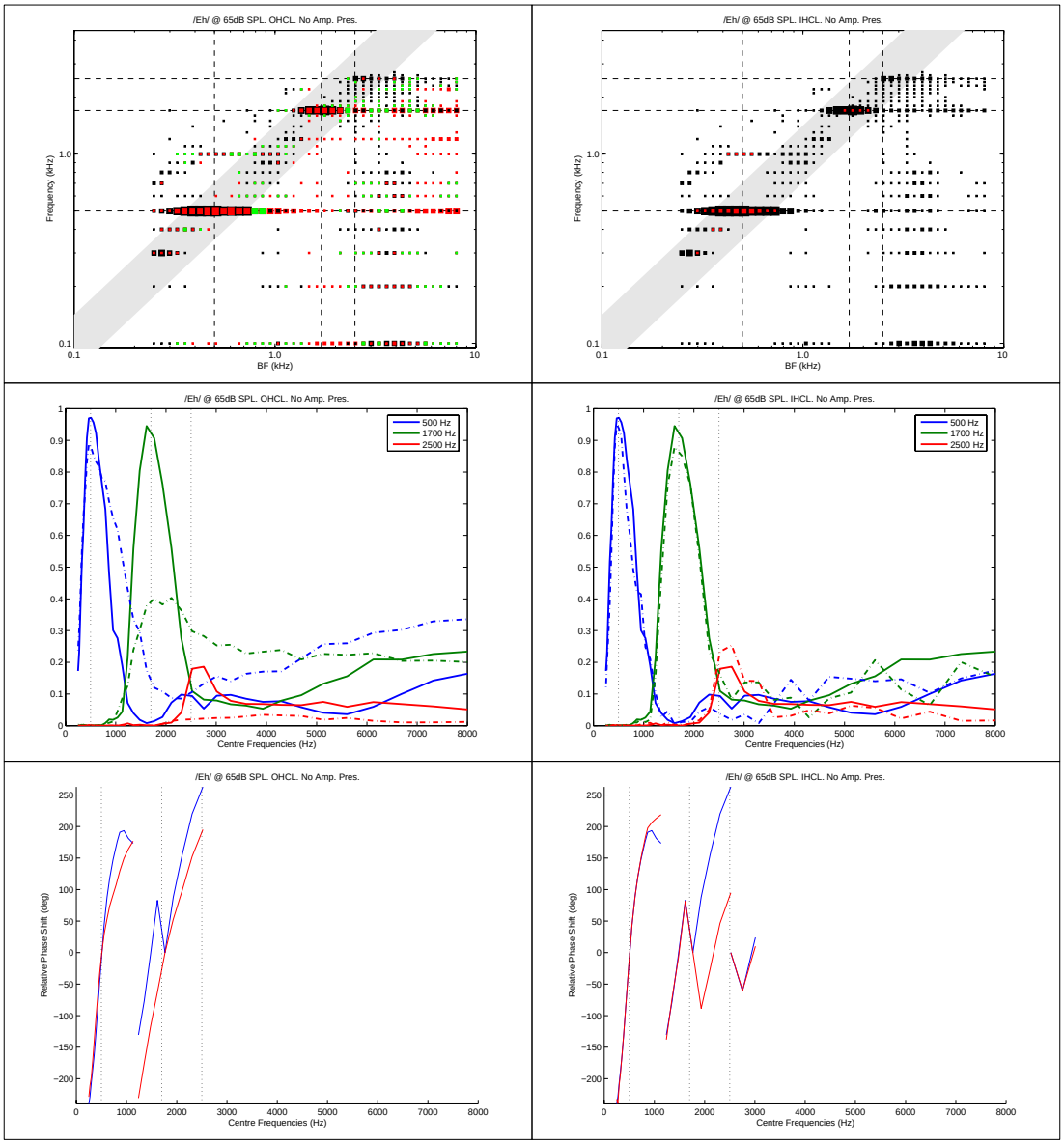


Figure 5.11: Box-plots, power ratios, and phase responses shown for solely outer hair cell impairment (left) and inner hair cell impairment (right). Box Plots: Black markers indicate normal response, red for impaired, and green markers for identical responses between normal and impaired responses. Power Ratios: Impaired responses are shown in dashed lines and normal in solid. Phase responses: Blue lines indicate normal response and red for impaired.

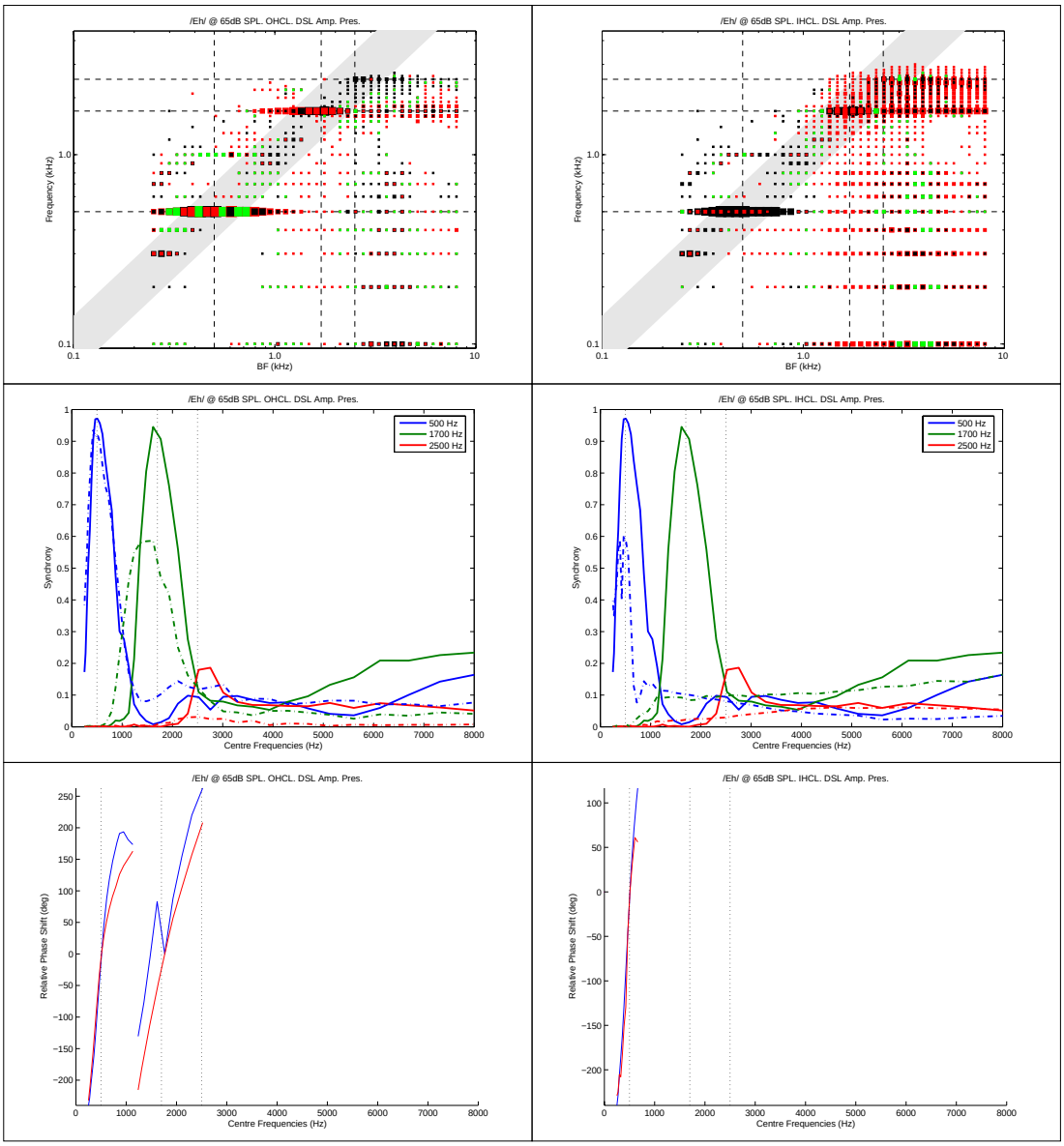


Figure 5.12: Box-plots, power ratios, and phase responses shown for solely outer hair cell impairment (left) and inner hair cell impairment (right), both after amplification has been applied. Box Plots: Black markers impaired normal response, red for impaired, and green markers for identical responses between normal and impaired responses. Power Ratios: Impaired responses are shown in dashed lines and normal in solid. Phase responses: Blue lines indicate normal response and red for impaired.

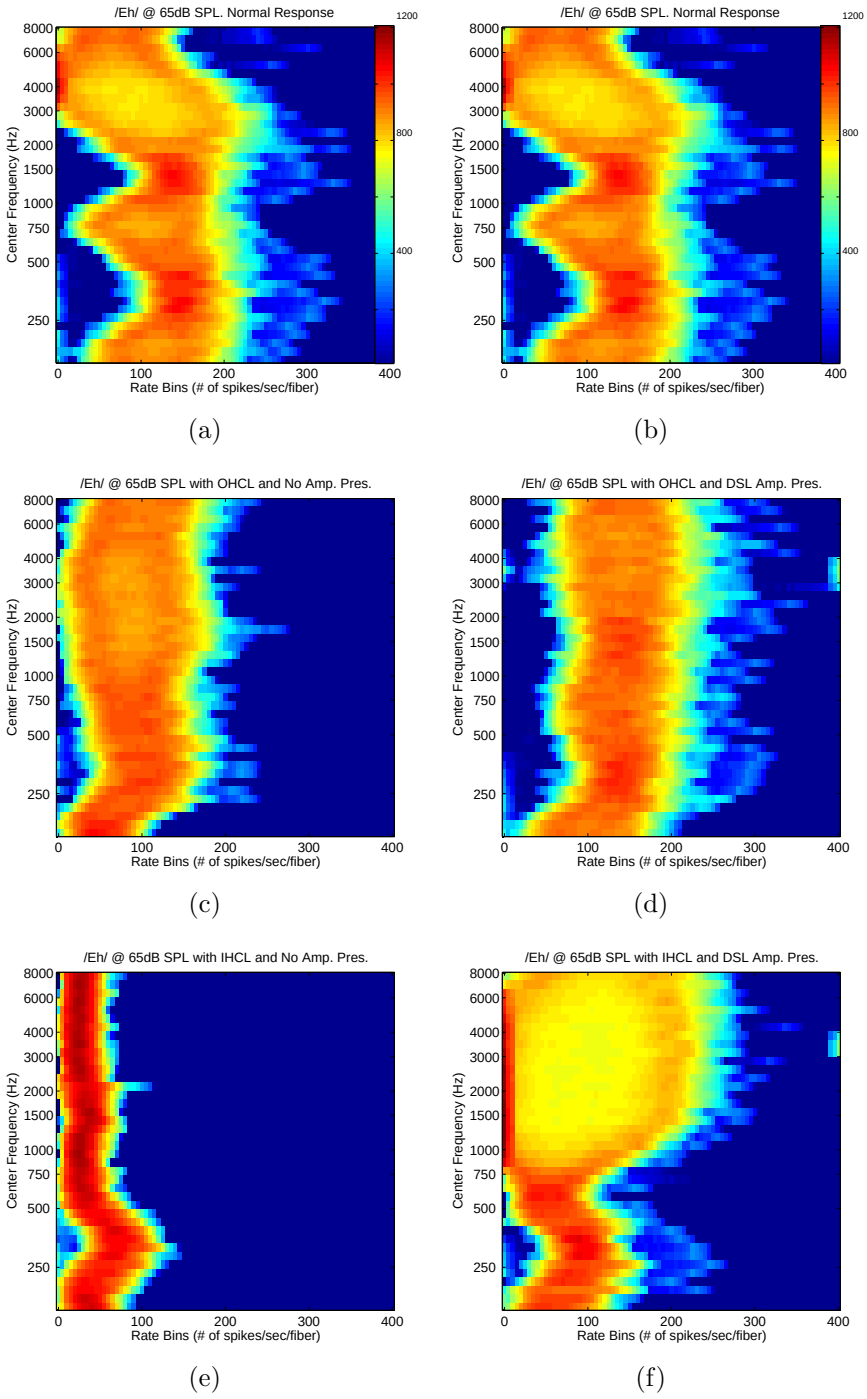


Figure 5.13: The figures above were generated in response to a 65dB SPL /ε/ vowel segment with a 100Hz fundamental frequency. The histograms are composed of 80 equally spaced bins of width 5 spike/s and low-pass filtered at a cut-off frequency of 250 Hz. Each histogram represents the distribution of bins with the colourbar in (a) indicating the magnitude of each bin. Figure values are represented as log base 10 for enhanced contrast.

particularly at the formant CF locations. The results indicate that to restore the rate response when only when IHC loss is present, high gains are necessary, albeit at a cost of spread of synchrony.

On the other hand by completely impairing the OHCs, we remove the cochlear amplifier effect, making the basilar membrane less responsive to sounds pressures. This effect is directly seen by a broadened and elevated tuning curve, and by an upward spread of synchrony in the power ratio plots. The simulations here indicate that adding amplification gains, such as the DSL algorithm, restores the box plots, but does not improve proper synchrony.

Finally, regardless of OHC or IHC impairment, the common consequence is that the neural response rate diminishes. This effect is demonstrated in Figs. 5.13 where either full OHC or IHC impairment without prescribed amplification shows much lower rate distribution when compared to the normal response. Amplification gains are required to bring the overall rates closer to that of normal.

5.7.2 Modified Error Metric

In the previous section we made visual comparisons between the normal and impaired neural responses solely when inner and outer hair cell impairment is considered. Unfortunately, calculating errors between normal and impaired neural responses using the the mean absolute error metric does not necessarily provide the most accurate compression gains, especially in certain cases of the spike timing neurogram and OHC impairment. The trouble here is that in certain adjustment cases, the mean absolute error metric is minimized only when extremely low compression gains are used in the impaired periphery, to the point where the impaired neural responses vanishes to nothing and the mean absolute error metric is applied just on the normal neural response. We therefore opt not to use neurograms as the basis for compression gain optimization and instead rely on box-plot information when comparing differences in spike timing information and histogram of rate response when comparing differences in the average discharge rate histogram.

