Fall 2013

Effects of Hearing Aid Amplification on Robust Neural Coding of Speech

Jonathan Daniel Boley
Purdue University

Graduate School ETD Form 9
(Revised 12/07)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Jonathan Daniel Boley

Entitled
Effects of Hearing Aid Amplification on Robust Neural Coding of Speech

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Michael Heinz
Chair

Thomas Talavage

Joshua Alexander

Edward Bartlett

Donna Fekete

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Michael Heinz

Approved by: George Wodicka 11/25/2013
Head of the Graduate Program Date

EFFECTS OF HEARING AID AMPLIFICATION ON ROBUST NEURAL CODING OF SPEECH

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Jonathan Daniel Boley

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2013

Purdue University

West Lafayette, Indiana

Dedicated to

11-03, 11-19, 11-20, 11-24, 11-25,

11-27, 11-29, 11-30, 11-31, 11-32,

11-33, 11-34, 11-35, 11-36, 11-37,

11-39, 11-44, 11-45, 11-46, 11-48,

11-49, 11-51, 12-01, 12-02, 12-04,

12-06, 12-07, 12-08, 12-17, 12-21,

12-22, 12-25, 12-28, 12-29

"One year in Italy with their eyes open would be worth more than three at Oxford; and six months in the fields with a platyscopic lens would teach them strange things about the world around them that all the long terms at Harrow and Winchester have failed to discover to them. But that would involve some trouble to the teacher. What a misfortune it is that we should thus be compelled to let our boys' schooling interfere with their education!" - Grant Allen (published 1894 in *Post-Prandial Philosophy*, p 129)

ACKNOWLEDGEMENTS

Many people helped make this dissertation possible. Michael Heinz gave me the opportunity to study neural coding in his lab, encouraged me to seek an NIH fellowship, and taught me much more than I could regurgitate in this document. Josh Alexander, Ed Bartlett, Donna Fekete, and Tom Talavage have provided useful feedback along the way. Jayaganesh Swaminathan, Sushrut Kale, Ananthakrishna Chintanpalli, Ken Henry, Mark Sayles, and Ann Hickox have provided useful insights during many fruitful conversations at lab meetings, conferences, random hallway conversations, etc. Michael Walls, Sushrut Kale, Jayaganesh Swaminathan, and Ken Henry have helped tremendously by helping with data collection, which was absolutely grueling work. Vidya Krull, Mark Sayles, and Michael Heinz provided very helpful feedback on early versions of this dissertation. It is a much better document due to their feedback.

I am grateful to have had the opportunity to work for GN ReSound for the past couple years as I finished this research, and I thank Andrew Dittberner, Michael Heinz, Andrew Brightman, and Donna Fekete for helping to make that happen. GN provided financial support during this time, but I am also grateful to the Biomedical Engineering department at Purdue for providing a fellowship for my first couple years, and to the National Institutes of Health for providing another fellowship (F31-DC010966).

Of course, my parents were very supportive. My dad instilled in me a love for learning early on, and taught me that learning should never end. My mom has always been very supportive and encouraging. Chip and Rhonda have also been very supportive and encouraging. I could not ask for better in-laws. My wife, Kelley, has by far provided the most support and encouragement, and put up with my academic endeavors for far too long.

TABLE OF CONTENTS

Page

Page

LIST OF FIGURES

Page

Page

ABSTRACT

Boley, Jonathan D. Ph.D., Purdue University, December 2013. Effects of Hearing Aid Amplification on Robust Speech Coding. Major Professor: Michael Heinz.

Hearing aids are able to restore some hearing abilities for people with auditory impairments, but background noise remains a significant problem. Unfortunately, we know very little about how speech is encoded in the auditory system, particularly in impaired systems with prosthetic amplifiers. There is growing evidence that relative timing in the neural signals (known as spatiotemporal coding) is important for speech perception, but there is little research that relates spatiotemporal coding and hearing aid amplification.

This research used a combination of computational modeling and neurophysiological experiments to characterize how hearing aids affect vowel coding in noise at the level of the auditory nerve. The results indicate that sensorineural hearing impairment degrades the temporal cues transmitted from the ear to the brain. Two hearing aid strategies (linear gain and wide dynamic-range compression) were used to amplify the acoustic signal. Although appropriate gain was shown to improve temporal coding for individual auditory nerve fibers, neither strategy improved spatiotemporal cues. Previous work has attempted to correct the relative timing by adding frequency-

dependent delays to the acoustic signal (e.g., within a hearing aid). We show that, although this strategy can affect the timing of individual auditory nerve responses, there is a fundamental limitation in the ability of this approach to improve the relative across-fiber timing (spatiotemporal coding) as intended.

We have shown that existing hearing aid technologies do not improve some of the neural cues that we think are important for perception, but it is important to understand these limitations. Our hope is that this knowledge can be used to develop new technologies to improve auditory perception in difficult acoustic environments.

CHAPTER 1. INTRODUCTION

The World Health Organization estimates that 360 million people have a disabling hearing loss (Stevens and Flaxman, 2013). These impairments can severely limit communication, and have been shown to reduce quality of life (Mulrow et al., 1990). However, experts suggest that regular use of a prescribed hearing aid can significantly reduce depression in addition to improving communication, cognitive function, and social and emotional well-being (Mulrow et al., 1992).

Despite the benefits, approximately half of hearing aid users still have difficulty listening in noisy environments (Edwards, 2007). In fact, listening in noise can be a very complex task and the best prosthetic hearing instruments today cannot restore a patient's listening abilities to normal. The research described here seeks to better understand some of the physiological reasons why hearing aids remain limited in their ability to restore normal speech perception in noise. We have used a combined computational and neurophysiological approach to address this issue, and have developed some computational modeling techniques for evaluating and fitting hearing aids in a quantitative manner.

1.1 Background

1.1.1 The Auditory System

The mammalian auditory system can be roughly divided into two interdependent systems: the peripheral and central systems. The peripheral auditory system includes everything from the outer ear to the auditory nerve, whereas the central auditory system includes all the brain structures that process auditory information.

The peripheral auditory system converts acoustic energy to electrical energy and can handle a surprisingly wide range of sound levels. For example, the loudest sound a human can hear without pain is approximately one million times the pressure of the softest audible sound (a range of 120 dB).

1.1.1.1 Gain Control

The auditory system uses a complex system of gain controls to operate over such a large dynamic range. An acoustic reflex controls the transmission of sounds through the middle ear (Møller, 1964), but another set of control systems modulates gain within the cochlea. In this physiological "algorithm", there are three primary sources of active cochlear gain adjustment, as shown in Figure 1.1. The outer hair cells (OHC) provide the first stage of gain control, reacting to the acoustic stimulus by changing length, thus amplifying the vibration of the organ of Corti (Brownell et al., 1985; Liberman et al., 2002). As Rhode (1971) showed in his seminal work, the gain of the cochlear amplifier is compressive (less gain for high input sound levels). The gain provided by the outer hair cells is also frequency dependent – it may be as high as 60 dB at the base of the cochlea

(in response to low-level high frequency stimuli), with little to no gain at the apex for low-frequency stimuli (Ruggero et al., 1997). Recio and colleagues (Recio et al., 1998) showed that compression could be seen in as little as 100 µs. Compression this fast would normally introduce severe distortion but, because the gain is applied to only a very localized region, any distortion is band-limited by cochlear filtering and would most likely be imperceptible.