When comparing box-plot information, the mean absolute error metric sometimes provides multiple local minima's almost equivalent in error. To solve this, we modify the mean absolute error metric slightly by combining the mean absolute error metric together with a correlation coefficient error metric. A correlation coefficient metric (Eqn. 5.4) works by calculating the similarity between two data sets. Two curves are related by a positive linear relationship when the coefficient is 1, a negative linear relationship when the coefficient is -1, and completely independent when the coefficient is 0.

$$\mathrm{CCR} = \frac{n \cdot \sum x \cdot y - (\sum x)(\sum y)}{\sqrt{n \cdot (\sum x^2) - (\sum x)^2} \cdot \sqrt{n \cdot (\sum y^2) - (\sum y)^2}} \tag{5.4}$$

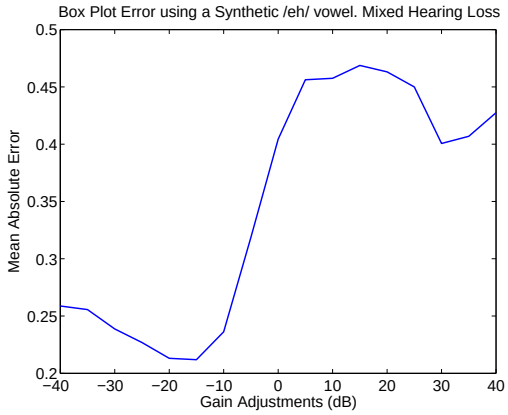
It is tempting to use the cross correlation coefficient metric alone over the mean absolute error metric, however the trouble with the correlation coefficient is that relative magnitude between curves is ignored. That is, the metric is maximized only when compared curves show a direct positive linear relationship. We overcome shortfalls associated with the mean absolute error and correlation coefficient metrics by combining both into a minimization metric as follows:

$$\mathrm{C} = \min[(1 - \alpha \cdot \mathrm{CCR}) \cdot \mathrm{MAE}] \tag{5.5}$$

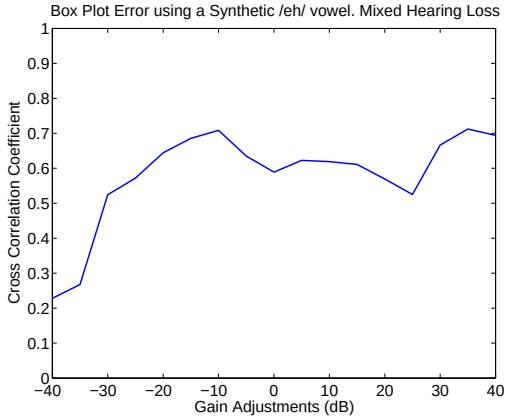
where we minimize[2] the cross-correlation by subtracting it from 1 and multiply the result to the mean absolute error metric. A scaling coefficient, α, is used to adjust the contribution of the cross-correlation coefficient in the combined error metric, although we choose α as 1 when using this metric throughout our simulations.

The combined error metric can be used to compare the neurograms, power ratios and phase responses (at each formant frequency), box plot, and histogram plots between the normal and impaired cases. Example error curves for all 3 metrics are shown in Figs. 5.14, 5.15, and 5.16 when comparing normal and impaired box-plots with a series of gain adjustments.

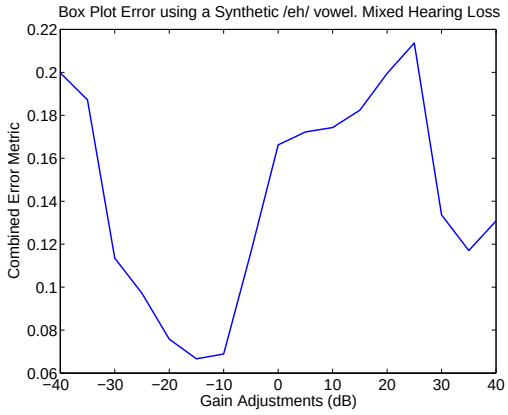
[2]Note that for our minimization, negative correlations are set to zero.



(a)

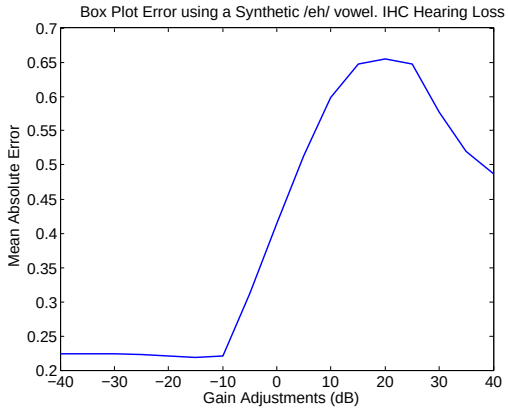


(b)

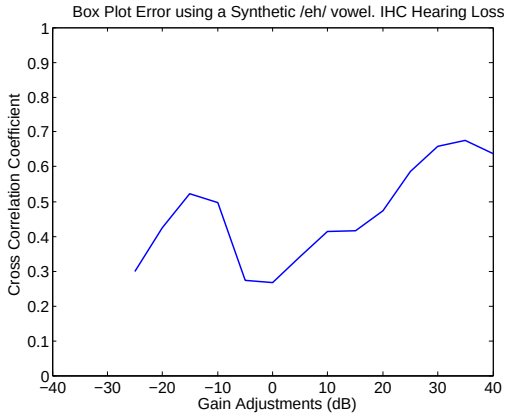


(c)

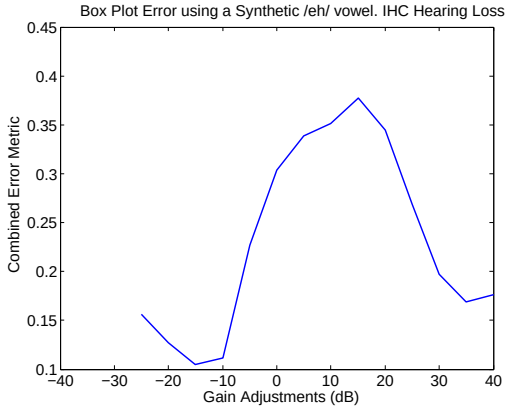
Figure 5.14: In the case of mixed hair cell impairment: error metric results after comparing box-plot representation of spike timing information from a normal and mildly impaired auditory periphery. Spike timing information was derived using a synthetic /ε/ vowel at 55 dB SPL. Note the clear local minima in the combined error metric.



(a)

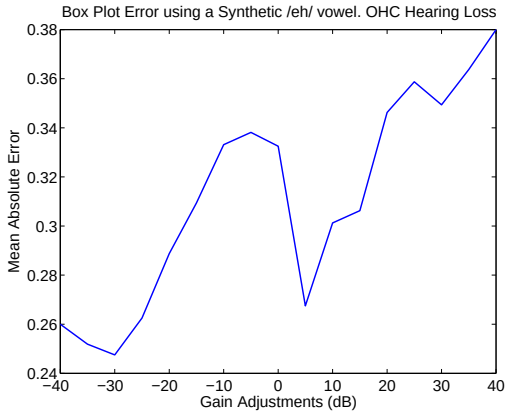


(b)

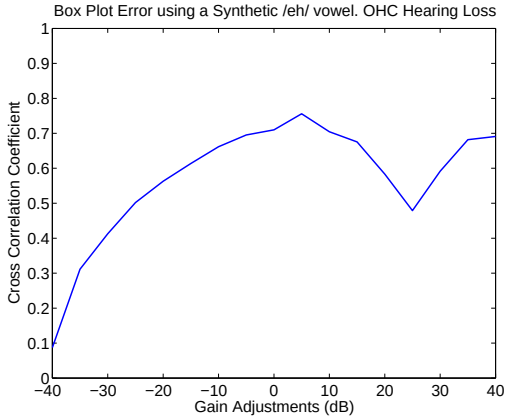


(c)

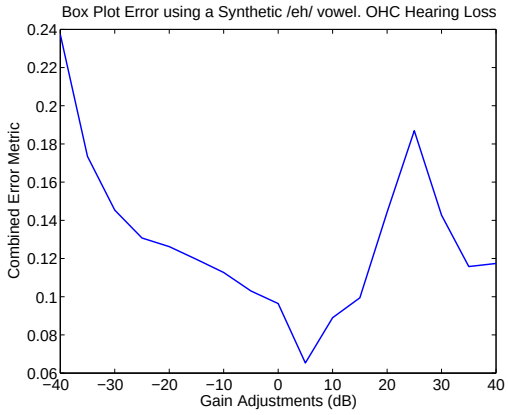
Figure 5.15: In the case of solely IHC impairment: error metric results after comparing box-plot representation of spike timing information from a normal and mildly impaired auditory periphery. Spike timing information was derived using a synthetic /ε/ vowel at 55 dB SPL. Note the clear local minima in the combined error metric.



(a)



(b)



(c)

Figure 5.16: In the case of solely OHC impairment: error metric results after comparing box-plot representation of spike timing information from a normal and mildly impaired auditory periphery. Spike timing information was derived using a synthetic /ε/ vowel at 55 dB SPL. Note the clear local minima in the combined error metric.

5.7.3 Optimal Gain Adjustments

We have proceeded with an initial investigation in visually comparing neural response characteristics of the auditory periphery with and without prescribed gains. Here we try to optimize the amplification gains using the gain optimization procedure described in section 5.5, but with an auditory periphery having either purely inner or outer hair cell impairment. The error metric follows from Eqn. 5.5. Furthermore, in this gain adjustment strategy, hearing loss was considered to be mild so that the full range of threshold shift can be accounted for by either inner or outer hair cell impairment.

The results from this gain adjustment experiment are shown in Fig. 5.17. Shown on each subplot are the optimal gains determined solely by OHC impairment, IHC impairment, and a mixture of both. As we have seen, mixed hair cell impairment shows positive optimal gain adjustments when the neurogram contains average discharge rates, and negative gain adjustments when the neurogram contain spike timing information.

When determining optimal gain adjustments using the fine timing neurogram with inner hair cell impairment alone, there is a need for lower gain adjustments, closer to the optimal gains determined for mixed hearing loss. This is because with lower gains, inner hair cells are capable of preserving the temporal response in the neurogram, including synchrony and phase capture. One caveat, however, is that because the inner hair cells are damaged, there is a direct impairment of the amount of neural activity derived from the hair cell (see appendix D). Therefore when optimizing using the average discharge rate neurogram, greater gains for IHC impairment are necessary to adequately restore the neural rate response.

Greater gains on the whole are required to compensate for outer hair cell impairment relative to mixed hair cell impairment, in both the average discharge rate or spike timing neurograms. This is partly due to having to compensate for additional amplification loss by removal of the cochlear amplifier. Furthermore, when looking at the spike timing neurogram, the neurogram is already suffering from spread of synchrony and phase changes. Therefore, adding additional compression gains does not worsen these properties substantially and may in fact help restore the rate component of the impaired spike-timing neurogram.

Interestingly, for the average discharge rate neurogram, purely inner hair cell impairment requires greater gains than outer hair cell impairment. This can be explained by referring to the rate level curve in Fig. 5.18. With inner hair cell damage, the slope of the rate level function lessens resulting in a maximum rate much lower than normal, and with outer hair cell impairment, the rate level function shifts to the right. When trying to restore the same rate response as in the normal auditory periphery, a greater gain is necessary to restore the normal response for inner hair cell damage than for outer hair cell damage.

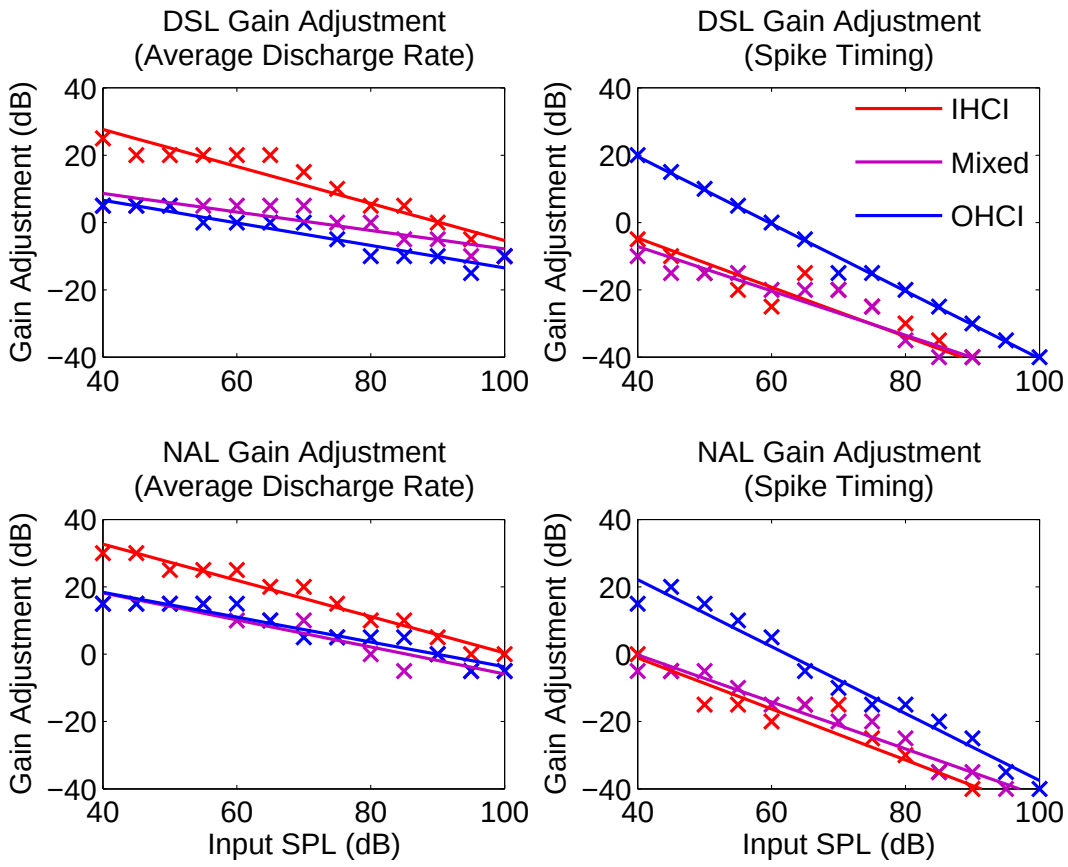


Figure 5.17: Red markers and lines depict solely inner hair cell impairment, purple for mixed, and blue for solely outer hair cell impairment.

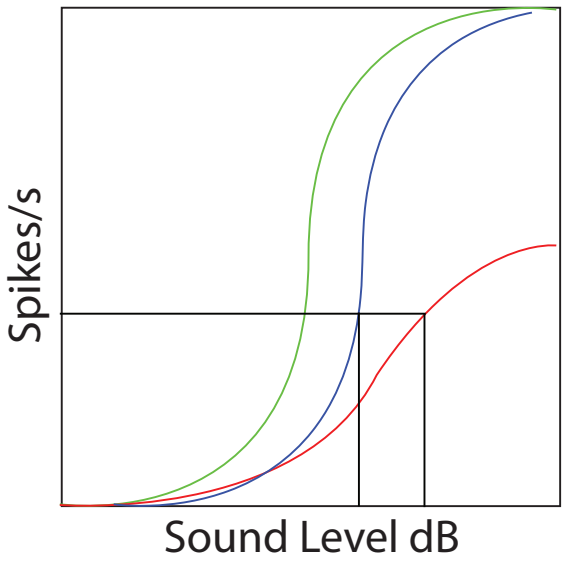


Figure 5.18: Auditory nerve rate response at an arbitrary CF. Green curve indicated normal rate response, blue curve when there is outer hair cell impairment, and the red curve when there is inner hair cell impairment. Note that inner hair cell impairment causes shallower response slope, and that outer hair cell impairment causes a right shift in the response.

5.8 Gain Adjustment Summary

Optimal gains are determined by fundamentally different properties of the two different types of neurograms. Mean discharge rate neurograms are optimally restored in the impaired auditory periphery based on the patterns of mean discharge rate across CF. This is in contrast to optimal restoration of fine timing neurograms, which depends on the degree of synchrony fibers have to the different speech frequencies, particularly formant frequencies, and the relative phase of synchronized responses in fibers responding to those frequencies.