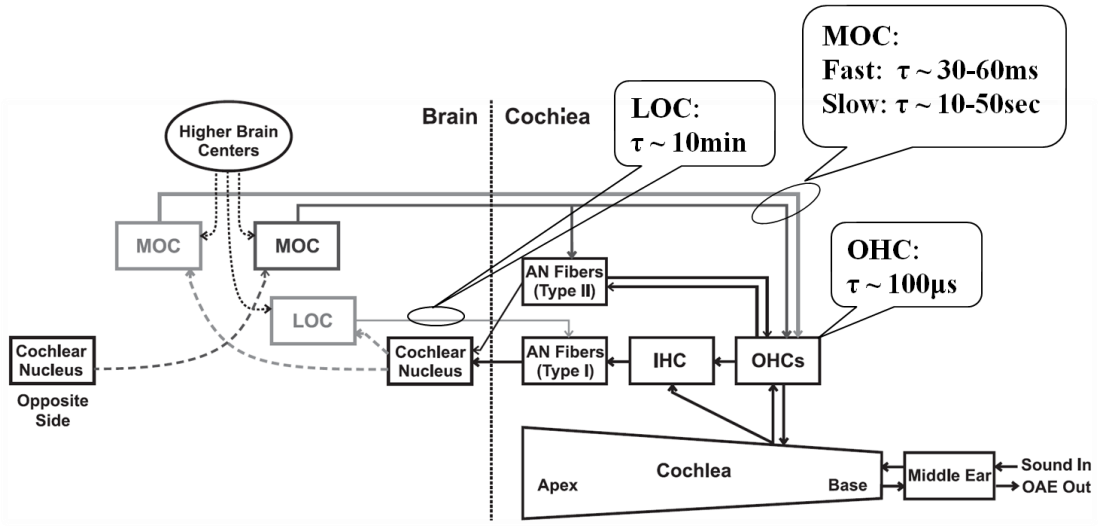

Figure 1.1 Physiological gain structure and time constants
Three primary gain control mechanisms (LOC – Lateral Olivo-Cochlear efferents, MOC – Medial Olivo-Cochlear efferents, OHC – Outer Hair Cells) are shown along with their respective time constants. [Adapted with permission from Lippincott Williams and Wilkins/Wolters Kluwer Health: Ear & Hearing (Guinan, 2006) , copyright 2006]

The second source of cochlear gain control is the Medial Olivo-Cochlear (MOC) system (see Figure 1.1). Cells near the medial superior olive in the brainstem project axons back into the cochlea and innervate the outer hair cells. This reflex is controlled by both ears (Guinan, 2006) and is known to have two distinct time courses. The fast

effect has a time constant of 30-60 ms, while the slow effect has a time constant closer to 10-50 seconds (Cooper and Guinan Jr., 2003). This moderately fast gain control mechanism appears to be, in part, a protective mechanism for high-intensity sounds (Maison and Liberman, 2000), and some scientists have suggested that the MOC system may serve to increase the signal-to-noise ratio in noisy conditions (Hienz et al., 1998). However, the strength of the MOC reflex appears to vary substantially from person to person, even among those with normal hearing (Backus and Guinan Jr., 2007).

A third and slower gain control system, the Lateral Olivo-Cochlear (LOC) system, uses a set of efferent fibers that come from the lateral superior olive. These cells receive signals from both ears and innervate the ipsilateral auditory nerve fibers. LOC efferents appear to be useful for slowly (τ ~10 min) balancing the output of the two ears, based on interaural level differences (Darrow et al., 2006; Groff and Liberman, 2003). There may also be some efferent control from higher-level brain structures (Mulders and Robertson, 2002), but little is known about such pathways.

1.1.1.2 Spectral Decomposition

As alluded to in the previous section, outer hair cells actively control basilar membrane vibration over a limited frequency range (Ruggero and Rich, 1991). In fact, the outer hair cells increase the frequency-dependent vibration that occurs in the cochlea due to the mechanical properties of the basilar membrane (Békésy and Bekesy, 1952). Even in a passive (dead or badly injured) cochlea, the basilar membrane resonates to high frequencies at the base and low frequencies at the apex. Inner hair

cells at various locations along the length of the cochlea therefore tend to transduce energy within a narrow frequency range. Information within auditory nerve fibers therefore reflects this tonotopicity.

1.1.2 Hearing Impairment

Hearing loss is often characterized by an inability to hear low intensity sounds. The audiogram is often used to describe the hearing loss, quantified as the behavioral threshold shift (relative to normal young subjects) for tones at various frequencies. Unfortunately, hearing impairment is not always as simple as an inability to detect low-intensity sounds, but the audiogram is the most common tool for diagnosing impaired hearing. (In this dissertation, "hearing loss" will refer to a simple audiometric threshold shift, whereas "hearing impairment" is meant to be more general. It is theoretically possible for two people to have identical audiograms, but different degrees of impairment.)

Hearing impairment can affect several auditory percepts, including loudness, pitch, localization, and speech perception (for a review, see Moore, 2007). One important aspect of hearing impairment is broadened tuning. Spectral tuning can be measured behaviorally with *psychophysical tuning curves* (PTCs), which are measured by holding a target sound at a low level (e.g., 10dB above threshold) and determining the level and frequency of a masker signal needed to mask the target (Zwicker, 1974). With normal hearing, the PTC usually has the shape of a narrow 'V', but at frequencies with increased

thresholds, this sharp tuning is absent (Leshowitz, 1975, 1976). This indicates that the frequency resolution of the auditory system is degraded with impairment.

Spectral tuning can also be measured physiologically in experimental animals by measuring the threshold of an auditory nerve fiber at different frequencies. The frequency which the fiber is most sensitive to is called the best frequency (BF), but this can change with level and/or impairment so we often refer to the characteristic frequency (CF), which is equivalent to the BF for a normal system at low levels. The CF does not change with level or impairment. Similar to PTCs, the bandwidth of a neural tuning curve is often characterized by the bandwidth 10dB above threshold. The bandwidth can be normalized by the BF to obtain a "quality factor", referred to as Q_{10}.