The results with the real and synthetic vowel /ε/ strongly support the conclusions made in the gain optimization strategy for the TIMIT sentences, showing positive optimal gain adjustments for the average discharge rate, and negative optimal gain adjustments for the fine timing neurogram, when mixed IHC and OHC impairment is considered (Refer to Table 5.3).

Inner hair cell impairment shows greater synchrony at lower sound pressure levels, however much weaker neural output (Refer to Table 5.4). Increasing the gains restores neural output to normal levels, but produces significant spreads synchrony as well.

Regardless of whether the neurogram contains spike timing information or average discharge rate, strict outer hair cell impairment shows a need for greater gain adjustments. The reason here is possibly two fold. First is that impairment of OHCs removes the cochlear amplifier effect and greater gains are requires to overcome this loss. Second is that since OHCs impairment already produces a broadened tuning curves, adding more compression gain does not worsen this and can help restore the rate component of the neurogram.

Table 5.3: Gain Adjustments for Mixed IHC and OHC Hearing Loss

DSL/NAL	**Sentence**	**Synthetic Vowel**	**Spoken Vowel**
Fine Timing Neurogram	Positive	Positive	Positive
Avg. Discharge Rate	Negative	Negative	Negative

Table 5.4: Change in optimal gains for biased hearing loss relative to those for mixed hearing loss

DSL/NAL	Synthetic Vowel	
	IHC Loss	**OHC Loss**
Fine Timing Neurogram	Similar	Greater
Avg. Discharge Rate	Greater	Similar

Chapter 6

Conclusions

6.1 Optimal Compression Gains

In this thesis we examined optimal compression gains in regards to linear hearing aid prescriptions. Our analysis sought to determine changes that occur in the neurogram as compression gains are applied, and how do those changes aid in selection of optimal gains. We determined that when the neurogram contains the average discharge rate representation, compression gains try to restore a similar rate distribution pattern as that from the normal auditory periphery. For that of mixed hearing loss, the result is that greater gains are necessary. When the neurogram contains spike-timing information, compression gains try to optimize for the degree of synchrony and phase capture, particularly those around formant frequencies. For mixed hearing loss, these properties are best restored when negative compression gains are considered.

6.2 Differences in Hearing Aid Prescriptions

We determined that the optimal compressions gains using the average discharge rate neurogram show that DSL requires smaller gain adjustments whereas the optimal average gains using the fine timing neurogram (Fig. 5.4) show that NAL-R requires smaller gain adjustments. Discrepancies seen in optimal gain adjustments suggest that linear hearing aids, such as the NAL-R or DSL, may be trying to optimize both the fine timing and average discharge rate neurograms by presenting amplification gains in the region between optimal gains determined by the two types of neurograms. It is interesting to note, however, that the optimal gains using the average discharge rate neurogram indicate that DSL requires smaller gain adjustments than NAL-R, but the reverse is true from the optimal gains using the fine timing neurogram. The difference here suggests that even though amplification schemes may be

optimizing for both fine and average discharge rate neurograms, DSL is weighted towards optimizing the average discharge neurogram, owing to its greater gains relative to NAL-R, whereas NAL-R is weighted towards optimizing the fine timing neurogram.

6.3 Future Works

The work presented here is just an initial investigation into optimal compression gains. For future work, we identify at least 3 avenues for exploration. Since in this thesis we used single-band compression gains to optimize linear hearing aids, future research can try to explore the effects of multi-band compression schemes on optimal gain adjustments. Second, since we assume that compression acts instantaneously on a given vowel or sentence, we can try to make compression schemes more realistic by including attack and release times that govern how quickly compression turns on and off. Finally, the results here were based on one synthetic /ɛ/ vowel. Avenues of exploration exist in working with different synthetic vowels or sentences.

Chapter 7

Appendix

A MATLAB Code for Section 5.4

A-1 Sentence_Gains

```
function Speech_Gains( name, spl, prescription, psth_compare_type,...
    big_save,loss_type )
% Speech_Gains( name, spl, prescription, psth_compare_type,...
% big_save,loss_type )
%
% Works on a speech sentence from the TIMIT corpus phone by phone to
% determine the optimal compression gain adjustment after hearing aid
% amplification prescription is applied. Speech_Gains computes the mean
% absolute difference between the gain adjusted neurogram and the normal
% neurogram after each phone segment.
%
% Input:
%      Variables
% ____________________
% name               : Name of wave and phoneme files, without extension.
% spl                : Sound pressure level to normalize speech waveform to.
% prescription       : Either DSL or NALR
% psth_compare_type  : Comparison using either FINE or AVG neural output.
% big_save           : Specify BIG to save neural output, or SMALL not to.
% loss_type          1: Mild, Gently Sloping
%                    2: Moderate, Flat Loss
%                    3: Moderate, Steeply Sloping
%                    4: Profound, Gently Sloping
%
% Data comes from the TIMIT corpus and has a corresponding phone file
% (.PHN) associated with each sentence. Phones files indicates the type of
```

```
% phone as well as the start and stop samples where the phones begins and
% ends.
%
% Output: Output variables are saved in an .MAT file. The gain optimization
% results are summaried in the cell variable 'total'.
%
%
% Example: Speech_Gains( 'SA1', 65, 'NALR', 'FINE', 'SMALL', 3 );

%  Checks Input Variables
%
%=========================================================================%
if nargin ≠ 6,
    disp('6 Inputs Required');
    return;
elseif ¬ischar(name) || (size(name,1) ≠ 1)
    disp('Invalid File Name');
    return;
elseif ¬(strncmpi(prescription, 'DSL',3) || ...
        strncmpi(prescription, 'NAL',3))
    disp('Improper Prescription');
    return;
elseif ¬(strncmpi(psth_compare_type, 'Fine',3) || ...
        strncmpi(psth_compare_type, 'Avg',3))
    disp('Improper PSTH Compare Type');
    return;
elseif ¬(strncmpi(big_save, 'Big',3) || strncmpi(big_save, 'Small',3))
    disp('Improper Save Specification');
    return;
elseif ¬isscalar(spl) || ¬isnumeric(spl)
    spl = 65;
    disp('Assuming a SPL of 65 dB');
elseif ¬isscalar(loss_type) || ¬isnumeric(loss_type)
    loss_type = 3;
    disp('Assuming a Audiogram #3');
end

%=========================================================================%
% Creates a save directory and switches into it.
%
%=========================================================================%
eval('startup');
root = pwd;
eval(['mkdir ' root '/Runs/']);
eval(['cd ' root '/Runs/']);

%=========================================================================%
%  Loads The Phenome List
%
```

```
%=========================================================================%
[start,stop,phn] = textread([name '.PHN'],'%f%f%s');
start = start + 1;

%=========================================================================%
%  Intitializes the threshold shifts at specified frequencies (freq)
%
%=========================================================================%
k = 10.^((-40:5:40)/20);%Corresponding scaling factors for adding -20:20db.
[FREQ, H_IMP, H_NOR] = Audiogram_Patterns(loss_type);
num = 30;

%=========================================================================%
% Load Speech data. Applies Head Gain, SPL normalization, and prescription%
%=========================================================================%
[data,FS] = readsph([name '.WAV']);
data = set_spl(data,spl)';

if strncmpi(prescription, 'DSL',3)
    data_pres = DSL(data,FS,FREQ,H_IMP);
    save_file = [name '_DSL' '_' int2str(spl) '_' psth_compare_type];
elseif strncmpi(prescription, 'NAL',3)
    data_pres = resample(hrtffilt(resample(data,32000,FS),32000),FS,32000);
    % Head Related Transfer Function (hrtffilt) adds outer ear gains.
    data_pres = NALR(data_pres,FS,FREQ,H_IMP);
    save_file = [name '_NAL' '_' int2str(spl) '_' psth_compare_type];
end

data = resample(hrtffilt(resample(data,32000,FS),32000),FS,32000);

%=========================================================================%
% Creates a save directory and switches into it.
%
%=========================================================================%
if strncmpi(big_save, 'Big',3)
    eval(['mkdir ' root '/Runs/' save_file]);
end

%=========================================================================%
% Intitializes Storage Arrays
%
%=========================================================================%
output_data_normal = []; % Concatenation of normal phones.
output_data_impaired = []; % Concatenation of optimal prescription phones.

psth_normal = []; % PSTH of string of normal phones through normal model.
psth_impaired = []; % PSTH of string of normal phones after prescription
                    %and though impaired model.
psth_freq = [];
```

```
psth_time = [];

cpn  = []; % Colour Limits PSTH of Data through normal model.
cpi = []; % Colour Limits PSTH of Data after prescription
          % and through impaired model.
climits = [];
total = {};
%=========================================================================%

for l = 1:length(start)

    %=====================================================================%
    % Obtain normal auditory PSTH of normal phoneme
%
    %=====================================================================%
    phone_normal = data(start(l):stop(l));%Takes a normal phoneme from data
    phone_normal = resample(phone_normal,100000,FS);
    input_speech = [output_data_normal phone_normal];

    [psth_normal, psth_freq, psth_time, cpn] = ...
        PSTH(input_speech, FREQ, H_NOR, FS, num, psth_compare_type);

    output_data_normal = [output_data_normal phone_normal]; % Store phoneme

    %=====================================================================%
    % Save PSTH of normal phonome
%
    %=====================================================================%
    if (strncmpi(big_save, 'Big',3) & (l ≤ 8))
        save_details = [name '_' int2str(l) '_' int2str(0)];
        eval(['save ' root '/Runs/' save_file '/' ...
            save_details ' psth_normal psth_freq psth_time']);
    end
    %=====================================================================%

    for i = 1:length(k)
        %=================================================================%
        % Obtain impaired auditory PSTH of Gain*NALR phoneme
%
        %=================================================================%
        phone_pres = data_pres(start(l):stop(l));
        phone_pres = resample(phone_pres,100000,FS);

        input_speech = [output_data_impaired k(i)*phone_pres];

        [psth_impaired, psth_freq, psth_time, cpi] = ...
            PSTH(input_speech, FREQ, H_IMP, FS, num, psth_compare_type);
        %=================================================================%
```

```
        % Save impaired auditory PSTH of Gain*NALR phoneme
%
        %=========================================================================%
        if (strncmpi(big_save, 'Big',3) & (l ≤ 8))
            save_details = [name '_' int2str(l) '_' int2str(i)];
            eval(['save ' root '/Runs/' save_file '/' ...
                save_details ' psth_impaired psth_freq psth_time']);
        end
        %=========================================================================%
        abs_error(l,i) = sum(sum(abs(psth_normal - psth_impaired)));

    end

    %=============================================================================%
    % Saves Optimal Gain Adjustment and Current Phenome SPL
%
    %=============================================================================%
    j = find(abs_error(l,:) == min(abs_error(l,:)));

    spls_normal_phoneme(l) = get_spl(phone_normal);
    spls_pres_phoneme(l) = get_spl(phone_pres);
    spls_pres_phoneme_gain(l) = get_spl(k(j)*phone_pres);
    gains(l) = 20*log10(k(j));
    phone{l} = phoneme_lut(phn(l));
    total = [total ; {name} {l} {phone{l}} {start(l)} {stop(l)} ...
        {data(start(l):stop(l))} {spls_normal_phoneme(l)} ...
        {data_pres(start(l):stop(l))} {spls_pres_phoneme(l)} ...
        {gains(l)} {abs_error(l,:)}];
    % Store phoneme
    output_data_impaired = [output_data_impaired k(j)*phone_pres];

    eval(['save ' save_file]);

    %=============================================================================%
    % Colour Limits for PSTH and STFT
%
    %=============================================================================%
    climits = [climits ; [cpn cpi] ];
    %=============================================================================%

    [l 20*log10(k(j)) spls_pres_phoneme(l)]

end
```

B MATLAB Code for Sections 5.5 - 5.7

B-1 PSTHmay

```
function [ psth_struct ] = PSTHmay(S, aud, binwidth, varargin)
%PSTHmay Computes the AN Neurogram.
%
% [ Neurogram_Strucutre ] =
% PSTHmay(data_structure, audi_structure, binwidth, plot)
%
% This function computes the Neurogram at 'binwidth' resolution using the
% spike timing information from the Zilany–Bruce Cat Auditory Model. The AN
% response is taken at 'CF_Points' different center frequency locations,
% spaced logarithmically along the basilar membrane. At each CF point, the
% function calculates the summed response of 50 nerve fibers organized as
% 60%, high, 20% medium, and 20% low spontaneous rate fibers. The Audiogram
% structure specifies the hearing loss profile, as +dB loss, at specific
% frequencies. In addition to calculating the Neurogram, phase response,
% power ratio, box plots, and histogram analysis are calculated where
% possible. An optional potting parameter, if 'y' plots the Neurogram.
%
% See Also make_data_struct, audiograms, fd_phsr, fd_boxp, fd_hist,
% psth_plot.

disp('PSTHmay. Faheem Dinath. June 7th 2008.')


%%
%=========================================================================%
%                                Constants
%
%=========================================================================%

fs = 100e3;  % May change if FS is greater than fs %
ts = 1/fs;
b = double(single(1/binwidth));
spont = [50 5 .1];
nrep = round(50*[0.6 0.2 0.2]);

aud
psth_freq = aud.F;
Cohc = aud.Cohc;
Cihc = aud.Cihc;

%%
%=========================================================================%
```

```
%                    Check for proper FS and Binwidth
%
%=========================================================================%

if ( b ≥ fs) && (S.FS < b )
    type = 'FINE';
    fs = b;
    dat = resample(S.data,fs,S.FS);
    disp(['Data resampled to fs : ' num2str(fs) 'Hz'])
    disp('Returning Fine Timing Information')
elseif ( b < fs ) && ( S.FS < fs )
    type = 'AVG';
    if ¬iswhole(binwidth*fs)
        disp('Make sure that binwidth*100e3 is a whole number')
        return;
    end
    dat = resample(S.data,fs,S.FS);
    disp(['Data resampled to fs : ' num2str(fs) 'Hz'])
    disp(['Returning Average Discharge Rate With Binwidth of '...
        num2str(binwidth) 's'])
elseif ( S.FS ≥ fs ) && ( b < fs )
    type = 'AVG';
    fs = S.FS;
    if ¬iswhole(binwidth*fs)
        disp('Make sure that binwidth*FS is a whole number')
        return;
    end
    dat = S.data;
    disp(['Data presented at : ' num2str(fs) 'Hz'])
    disp(['Returning Average Discharge Rate With Binwidth of '...
        num2str(binwidth) 's'])
elseif ( S.FS ≥ b ) && ( b ≥fs )
    type = 'FINE';
    fs = S.FS;
    dat = S.data;
    disp(['Data presented at : ' num2str(fs) 'Hz'])
    disp('Returning Fine Timing Information')
else
    disp('Something is wrong with FS and binwidth.')
    return;
end

data = make_data_struct( dat, fs, S.SPL );
data.formants = S.formants;                 % Keep the same
data.approx_formants = S.approx_formants;   % Keep the same

if  isfield(S, 'data_orig') % If there's been amplification.
    data_orig = S.data_orig;
    S = rmfield(S,'data_orig');
```

```
    data_pres = S;
else                          % Otherwise there hasn't been amplification.
    data_orig = S;
end

%%
%=========================================================================%
%                              Filter Design
%
%=========================================================================%

if strcmp(type, 'AVG')
    Ap = 0.1;       % Passband Ripple
    Ast = 100;      % Amplitude below passband
    Fp = 250*binwidth;    % Cut-off at 250
    Fst = 350*binwidth;   % to 350 Hz
    disp('Low-Pass Filtering psth with cut-off @ 250 - 350 Hz')
elseif strcmp(type, 'FINE')
    Ap = 0.1;
    Ast = 100;
    Fp = 6000*binwidth;
    Fst = 8000*binwidth;
    disp('Low-Pass Filtering psth with cut-off @ 6000 - 8000 Hz')
end
% d = fdesign.lowpass(Fp,Fst,Ap,Ast);
% hd = design(d,'butter','matchexactly','passband');
% [h,t] = impz(hd);
[B,A] = butter(6,Fp,'low');
% freqz(B,A)

%%
%=========================================================================%
%                                Main  Loop
%
%=========================================================================%

%=+++++++++++++++++++++++++++% Clip Data %===+++++++++++++++++++++++++++=%
binw = round(binwidth*fs);              % How many samples are in each bin.
remainder = mod(length(dat),binw);      % Number of samples to cut off end.
dat = dat(1:end-remainder);             % Make length multiple of binwidth.
len = length(dat)/binw;                 % New length of binned data.
%+++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++++%

for i = 1:length(psth_freq)
    [a,a,a,a,a,a,a,a,psth500k_a] = zbcatmodel(dat,psth_freq(i),nrep(1),...
        ts,length(dat)*ts,Cohc(i),Cihc(i),spont(1)); % High sr
    [a,a,a,a,a,a,a,a,psth500k_b] = zbcatmodel(dat,psth_freq(i),nrep(2),...
        ts,length(dat)*ts,Cohc(i),Cihc(i),spont(2)); % Medium sr
    [a,a,a,a,a,a,a,a,psth500k_c] = zbcatmodel(dat,psth_freq(i),nrep(3),...
```

```
        ts,length(dat)*ts,Cohc(i),Cihc(i),spont(3)); % Low sr

    psth500k = psth500k_a + psth500k_b + psth500k_c;
    clear a psth500k_a psth500k_b psth500k_c

    pr = sum(reshape(psth500k,binw,len),1)/sum(nrep)/binwidth;
                                        % psth in units of spikes/s/fiber
%     pr = filter(hd, pr);
    pr = filtfilt(B,A,pr);
    psth(i,:) = single(pr);
end

psth_struct.type = type;
psth_struct.psth = psth(:,1:len);
psth_struct.psth_time = [0:(len-1)]*binwidth;
psth_struct.psth_freq = psth_freq;
psth_struct.psth_mnmx =[min(psth_struct.psth(:)) max(psth_struct.psth(:))];
psth_struct.binwidth = binwidth;
psth_struct.data_struct = data;
psth_struct.data_orig = data_orig;
if exist('data_pres')
    psth_struct.data_pres = data_pres;
end
psth_struct.audiogram_struct = aud;
[psth_struct.F, psth_struct.PSTH, psth_struct.PHASE] =...
    quickfft( psth_struct.psth, 1/binwidth );

%%
%=========================================================================%
%                       PR, PH, BOX, HIST & Plots
%
%=========================================================================%

if strcmp(psth_struct.type, 'FINE')
    psth_struct.psth_norm = 1/sum(nrep)/binwidth;   % Normalization Coeff.
    disp(['Calculating Box plot, Power Ratio, and Phase Responses @ '...
        num2str(psth_struct.data_struct.approx_formants) ' Hz']);
    psth_struct.phsr_freq = {};
    psth_struct.phsr = {};
    psth_struct.pwrr = [];
    for i = 1:length(psth_struct.data_struct.approx_formants)
        [psth_struct.phsr_freq{i}, psth_struct.phsr{i}, ...
            psth_struct.pwrr(i,:)] = fd_phsr( psth_struct, ...
            psth_struct.data_struct.approx_formants(i) );
    end
    [ psth_struct.boxp_freq, psth_struct.boxp] = fd_boxp( psth_struct );
    psth_struct.boxp_mnmx = [min(psth_struct.boxp(:)) ...
        max(psth_struct.boxp(:))];
elseif strcmp(psth_struct.type, 'AVG')
```

```
    disp('Calculating the Histogram Response')
    [ psth_struct.hist_bins, psth_struct.hist ] = fd_hist( psth_struct );
    psth_struct.hist_mnmx = [min(psth_struct.hist(:)) ...
        max(psth_struct.hist(:))];
end

if strncmpi(varargin{1},'y',1)
    psth_plot( psth_struct );
end

%%
%=========================================================================%
%                            Is a Whole Number?
%
%=========================================================================%
function [ flag ] = iswhole( number )
flag = (number == round(number));
```

B-2 ampl_pres

```
function [ data_struct_new ] = ampl_pres( data_struct, audi_struct, pres_type )
% AMPL_PRES applies either NAL-R or DSL amplification prescriptions
%
% [ data_struct_out ] = ampl_pres( data_struct_in, audi_struct, pres_type )
% uses the 'audiograms' data structure to apply appropriate NAL-R or DSL
% hearing aid prescription gains to the data in data_struct_in.
% Amplificiation prescription is specified by setting 'pres_type' as either
% 'NAL' or 'DSL'. The resulting data is contained in the structure
% data_struct_out along with the original data.
%
% See Also make_data_struct

if strcmp(audi_struct.type, 'Normal')
    disp('Warning: You are appling amplification gains to a normal AP.');
end

%%
%=========================================================================%
%                                Variables
%
%=========================================================================%
data_length = length(data_struct.data);
window = 256;
shift = 128;
fft_length = 2^(ceil(log2(window))+1);
zero_padding = fft_length - window;
f = -0.5:1/fft_length:0.5; f = f*data_struct.FS; f(end) = [];
%=========================================================================%
```

```
if strcmp(pres_type, 'DSL')
    %%
    %=========================================================================%
    %                        DSL Amplification Scheme
%
    %=========================================================================%
F = [250 500 750 1000 1500 2000 3000 4000 6000 8000];
H = [0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95 100 105 110];
Z = [[0  3  5  7  9  12 14 17 20 22 25 29 32 36 39 43 47 51 55 59 62 66 68]'...
     [2  4  6  8  11 13 15 18 20 23 26 29 32 35 38 42 45 48 52 55 59 62 66]'...
     [3  5  7  10 12 14 17 19 22 25 28 31 34 37 40 43 47 50 54 57 61 64 68]'...
     [3  5  8  10 13 15 18 21 24 27 30 33 36 40 43 46 50 53 57 60 64 68 71]'...
     [5  8  10 13 15 18 20 23 26 29 32 35 38 42 45 48 52 55 59 62 66 70 73]'...
     [12 15 17 19 22 24 27 30 33 36 39 42 46 49 52 56 59 63 66 70 73 77 80]'...
     [16 18 20 23 25 28 30 33 36 39 42 45 48 52 55 59 62 66 69 73 76 80 83]'...
     [14 17 19 21 24 27 29 32 35 38 41 45 48 51 55 58 62 65 69 73 76 80 84]'...
     [8  11 14 17 20 23 26 29 32 36 39 43 46 50 54 58 61 65 69 73 76 81 85]'...
     [6  10 13 16 19 22 25 27 30 35 38 42 45 49 53 58 60 64 68 72 75 80 85]'];

    REAG_dB = interp2(F,H,Z,audi_struct.F,audi_struct.H, 'cubic');
    REAG_SC = 10.^(REAG_dB/20);

    % PLOTTING %
%     plot(audi_struct.F, REAG_dB)
%     title('DSL Real Ear Unadided Amplification Gains')
%     xlabel('Center Frequency (CF)')
%     ylabel('Amplification Gain (dB)')
%     xlim([0 8000]);

    % NO NEED TO APPLY THE HEAD RELATED TRANSFER FUNCTION %
    data = data_struct.data;
    gains = fftshift(interp1([0 audi_struct.F],[1 REAG_SC],abs(f),'cubic'));
    %=========================================================================%

elseif strcmp(pres_type, 'NAL')
    %%
    %=========================================================================%
    %                      NAL-R Amplification Scheme
%
    %=========================================================================%

    % The NAL-R formular gives a constant for each F:
    F = [250 500 750 1000 1500 2000 3000 4000 6000 8000];
    k = [-17  -8  -3    1    0   -1   -2   -2   -2   -2];

    % The NAL-R formular averages loss at F = 500, 1000, and 2000 Hz
    H = audi_struct.orig_H;     % This should corrrespond to the F vector.
    H3FA = (H(2) + H(4) + H(6))/3;
```

```
    X = 0.15*H3FA;              % Constant
    IG_dB = X + 0.31*H + k;     % Gains
    IG_dB(IG_dB<0) = 0;         % My addition: Remove losses.
    IG_dB = interp1(F, IG_dB, audi_struct.F, 'cubic');
    IG_SC = 10.^(IG_dB/20);

    % PLOTTING %
%     plot(audi_struct.F, IG_dB)
%     title('NAL-R Insertion Amplification Gains')
%     xlabel('Center Frequency (CF)')
%     ylabel('Amplification Gain (dB)')
%     xlim([0 8000]);

    % MUST APPLY THE HEAD RELATED TRANSFER FUNCTION %
    data = hrtffilt(data_struct.data,data_struct.FS);
    gains = fftshift(interp1([0 audi_struct.F],[1 IG_SC],abs(f),'cubic'));
    %=========================================================================%

else
    disp('No Amplification Prescription Applied.')
    data_struct_new = data_struct;
    return
end

%%
%=============================================================================%
%                       The Zero-Phase Filtering Loop
%
%=============================================================================%
data = stratify(data, window, 'shift', shift);
for i = 1:size(data,1)
    if i == 1
        w = hanning(window); w(1:ceil(window/2)) = 1;
        x = data(i,:) .* w';
    elseif i == size(data,1)
        w = hanning(window); w((ceil(window/2)+1):end) = 1;
        x = data(i,:) .* w';
    else
        x = data(i,:).*hanning(window)';
    end
    x = [x zeros(1,zero_padding)];
    X = fft(x).*gains;
    x = ifft(X);
    data(i,:) = x(1:window);
end
data = overlapandadd(data,window,shift);
data = data(1:data_length);
%=============================================================================%
```

```
data_struct_new = make_data_struct( data, data_struct.FS, get_spl(data) );
data_struct_new.formants = data_struct.formants;                % Keep the same
data_struct_new.approx_formants = data_struct.approx_formants;  % Keep the same
% Make fundamental freq the same too?? Seems to stay the same..
data_struct_new.pres_type = pres_type;
data_struct_new.pres_gain = fftshift(gains);
data_struct_new.pres_freq = f;
% Copy over the old FFT's too??
% data_struct_new.data_orig = data_struct.data;
% data_struct_new.SPL_orig = data_struct.SPL;
data_struct_new.data_orig = data_struct;
```

B-3 audiograms

```
function audiogram = audiograms(num, CF_Points, varargin)
%AUDIOGRAMS returns a sample audiogram profile.
%
% Audiogram_Strucutre = audiograms(LookUpNumber, CF_Points) uses a 15 row
% look-up take to return a sample audiogram. LookUpNumber number 1 contains
% a normal audiogram profile with an overall zero dB loss. Each audiogram
% profile contains at least 10 logarithmically spaced hearing losses. The
% parameter CF_Points specifies the number of threshold points along the
% audiogram profile, with 10 being the least.
%
% REF: Susan Scollie, Richard Seewald, Leonard Cornelisse, Sheila Moodie,
% Marlene Bagatto, Diana Laurnagaray, Steve Beaulac and John Pumford. The
% Desired Sensation Level Multistage Input/Output Algorithm. Trends Amplif.
% 2005;9(4):159-97.

AG = [[ 0  0  0  0  0  0  0  0  0  0];...
      [20 20 25 30 40 45 50 50 50 50];...
      [25 30 45 55 65 80 85 90 90 90];...
      [30 30 30 30 30 30 30 30 30 30];...
      [30 30 35 35 40 40 45 45 50 50];...
      [30 30 35 40 45 50 60 70 75 75];...
      [30 35 40 45 50 60 80 90 90 90];...
      [40 40 40 40 40 40 40 40 40 40];...
      [50 50 50 50 50 50 50 50 50 50];...
      [50 50 55 55 60 60 65 65 70 70];...
      [50 50 50 55 65 70 80 90 95 95];...
      [70 70 70 70 70 70 70 70 70 70];...
      [70 70 75 75 80 80 85 85 90 90];...
      [90 90 90 90 90 90 90 90 90 90];...
      [80 85 90 95 100 110 110 110 110 110]];

type = {'Normal', 'Mild', 'Moderate', 'Mild', 'Mild', 'Moderate',...
    'Moderate', 'Moderate', 'Moderate', 'Moderate', 'Moderate',...
```

```
    'Moderate', 'Severe', 'Severe', 'Profound',};

if num < 1
    num = 1;
end

if num > 15
    num = 15;
end

%%
t = type{num};
F = [250 500 750 1000 1500 2000 3000 4000 6000 8000];

if exist('CF_Points') && ( CF_Points > 10 )
    FREQS = logspace( log10( F(1) ), log10( F(end) ), CF_Points );
    H = interp1(F,AG(num,:),FREQS,'cubic'); % 'linear' causes problems!
else
    FREQS = F;
    H = AG(num,:);
end

if ¬isempty(varargin)
    IOHC_loss = varargin{1};
else
    IOHC_loss = 'Mixed';
end

if strcmp(IOHC_loss, 'OHCL')
    [Cohc,Cihc]=fitaudiogram(FREQS,H,H);
    disp('Impariment due to Outer Hair Cell Loss');
elseif strcmp(IOHC_loss, 'IHCL')
    [Cohc,Cihc]=fitaudiogram(FREQS,H,zeros(size(H)));
    disp('Impariment due to Inner Hair Cell Loss');
else
    IOHC_loss = 'Mixed';
    [Cohc,Cihc]=fitaudiogram(FREQS,H);
    disp('Mixed Inner & Outer Hair Cell Loss');
end

%%
audiogram.type = t;
audiogram.F = round(FREQS);
audiogram.H = H;
audiogram.Cohc = Cohc;
audiogram.Cihc = Cihc;
audiogram.orig_F = F;
audiogram.orig_H = AG(num,:);
audiogram.IOHC_loss = IOHC_loss;
```

```
%%
if nargout == 0
    F = floor(log2(F/125)/0.5)/2 + 1;
    figure;
    plot(F, -AG(num,:),'rx', 'Markersize', 20, 'Linewidth', 4);
    ylim([-120 10]); set(gca, 'XScale', 'linear')
    hold on;

    line(3.5*ones(1,2),[10 -120],'color','k','Linestyle','--');
    line(4.5*ones(1,2),[10 -120],'color','k','Linestyle','--');
    line(5.5*ones(1,2),[10 -120],'color','k','Linestyle','--');
    line(6.5*ones(1,2),[10 -120],'color','k','Linestyle','--');

    line([0 8000], 10*ones(1,2),'color','k');
    line([0 8000], -10*ones(1,2),'color','k');
    line([0 8000], -30*ones(1,2),'color','k');
    line([0 8000], -50*ones(1,2),'color','k');
    line([0 8000], -70*ones(1,2),'color','k');
    line([0 8000], -90*ones(1,2),'color','k');
    line([0 8000], -110*ones(1,2),'color','k');
    line([0 8000], -120*ones(1,2),'color','k');

    hold off;
    set(gca,'XAxisLocation','top');
    set(gca, 'XTick', [1 2 3 4 5 6 7]);
    set(gca, 'XTickLabel', [125 250 500 1000 2000 4000 8000]);
    set(gca, 'YTick', fliplr([0 -20 -40 -60 -80 -100 -120]));
    set(gca, 'YTickLabel', fliplr([0 20 40 60 80 100 120]));
    xlim([1 7]);
    ylim([-120 10]);
    grid;
    set(gca, 'GridLineStyle', '-');
    xlabel('Frequency (Hz)', 'FontSize', 16);
    ylabel('Hearing Threshold Loss (dB)', 'FontSize', 16);
    title([t ' Hearing Loss'], 'FontSize', 16);
end
```

B-4 fd_boxp

```
function [ f, ft_f0 ] = fd_boxp(psth_struct, varargin);
% FD_BOXP computes the strength of AN fiber phase locking to individual
% frequency components of a vowel.
%
% [ f, box_data] = fd_boxp(Neurogram_Strucutre) receives a neurogram data
% structure which it uses to compute the box plot. An 80ms Hamming window
% is isolates a portion of the neurogram in 20 ms from the stimulus onset.
% An offset is requied to avoid any adaptation irregularities and a
```

```
% windowed response is taken because it is assumed that the stimulus is
% perioid, ie, a vowel. The Fourier transform of the windowed neurogram is
% computed and normalized to express the transform components in units of
% spikes/s. The magnitude of the resulting transform components are the
% synchonized rates.
%
% REF: Roger L. Miller et al. Effects of acoustic trauma on the
% representation of the vowel /e/ in cat auditory nerve fibers. J. Acoust
% Soc. Am. 1997.

disp('fd_boxp. Faheem Dinath. May 29th 2008.')

% Make sure that we're looking at fine timing PSTH
if ¬strcmp(psth_struct.type, 'FINE')
    f = [];
    ft_f0 = [];
    return
end

%=========================================================================%
if (nargout == 0) | strncmpi(varargin,'y',1)
    h1 = figure;
    h3 = fill([0.1 10 10 0.1],...
        [0.1*2^(1/2) 10*2^(1/2) 10/2^(1/2) 0.1/2^(1/2)],0.9*ones(1,3));
    set(h3,'edgecolor','none')
    set(gca,'xscale','log','yscale','log')
    xlabel('BF (kHz)')
    ylabel('Frequency (kHz)')
    axis([0.1 10 0.095 4.5])
    set(gca,'xticklabel','0.1|1.0|10')
    set(gca,'yticklabel','0.1|1.0')
    hold on
    plot([0.5 0.5],[0.095 4.5],'k--')
    plot([1.7 1.7],[0.095 4.5],'k--')
    plot([2.5 2.5],[0.095 4.5],'k--')
    plot([0.1 10],[0.5 0.5],'k--')
    plot([0.1 10],[1.7 1.7],'k--')
    plot([0.1 10],[2.5 2.5],'k--')
    wysiwyg
end
%=========================================================================%

psth_freq = psth_struct.psth_freq;
psth_time = psth_struct.psth_time;
binwidth = psth_struct.binwidth;

for cfslp=1:length(psth_freq)

    cf=psth_freq(cfslp);
```

```
    if psth_time(end) < 0.1
        disp('Data run-time too short; must be greater than 100 ms');
        return
    end

    [dummyy, tonset] =min(abs(psth_time-20e-3));

    toffset = tonset+ 80e-3/binwidth;
    toffset = round(toffset) - 1;

%     Hamming window starting 20ms after STIMULUS onset
    p = psth_struct.psth(cfslp, tonset:toffset);
    w = hamming(length(p))';

    [f, MX, P] = quickfft( w.*p, 1/binwidth );
    ft(cfslp,:)=MX/sqrt(sum(w.^2)/length(w));

    f0 = psth_struct.data_struct.fundamental;

    [mn f_min] = min(abs(f - f0));
    [mn f_max] = min(abs(f - 4e3));
    f_ind = f_min:f_min-1:f_max;
    [ freq, val ] = max_value_around( psth_struct.data_struct, f(f_ind));

    for i = 1:length(freq)
        [v add] = min( abs( f - freq(i) ) );
        f_ind(i) = add;
    end

%     f_ind = find(f==100):find(f==100)-1:find(f==4e3);

    if (nargout == 0) | strncmpi(varargin,'y',1)
        figure(h1)
        h2 = loglog(cf/1e3,f(f_ind)/1e3,'ks','markerfacecolor','k');
        for lp=1:length(h2)
            if ft(cfslp,f_ind(lp))<15
                set(h2(lp),'marker','none')
            elseif ft(cfslp,f_ind(lp))<30
                set(h2(lp),'markersize',2)
            elseif ft(cfslp,f_ind(lp))<45
                set(h2(lp),'markersize',4)
            elseif ft(cfslp,f_ind(lp))<60
                set(h2(lp),'markersize',6)
            elseif ft(cfslp,f_ind(lp))<75
                set(h2(lp),'markersize',8)
            elseif ft(cfslp,f_ind(lp))<90
                set(h2(lp),'markersize',10)
            elseif ft(cfslp,f_ind(lp))<105
```

```
                set(h2(lp),'markersize',12)
            elseif ft(cfslp,f_ind(lp))<120
                set(h2(lp),'markersize',14)
            else
                set(h2(lp),'markersize',16)
            end
        end
    end

    ft_f0(cfslp,:) = ft(cfslp,f_ind);
end

ft_f0(ft_f0≤15) = 0;
ft_f0((ft_f0>15) & (ft_f0≤30)) = 1;
ft_f0((ft_f0>30) & (ft_f0≤45)) = 2;
ft_f0((ft_f0>45) & (ft_f0≤60)) = 3;
ft_f0((ft_f0>60) & (ft_f0≤75)) = 4;
ft_f0((ft_f0>75) & (ft_f0≤90)) = 5;
ft_f0((ft_f0>90) & (ft_f0≤105)) = 6;
ft_f0((ft_f0>105) & (ft_f0≤120)) = 7;
ft_f0(ft_f0>120) = 8;
f = f(f_ind);
```

B-5 fd_hist

```
function [ bins, psth_hist ] = fd_hist( psth_struct, varargin )
% FD_HIST computes the count distribution of rates in the neurogram by
% placing rates into 200 bins.
%
% [ bins, histogram ] = fd_hist( Neurogram_Struct, varargin )

if ¬strcmp(psth_struct.type, 'AVG')
    bins = [];
    psth_hist = [];
    return
end

fibre = size(psth_struct.psth,1);
psth_hist = [];

for i = 1:fibre
    [a bins] = hist(psth_struct.psth(i,:), 0:5:200);
    psth_hist(i,:) = smooth(a);
end

if (nargout == 0) | strncmpi(varargin,'y',1)
    figure;
    imagesc(bins, psth_struct.psth_freq, psth_hist); axis xy
```

```
    set(gca, 'YScale', 'log');
    set(gca, 'YTick', [250 500 750 1000 1500 2000 3000 4000 6000 8000]);
    set(gca, 'YTickLabel', [250 500 750 1000 1500 2000 3000 4000 6000 8000]);
    ylabel('Center Frequency (Hz)')
    xlabel('Rate Bins (spikes/sec/fiber)')
    title('Histogram Response')
end
```

B-6 fd_phsr

```
function [freq, phase, pr] = fd_phsr(psth_struct, CF, varargin)
% FD_PHSR computes the phase response.
%
% [freq, phase, pr] = fd_phsr(Neurogram_Struct, CF, varargin)
% uses the neurogram to calculate the phase of a partiuclar speech
% frequency component in the region around a fiber with the same CF. The
% region is determined using the power ration, as a continuois adjacent
% segment where the power ratio is greater than 0.1. This function
% calculates both the phase and power ratio repsonse for the specified CF.

% Make sure that we're looking at fine timing PSTH
if ¬strcmp(psth_struct.type, 'FINE')
    freq = [];
    phase = [];
    return
end

pr = fd_pwrr(psth_struct, CF, varargin);

disp('fd_phsr. Faheem Dinath. May 29th 2008.')

psth_freq = psth_struct.psth_freq;

% / Find Continuous Region Greater than 0.1 \
%=========================================================================%
[temp pos] = min(abs(psth_freq - CF));

if pr(pos) ≤0.1
    freq = [];
    phase = [];
    return;
end

BW = pr>0.1;
[BW,num] = bwlabel(BW);
num = BW(pos);
BW = find(BW == num);     % Binary Vector indicating where region lie.
```

```
%=========================================================================%

for i = 1:length(BW)
    [temp pos] = min( abs( psth_struct.F - CF ) );
    phase(i) = psth_struct.PHASE(BW(i), pos);
end

% / Unwrap Phase \
%=========================================================================%
freq = psth_freq(BW);
[temp pos] = min(abs(freq-CF));
phase = phase*2*pi/180; % Convert to Radians
phase = mod(phase,2*pi);
phase = unwrap(phase);
phase = phase - phase(pos);
phase = phase*180/2/pi; % Convert to Degrees
%=========================================================================%

% / Plot PR and Phase Response \
%=========================================================================%
if (nargout == 0) | strncmpi(varargin,'y',1)

    figure;
    plot(gca,psth_struct.psth_freq,pr);
    drawnow;
    xlabel('Centre Frequencies (Hz)')
    title(['Power Ratio at ' num2str(CF) ' Hz'])
    ylim([0 1]);

    figure;
    plot(gca,freq,phase);
    drawnow;
    ylabel('Relative Phase Shift (deg)')
    xlabel('Centre Frequencies (Hz)')
    title(['Phase Response at ' num2str(CF) ' Hz'])

end
%=========================================================================%
```

B-7 fd_pwrr

```
function pr = fd_pwrr(psth_struct, fx, varargin)
%[PR, H] = fd_pwrr(Neurogram_Struct, fx) computes the
%strength of AN fiber phase locking to individual frequency components of a
%vowel.
%
%   fdpowerratio recieves a peritimulus time historgram, psth, its
%   corresponding center frequency axis, psth_freq, and the binwidth used
```

```
%   to compute the PSTH from the spike timing information. The variable fx
%   is the frequency component of the vowel to which phase locking is
%   measured. fdpowerratio returns the handle of the figure plot, H, and
%   figure data, PR.
%
%   The power ratio is defined as the sum of power in the AN fft response
%   at the frequency fx and its harmonics, divided by the total power in
%   the response. Because phase locking in cats is not observed above 5kHz,
%   the summations are limited to frequency componenets below 5kHz.
%
%
%                      u
%                     ___
%                     \      ( R^2(m*fx) )
%                     /
%                     ___
%                     m=1
%   PR(fx) =      ------------------------
%                      v
%                     ___
%                     \      ( R^2(n*f0) )
%                     /
%                     ___
%                     n=1
%
%   with u<4 and (u*fx) ≤ 5kHz
%   and v=50 and (v*f0) ≤ 5kHz
%
%   NOTE: fdpowerration assumes that vowels have fundamental harmonics, f0,
%   equalto 100hz and that fx is a multiple of the fundamental harmonic.
%
%   REF: Roger L. Miller et al. Effects of acoustic trauma on the
%   representation of the vowel /e/ in cat auditory nerve fibers.
%   J. Acoust Soc. Am. 1997.

disp('fd_pwrr. Faheem Dinath. May 29th 2008.')

% Make sure that we're looking at fine timing PSTH
if ¬strcmp(psth_struct.type, 'FINE')
    pr = [];
    return
end

f0 = psth_struct.data_struct.fundamental;
f0_har = floor(5000/f0);

fx_har = floor(5000/fx);
if fx_har > 3
    fx_har = 3;
```

```
end


for i = 1:f0_har
    [ freq_f0(i), val ] = max_value_around( psth_struct.data_struct, i*f0);
end

for i = 1:fx_har
    [ freq_fx(i), val ] = max_value_around( psth_struct.data_struct, i*fx);
end

% [PSTH_F, PSTH, PSTH_P] = quickfft( psth, 1/binwidth );
% PSTH = abs(PSTH).^2;    % Looking at the Power!

PSTH_F = psth_struct.F;
PSTH = abs(psth_struct.PSTH).^2;

for i = 1:size(PSTH,1)

    sum_n(i) = 0;
    for j = 1:fx_har
        [m1 m2] = min( abs( PSTH_F - freq_fx(j) ) );
        sum_n(i) = sum_n(i) + PSTH(i,m2);
    end

    sum_d(i) = 0;
    for j = 1:f0_har
        [m1 m2] = min( abs( PSTH_F - freq_f0(j) ) );
        sum_d(i) = sum_d(i) + PSTH(i,m2);
    end

end

pr = sum_n./sum_d;

if (nargout == 0) | strncmpi(varargin,'y',1)
    figure;
    plot(gca,psth_struct.psth_freq,pr);
    drawnow;
    xlabel('Centre Frequencies (Hz)')
    title(['Power Ratio at ' num2str(fx) ' Hz'])
    ylim([0 1]);
end

%=========================================================================%
```

B-8 get_form

```
function [formants] = get_form(data,FS)
% GET_FORM Approximates formants using linear predictive coding.
%
% [formants] = get_form(data,FS) returns the max 3 formants from data.
%
% Modified from Introduction to Computer Programming with MATLAB, Lecture
% 10: Speech Signal Analysis, UCL Department of Phonetics and Linguistics
% http://www.phon.ucl.ac.uk/courses/spsci/matlab/lect10.html

% get Linear prediction filter
ncoeff=2+FS/1000;           % rule of thumb for formant estimation
a=lpc(data,ncoeff);

% find frequencies by root-solving
r=roots(a);                  % find roots of polynomial a
r=r(imag(r)>0.01);           % only look for roots >0Hz up to fs/2
ffreq=sort(atan2(imag(r),real(r))*FS/(2*pi));

% % plot frequency response
[h,f]=freqz(1,a,2^16,FS);
h = 20*log10(abs(h));
% plot(f,h+eps);
% legend('LP Filter');
% xlabel('Frequency (Hz)');
% ylabel('Gain (dB)');


[XMAX,IMAX,XMIN,IMIN] = extrema(h);
formants = sortrows([f(IMAX) XMAX]);

% Remove formants less than 100 Hz
if formants(1,1) < 100
    formants(1,:) = [];
end

formants = formants(1:3,1)';
```

B-9 get_fund

```
function [ Fx ] = get_fund( data, FS )
% GET_FUND Approximates formants using a cross-correlation method.
%
% [formants] = get_fund(data,FS) assumes the fundamental lies between 50 -
% 300 Hz.
%
% Modified from Introduction to Computer Programming with MATLAB, Lecture
% 10: Speech Signal Analysis, UCL Department of Phonetics and Linguistics
% http://www.phon.ucl.ac.uk/courses/spsci/matlab/lect10.html
```

```
[f, X] = newfft(data, FS);
Δ_f = f(end) - f(end-1);
f_low = round(50/Δ_f);   % Fundamental above 50 Hz
f_high = round(300/Δ_f); % Fundamental below 300 Hz

r=xcorr(X,f_high,'coeff');
r=r(f_high+1:2*f_high+1);
d = 0:(f_high);
[fmax,fx]=max(r(f_low:f_high));

if (fmax < 0.5)
    Fx = NaN;
else
    Fx = d(f_low+fx-1)*Δ_f;
end
```

B-10 make_data_struct

```
function [ data_struct ] = make_data_struct( varargin )
% MAKE_DATA_STRUCT makes a data strucure from a speech data file.
%
% make_data_struct(data, FS, spl)
% make_data_struct(data, FS, spl, audiogram_struct, prescription_type)
% where audiogram_struct is the hearing loss profile and prescription_type
% is either 'NAL-R' or 'DSL'.
%
% DATA_STRUCT
% |
% |-> data           : data itself
% |-> FS             : data sampling frequency
% |-> SPL            : data sound pressure level
% |-> fundamental    : fundamental frequency of data
% |-> F              : fourier frequencies of data
% |-> DATA           : fourier coefficients of data
% |-> PHASE          : fourier phase of data
% |-> formants       : predicted Formants
% |-> approx_formants: formants as a multiple of fundamental frequency
%
% See Also audiograms

if length(varargin) < 2
    disp('Error, not enough inputs');
    data_struct = struct([]);
    return;
elseif (length(varargin) == 3) && isnumeric(varargin{1})
    x = varargin{1};
    FS = varargin{2};
```

```
    spl = varargin{3};
    audi_struct = [];
    pres_type = [];
elseif (length(varargin) == 5) && isnumeric(varargin{1}) &&...
        isstruct(varargin{4}) && ischar(varargin{5})
    x = varargin{1};
    FS = varargin{2};
    spl = varargin{3};
    audi_struct = varargin{4};
    pres_type = varargin{5};
else
    disp('Check your inputs!')
    data_struct = struct([]);
    return;
end

data_struct.data = set_spl(x(:)',spl);
data_struct.FS = FS;
data_struct.SPL = spl;
data_struct.fundamental = get_fund(x,FS);
data_struct.formants = get_form(x,FS);
[data_struct.F, data_struct.DATA, data_struct.PHASE ] = ...
    quickfft( data_struct.data, data_struct.FS );
data_struct.approx_formants = ...
    round(data_struct.formants/data_struct.fundamental)*data_struct.fundamental;

if isstruct(audi_struct) && ischar(pres_type)
    data_struct = ampl_pres( data_struct, audi_struct, pres_type );
end
```

B-11 make_psth_struct

```
function psth = make_psth_struct(data, FS, spl, loss, pres, CFcount,...
    IOHC_loss, binwidth)
audi = audiograms(loss, CFcount, IOHC_loss);
data = make_data_struct( data, FS, spl, audi, pres );
psth = PSTHmay(data, audi, binwidth, 'n');
end
```

B-12 psth_err_mean

```
function [ error ] = psth_err_mean( psth_struct1, psth_struct2, varargin )
% PSTH_ERR_MEAN reads two psth structures and returns the error metrics
% between them using mean absolute error.
%
% error = psth_err_mean( psth_struct1, psth_struct2, varargin )
%
```

```
% varargin can either be 'make' or the output error variable itself. On
% first run varargin has to be set to 'make' and on subsiquent runs the
% output variable. Each time psth_err_mean is called, it concatinates the
% error values on the error variable.

if strcmp(psth_struct1.type, 'FINE') && strcmp(psth_struct2.type, 'FINE')
    if (nargin == 2)
        disp('Not enough input variables')
    elseif strcmp(varargin{1}, 'make')
        error = struct('phsr', [], 'pwrr', [], 'psth', [], 'boxp', [],...
            'SPL', [], 'SPL_uniq', [], 'ADJ', [], 'phsr_opti', [],...
            'pwrr_opti', [], 'psth_opti', [], 'boxp_opti', []);
        error.approx_formants = psth_struct1.data_orig.approx_formants;
        error.type = psth_struct1.type;
        error.meth = 'mean';
        return
    else
        error = varargin{1};
    end

    % Make the ranges of phase the same %
    [ phsr_freq1, phsr_freq2, phsr1, phsr2 ] = ...
        same_phsr_range( psth_struct1, psth_struct2 );

    for i = 1:length(psth_struct1.phsr)
        phsr(i)=mean(abs(phsr1{i} - phsr2{i}));
        pwrr(i)=mean(abs(psth_struct1.pwrr(i,:) - psth_struct2.pwrr(i,:)));
    end

    error.phsr = [error.phsr phsr'];
    error.pwrr = [error.pwrr pwrr'];
    error.psth = [error.psth ...
        mean(abs(psth_struct1.psth(:) - psth_struct2.psth(:)))];
    error.boxp = [error.boxp ...
        mean(abs(psth_struct1.boxp(:) - psth_struct2.boxp(:)))];
    error.ADJ  = [error.ADJ ...
         (psth_struct2.data_orig.SPL - psth_struct1.data_orig.SPL)];
    error.SPL  = [error.SPL ...
        psth_struct1.data_orig.SPL];

    [error.phsr_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.phsr, error.ADJ, 'min' );
    [error.pwrr_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.pwrr, error.ADJ, 'min' );
    [error.psth_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.psth, error.ADJ, 'min' );
    [error.boxp_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.boxp, error.ADJ, 'min' );
```

```
elseif strcmp(psth_struct1.type, 'AVG') && strcmp(psth_struct2.type, 'AVG')
    if (nargin == 2)
        disp('Not enough input variables')
    elseif strcmp(varargin{1}, 'make')
        error = struct('hist', [], 'psth', [], 'SPL', [], ...
            'SPL_uniq', [], 'ADJ', [], 'hist_opti', [], 'psth_opti', []);
        error.approx_formants = psth_struct1.data_orig.approx_formants;
        error.type = psth_struct1.type;
        error.meth = 'mean';
        return
    else
        error = varargin{1};
    end

    error.hist = [error.hist ...
        mean(abs(psth_struct1.hist(:) - psth_struct2.hist(:)))];
    error.psth = [error.psth ...
        mean(abs(psth_struct1.psth(:) - psth_struct2.psth(:)))];
    error.ADJ  = [error.ADJ ...
        (psth_struct2.data_orig.SPL - psth_struct1.data_orig.SPL)];
    error.SPL  = [error.SPL psth_struct1.data_orig.SPL];

    [error.hist_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.hist, error.ADJ, 'min' );
    [error.psth_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.psth, error.ADJ, 'min' );
else
    disp('Unknown Structure Type');
end
```

B-13 psth_err_xcorr

```
function [ error ] = psth_err_xcorr( psth_struct1, psth_struct2, varargin )
% PSTH_ERR_XCORR reads two psth structures and returns the error metrics
% between them using cross correlation at zero lag.
%
% error = psth_err_mean( psth_struct1, psth_struct2, varargin )
%
% varargin can either be 'make' or the output error variable itself. On
% first run varargin has to be set to 'make' and on subsiquent runs the
% output variable. Each time psth_err_xcorr is called, it concatinates the
% error values on the error variable.

if strcmp(psth_struct1.type, 'FINE') && strcmp(psth_struct2.type, 'FINE')
    if (nargin == 2)
        disp('Not enough input variables')
    elseif strcmp(varargin{1}, 'make')
        error = struct('phsr', [], 'pwrr', [], 'psth', [], 'boxp', [],...
```

```
            'SPL', [], 'SPL_uniq', [], 'ADJ', [], 'phsr_opti', [],...
            'pwrr_opti', [], 'psth_opti', [], 'boxp_opti', []);
        error.approx_formants = psth_struct1.data_orig.approx_formants;
        error.type = psth_struct1.type;
        error.meth = 'xcorr';
        return
    else
        error = varargin{1};
    end

    % Make the ranges of phase the same %
    [ phsr_freq1, phsr_freq2, phsr1, phsr2 ] = ...
        same_phsr_range( psth_struct1, psth_struct2 );

    for i = 1:length(psth_struct1.phsr)
        phsr(i) = xcorr(phsr1{i}, phsr2{i},0,'coeff');
        pwrr(i) = xcorr(psth_struct1.pwrr(i,:),...
            psth_struct2.pwrr(i,:),0,'coeff');
    end
    phsr(isnan(phsr)) = 0;
    error.phsr = [error.phsr phsr'];
    error.pwrr = [error.pwrr pwrr'];
    error.psth = ...
        [error.psth xcorr(psth_struct1.psth(:),...
        psth_struct2.psth(:),0,'coeff')];
    error.boxp = ...
        [error.boxp xcorr(psth_struct1.boxp(:),...
        psth_struct2.boxp(:),0,'coeff')];
    error.ADJ  = [error.ADJ ...
        (psth_struct2.data_orig.SPL - psth_struct1.data_orig.SPL)];
    error.SPL  = [error.SPL psth_struct1.data_orig.SPL];

    [error.phsr_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.phsr, error.ADJ, 'max' );
    [error.pwrr_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.pwrr, error.ADJ, 'max' );
    [error.psth_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.psth, error.ADJ, 'max' );
    [error.boxp_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.boxp, error.ADJ, 'max' );

elseif strcmp(psth_struct1.type, 'AVG') && strcmp(psth_struct2.type, 'AVG')
    if (nargin == 2)
        disp('Not enough input variables')
    elseif strcmp(varargin{1}, 'make')
        error = struct('hist', [], 'psth', [], 'SPL', [],...
            'SPL_uniq', [], 'ADJ', [], 'hist_opti', [], 'psth_opti', []);
        error.approx_formants = psth_struct1.data_orig.approx_formants;
        error.type = psth_struct1.type;
```

```
        error.meth = 'xcorr';
        return
    else
        error = varargin{1};
    end

    error.hist = [error.hist xcorr(psth_struct1.hist(:),...
        psth_struct2.hist(:),0,'coeff')];
    error.psth = [error.psth xcorr(psth_struct1.psth(:),...
        psth_struct2.psth(:),0,'coeff')];
    error.ADJ  = [error.ADJ ...
        (psth_struct2.data_orig.SPL - psth_struct1.data_orig.SPL)];
    error.SPL  = [error.SPL psth_struct1.data_orig.SPL];

    [error.hist_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.hist, error.ADJ, 'max' );
    [error.psth_opti error.SPL_uniq] = ...
        find_mxmn( error.SPL, error.psth, error.ADJ, 'max' );
else
    disp('Unknown Structure Type');
end
```

C Additional Figures for Section 5.5.1

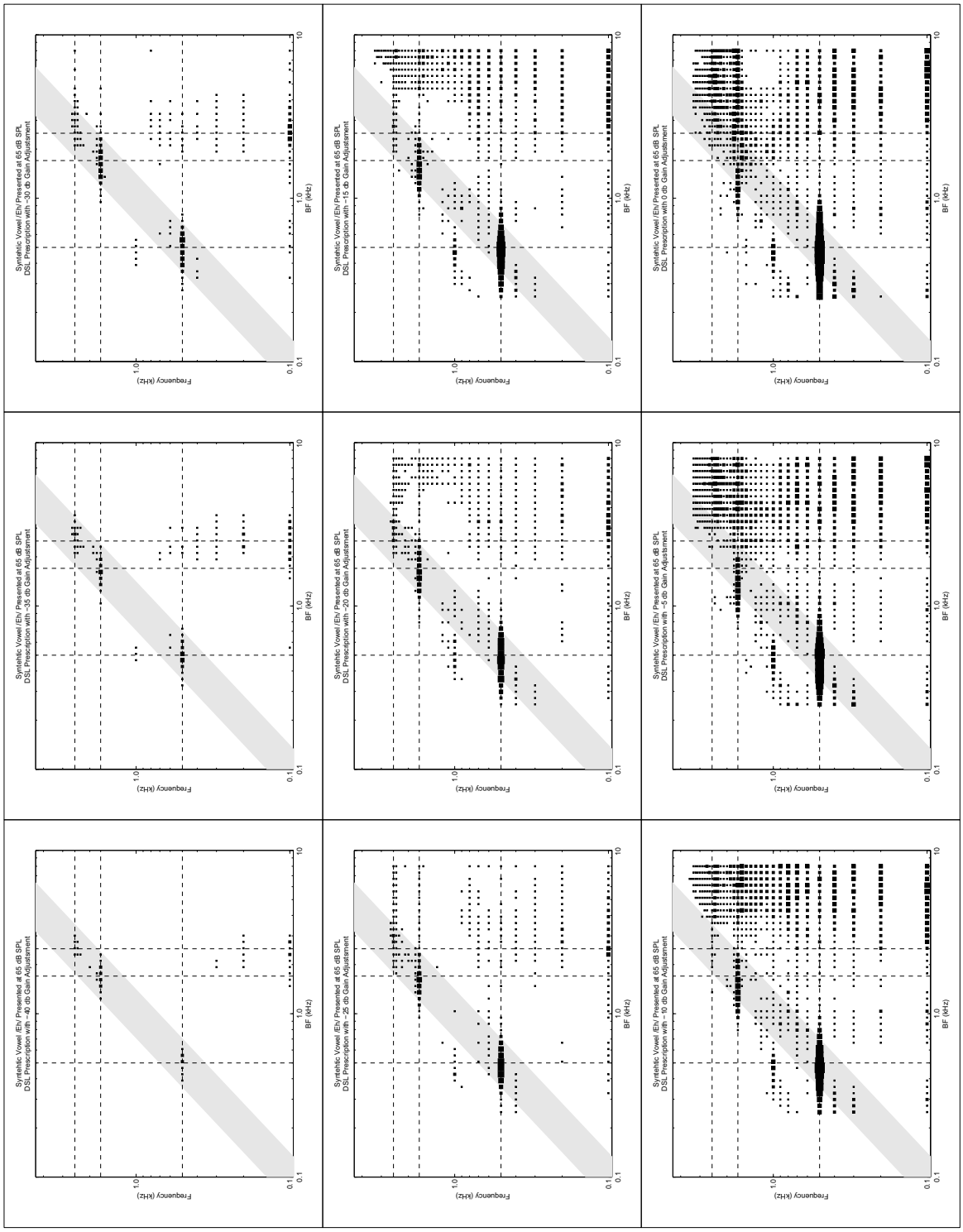


Figure 7.1: Box plots from a mildly impaired auditory periphery of an /ɛ/ vowel with both DSL amplification and adjustment gains from -40 to 0 db applied. Figures are rotated 90 deg counter clockwise.

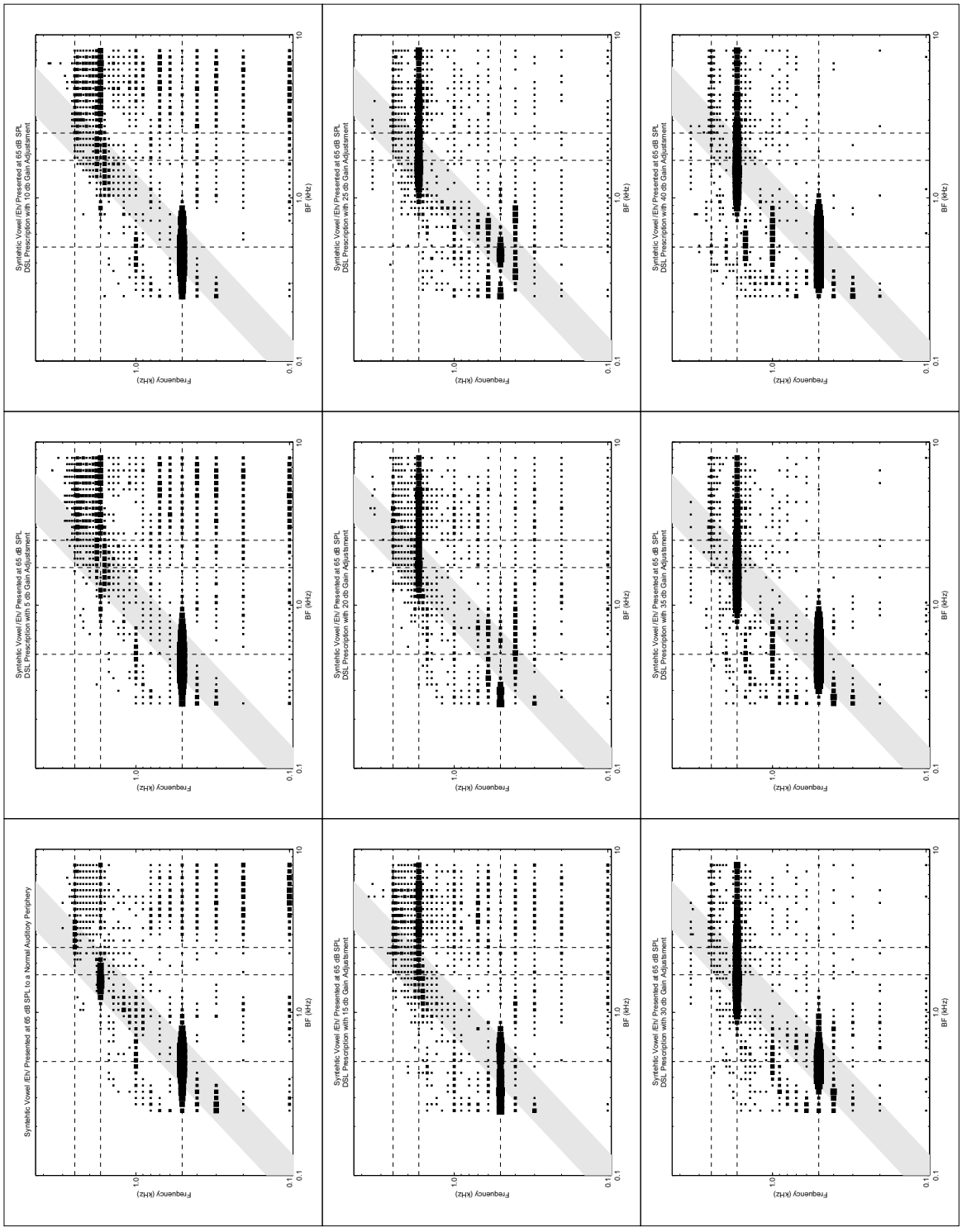


Figure 7.2: Box plots from a mildly impaired auditory periphery of an /ε/ vowel with both DSL amplification and adjustment gains from 5 to 40 db applied. Figures are rotated 90 deg counter clockwise. Normal Box Plot show on top left.

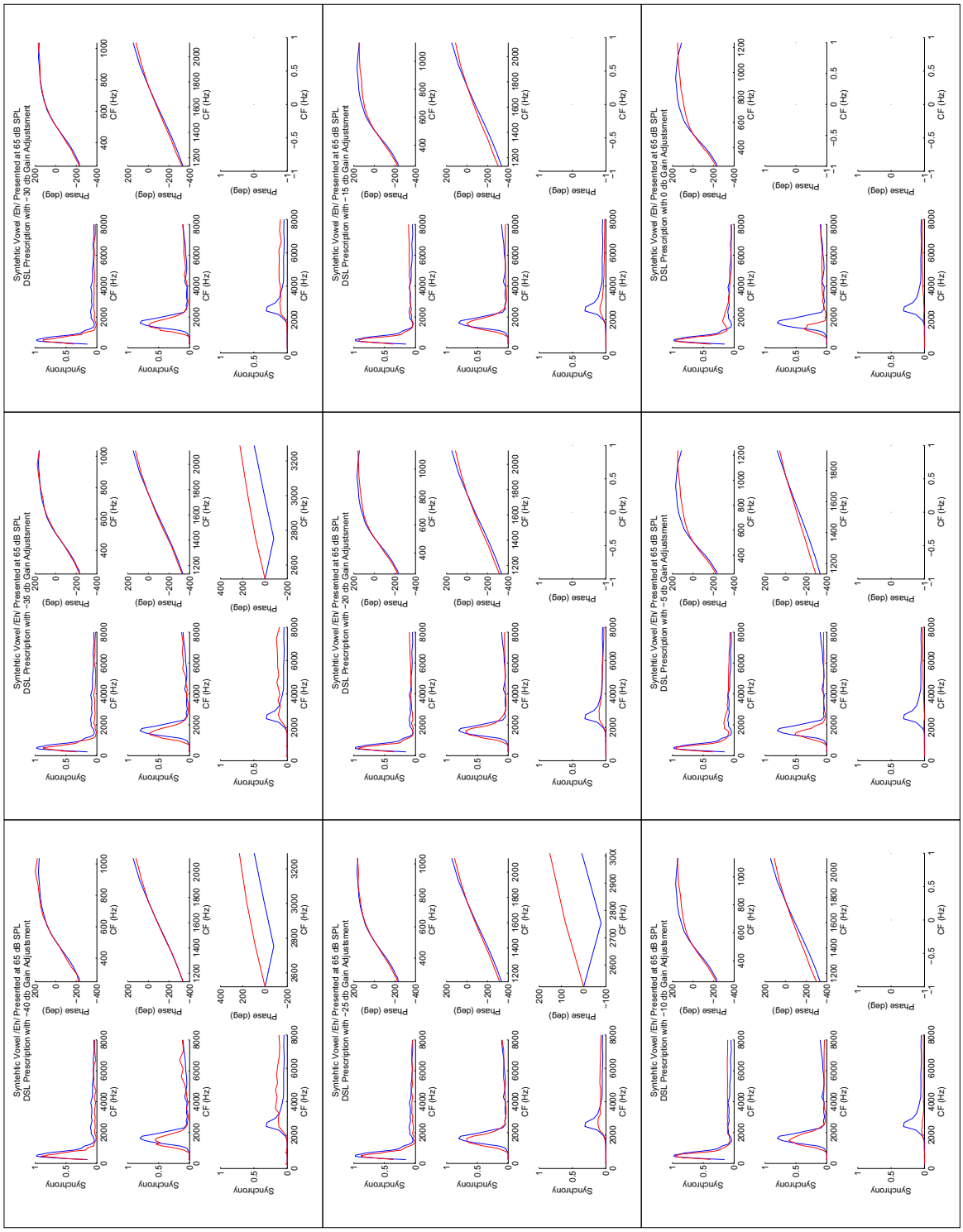


Figure 7.3: Power ratios and phase responses from a mildly impaired AP (red) of an /ε/ vowel with both DSL amplification and adjustment gains from -40 to 0 db applied. Normal AP response is shown in blue. Responses are taken at formants 500, 1700, and 2500 Hz.

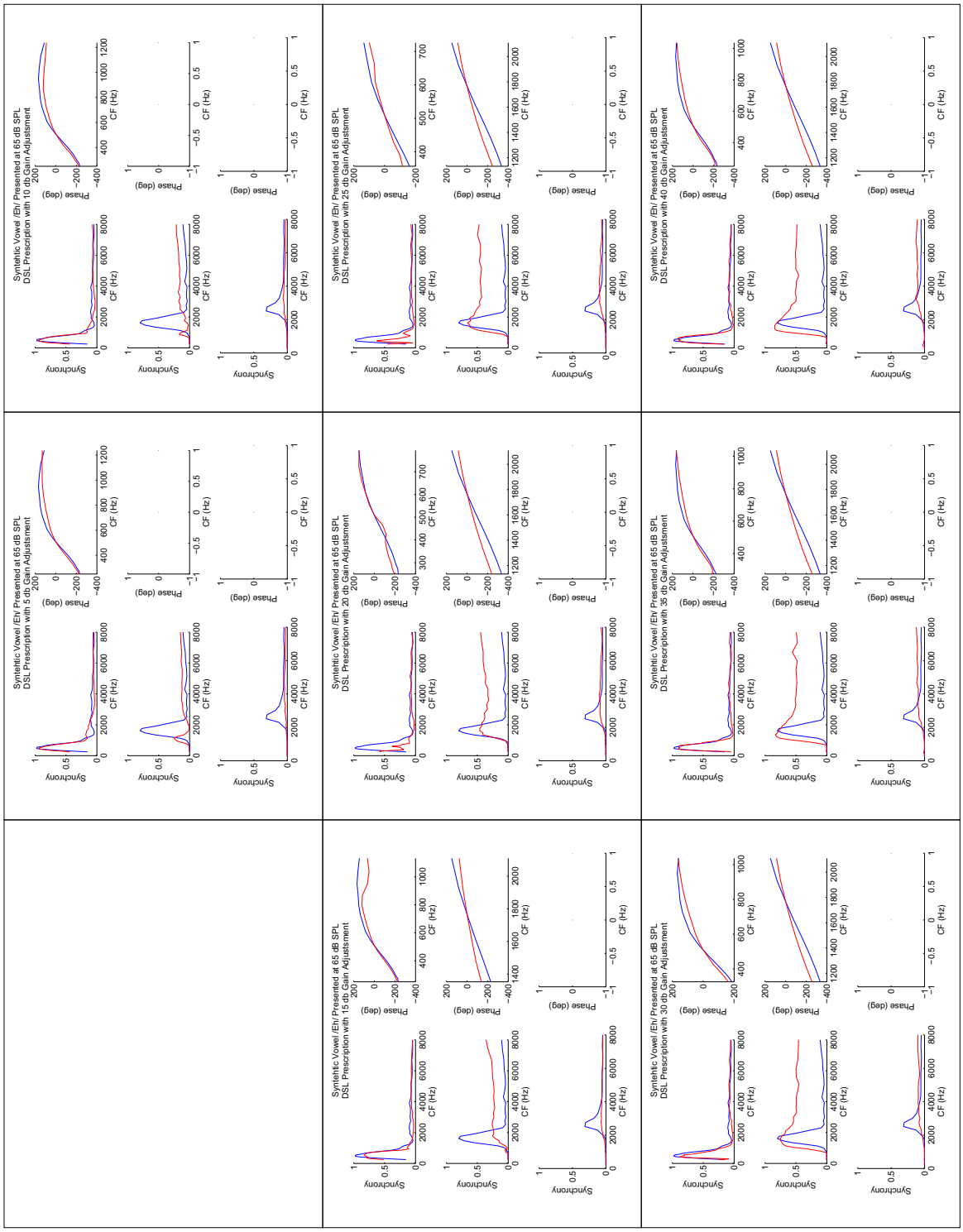


Figure 7.4: Power ratios and phase responses from a mildly impaired AP (red) of an /ɛ/ vowel with both DSL amplification and adjustment gains from 5 to 40 db applied. Normal AP response is shown in blue. Responses are taken at formants 500, 1700, and 2500 Hz.

D Additional Figures for Section 5.7.3

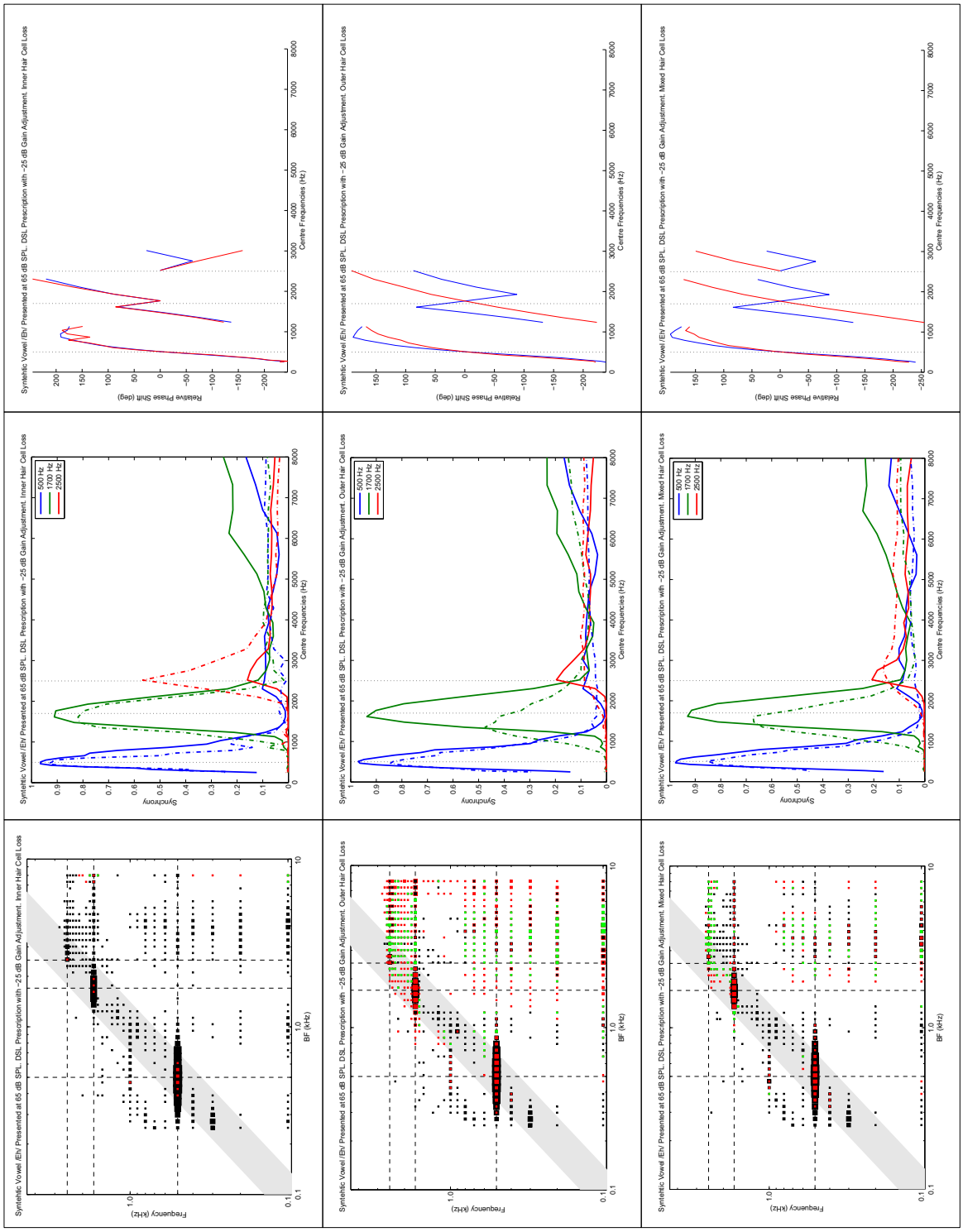


Figure 7.5: Sample optimizations with DSL and -25 db gain adjustment. Box Plots: normal response in black, impaired in red, and similarities in green. Power Ratios: Impaired responses in dashed lines and normal in solid. Phase responses: normal response in blue and impaired in red.

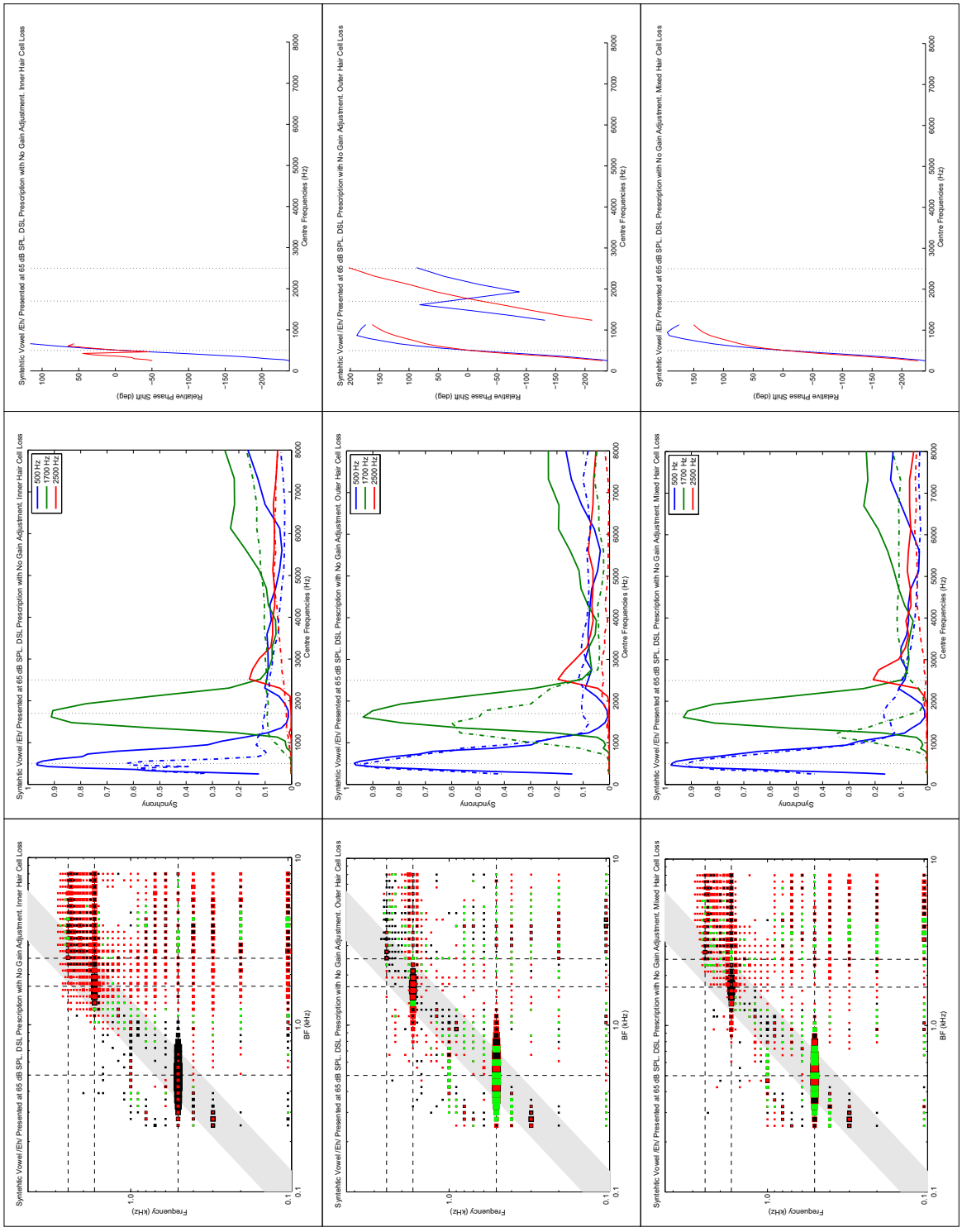


Figure 7.6: DSL prescription with no additional gain adjustment.

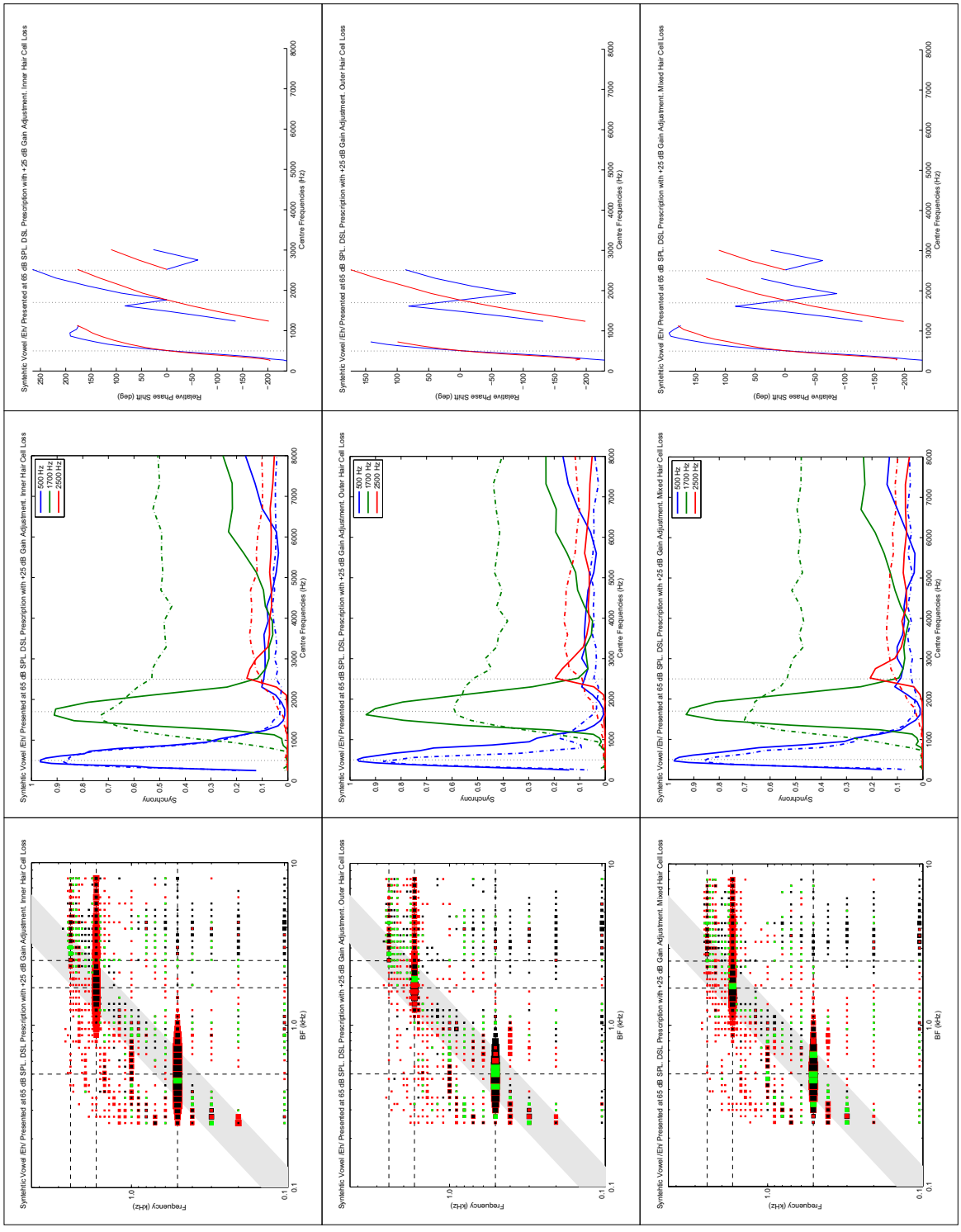


Figure 7.7: Sample optimizations with DSL and +25 db gain adjustment.

Bibliography

British Society of Audiology (1988). "British society of audiology recommendation," British Journal of Audiology **22**, 123.

Bruce, I. C. (2007). "Hearing aids," in *Encyclopedia of Biomaterials and Biomedical Engineering*, edited by G. L. Bowlin and G. Wnek (Taylor & Francis, Oxfordshire, UK).

Bruce, I. C., Sachs, M. B., and Young, E. D. (2003). "An auditory-periphery model of the effects of acoustic trauma on auditory nerve responses," J. Acoust. Soc. Am. **113**, 369–388.

Byrne, D. and Dillon, H. (1986). "The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid," Ear and Hearing **7**, 257–265.

Byrne, D. and Tonisson, W. (1976). "Selecting the gain of hearing aids for persons with sensorineural hearing impairments.," International Journal of Audiology **5**, 5159.

Carney, L. H. (1993). "A model for the responses of low-frequency auditory-nerve fibers in cat," The Journal of the Acoustical Society of America **93**, 401–417.

Carney, L. H. (1994). "Spatiotemporal encoding of sound level: models for normal encoding and recruitment of loudness.," Hearing Research **76**, 3144.

Dillon, H. (2001). *Hearing Aids* (Thieme Medical Publishers, New York, NY).

Drake, R., Vogl, W., and Mitchell, A. (2004). *Gray's Anatomy for Students* (Churchill Livingstone, Oxfordshire, UK), 1ed ed.

Eggermont, J. and Spoor, A. (1973). "Cochlear adaptation in guinea pigs. a quantitative description.," Audiology **12**, 193–220.

Fettiplace, R. (2006). "Active hair bundle movements in auditory hair cells," J. Physiol. **576**, 29–36.

Garner, W. R. (1950). "Reviewed work, theory of hearing by earnest glen wever," The Quartery Revew of Biology **25**, 101–103.

Gelfand, S. A. (1998). *Hearing: an introduction to psychological and physiological acoustics* (Marcel Dekker, New York, NY), 3ed ed.

Haenggeli, A., Zhang, J., Vischer, M., Pelizzone, M., and Rouiller, E. (1998). "Electrically evoked compound action potential (ecap) of the cochlear nerve in response to pulsatile electrical stimulation of the cochlea in the rat: effects of stimulation at high rates.," Audiology **37**, 353–371.

Heinz, M. G., Issa, J. B., and Young, E. D. (2005). "Auditory-nerve rate responses are inconsistent with common hypotheses for the neural correlates of loudness recruitment," JARO: Journal of the Association for Research in Otolaryngology **6**, 91105.

Henderson, D. and Hamernik, R. P. (1995). "Biologic bases of noise-induced hearing loss," Occup Med **10**, 513–534.

Javel, E. (1996). "Long-term adaptation in cat auditory-nerve fiber responses.," jasa **99**, 1040–1052.

Kamm, C., Dirks, D., and Mickey, M. (1978). "Effect of sensorineural hearing loss on loudness discomfort level and most comfortable loudness.," Journal of Speech and Hearing Research **21**, 668–681.

Kiang, N. Y., Liberman, M. C., Sewell, W. F., and Guinan, J. J. (1986). "Single unit clues to cochlear mechanisms.," Hearing Research **22**, 171–182.

Killian, M., Klis, S., and Smoorenburg, G. (1994). "Adaptation in the compound action potential response of the guinea pig viiith nerve to electric stimulation.," Hear. Res. **81**, 66–82.

Liberman, M. C. and Kiang, N. Y.-S. (1978). "Acoustic trauma in cats: Cochlear pathology and auditory-nerve activity.," Acta. Otolaryngol. Suppl. **358**, 1–63.

Lybarger, S. F. (1944). "U.s. patent application sn 532,278," US Patent.

Lybarger, S. F. (1978). "Selective amplification–a review and evaluation," J. Am. Audiol. Soc. **3**, 258–266.

Miller, R. L., Schilling, J. R., Franck, K. R., and Young, E. D. (1997). "Effects of acoustic trauma on the representation of the vowel /ɛ/ in cat auditory nerve fibers," J. Acoust. Soc. Am. **101**, 3602–3616.

Netter, F. H. (2002). *Atlas of Neuroanatomy and Neurophysiology* (Icon Custom Communications, Teterboro, NJ).

Nourski, K., Abbas, P., and Miller, C. (2006), "Effects of remaining hair cells on cochlear implant function," Technical report, (University of Iowa, Iowa City, Iowa, USA). 15th Quarterly Progress Report, Neural Prosthesis Program.

Oxenham, A. J. and Bacon, S. P. (2003). "Cochlear compression: Perceptual measures and implications for normal and impaired hearing," Ear and Hearing **24**, 352366.

Patterson, G. E. and Barney, H. L. (1952). "Control methods used in a study of vowels," jasa **24**, 175–184.

Reale, R. A. and Imig, T. J. (2004). "Tonotopic organization in auditory cortex of the cat," The Journal of Comparative Neurology **192**, 265–291.

Roberts, D. (2002). *Signals and Perception: The Fundamentals of the Human Senses* (Palgrave Macmillan, Houndmills, UK), 1ed ed.

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

Sachs, M. B., Bruce, I. C., Miller, R. L., and Young, E. D. (2002). "Biological basis of hearing-aid design," Annals of Biomedical Engineering **30**, 157–168.

Seewald, R. C., Ross, M., and Spiro, M. K. (1985). "Selecting amplification characteristics for young hearing-impaired children.," Ear and Hearing **6**, 48–53.

Shaw, E. A. G. (1974). "The external ear," in *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff (Springer, Berlin), Vol. 5/1, pp. 455–490.

Smith, R. (1977). "Short-term adaptation in single auditory nerve fibers: some post-stimulatory effects," Journal of Neurophysiology **40**, 1098–1111.

von Bksy, G. (1960). *Experiments in Hearing* (McGraw-Hill, New York, NY).

Warren, R. M. (1999). *Auditory Perception: A New Analysis and Synthesis* (Cambridge University Press, New York, NY).

Westerman, L. and Smith, R. (1984). "Rapid and short-term adaptation in auditory nerve responses.," Hear Res. **15**, 249–60.

Wiener, F. M. and Ross, D. A. (1946). "The pressure distribution in the auditory canal in a progressive sound field," J. Acoust. Soc. Am. **18**, 401–408.

Yarmush, M. L., Toner, M., Plonsey, R., and Bronzino, J. D. (2003). *Biotechnology for Biomedical Engineers* (CRC, New York, NY).

Young, E. D. and Sachs, M. B. (1979). "Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers," J. Acoust. Soc. Am. **66**, 1381–1403.

Zilany, M. S. A. and Bruce, I. C. (2006). "Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery," J. Acoust. Soc. Am. **120**, 1446–1466.

Zilany, M. S. A. and Bruce, I. C. (2007). "Representation of the vowel /ε/ in normal and impaired auditory nerve fibers: Model predictions of responses in cats," J. Acoust. Soc. Am. **122**, 402–417.