Liberman and Dodds (1984a) showed that damaged outer hair cells are associated with broad neural tuning curves and elevated thresholds, while damaged inner hair cells are associated with only elevated auditory nerve thresholds (i.e., without broadened tuning). Hearing loss associated with impaired hair cells (often due to noise exposure and/or aging) is referred to as sensorineural hearing loss (SNHL). The typical perceptual model of SNHL assumes that most of the impairment is due to damaged outer hair cells that normally amplify low intensity sounds but apply less gain to sounds that are already high intensity (Moore and Glasberg, 2004). Note, however, that this model may underestimate the contribution of inner hair cell dysfunction to behavioral threshold shifts (Moore and Glasberg, 2004; Schuknecht, 1993).

Damage to the cochlea results in several changes to the auditory nerve responses, including increased threshold, shifted best frequency, reduced spontaneous rate,

broadened tuning, and abnormal rate-level functions (Heinz and Young, 2004; Kiang et al., 1976; Liberman and Dodds, 1984a; Liberman and Kiang, 1984; Wang et al., 1997). Although within-fiber phase locking to tones has been reported to be degraded following sensorineural hearing loss (Woolf et al., 1981), most evidence suggests phase locking remains strong to tones in quiet (Harrison and Evans, 1979; Heinz et al., 2010; Miller et al., 1997) and only degrades in the presence of background noise (Henry and Heinz, 2012).

Sounds may be encoded in several ways. The relative firing rates across different cochlear positions (*rate-place coding*) can provide information about auditory stimuli. Similarly, differences in the strength and/or frequency of phase locking across different cochlear positions (*temporal-place coding*) can provide different information about auditory stimuli. Differences in the phase of phase locking can also encode information, and this has been referred to as *spatiotemporal coding*.

Rate-place coding in normal-hearing animals has been shown to be sufficient for vowel identification in quiet (May et al., 1996), but temporal-place coding is more robust to increased levels and background noise (Delgutte and Kiang, 1984; Geisler and Gamble, 1989; Sachs et al., 1983; Silkes and Geisler, 1991; Young, 2008). Noise-induced hearing loss (NIHL) degrades both rate-place and temporal-place representations of vowels, although temporal-place coding remains in some conditions for which rate-place coding is lost (Geisler, 1989; Miller et al., 1999a; Palmer and Moorjani, 1993). Some research suggests that *spatiotemporal coding* may be important for speech (Deng and Geisler, 1987; Heinz, 2007; Shamma, 1985a) as well as for pitch, intensity,

localization, and masking (Carney et al., 2002; Heinz et al., 2001a; Joris et al., 2006b; Larsen et al., 2008; Shamma and Klein, 2000). Monaural coincidence neurons, similar to the binaural units found in the medial superior olive (Goldberg and Brown, 1969), could theoretically decode temporal-place and spatiotemporal cues by comparing responses of fibers with similar (but not identical) characteristic frequencies. In fact, evidence suggests that neurons in the ventral cochlear nucleus show enhanced temporal coding (Joris et al., 1994a, 1994b; Rothman et al., 1993) and globular bushy cells in particular appear to perform this cross-frequency monaural coincidence detection (Carney, 1990; Wang and Delgutte, 2012).

1.1.3 Hearing Aid Design

1.1.3.1 General Design Principles

Modern digital hearing aids have numerous algorithms available, which can be split into three categories: modeling, cleaning, and managing. Modeling algorithms make up for some of the hearing loss by applying frequency-dependent gain to the signal. Cleaning algorithms try to improve the signal-to-noise ratio (e.g., directional microphones and noise reduction). Managing algorithms reduce any artifacts caused by other algorithms (e.g., feedback reduction when too much gain is applied).

The most fundamental function of a hearing aid is to amplify sound. Whether the goal is simply to restore audibility (Scollie et al., 2005), equalize loudness (Moore, 2000; Moore et al., 1999a), or to improve speech intelligibility (Byrne et al., 2001; Dillon, 2001), the gain of most modern hearing aids is nonlinear. Just as a normal (nonlinear) cochlea

would, a hearing aid can apply substantial gain to low intensity sounds and less gain to sounds that are already of high intensity. As shown in Figure 1.2, the slope of the input/output function of a nonlinear hearing aid is less than unity. The gain in this example is decreased as the input level increases, thus compressing the dynamic range of the sound presented to the ear. Hearing impaired listeners often have steeper than normal loudness growth (known as loudness recruitment; see for example Moore, 2007) but perceive loud sounds normally. Therefore, the compressive gain of a hearing aid is designed to restore nonlinearity by amplifying soft sounds but minimally affecting more intense sounds.

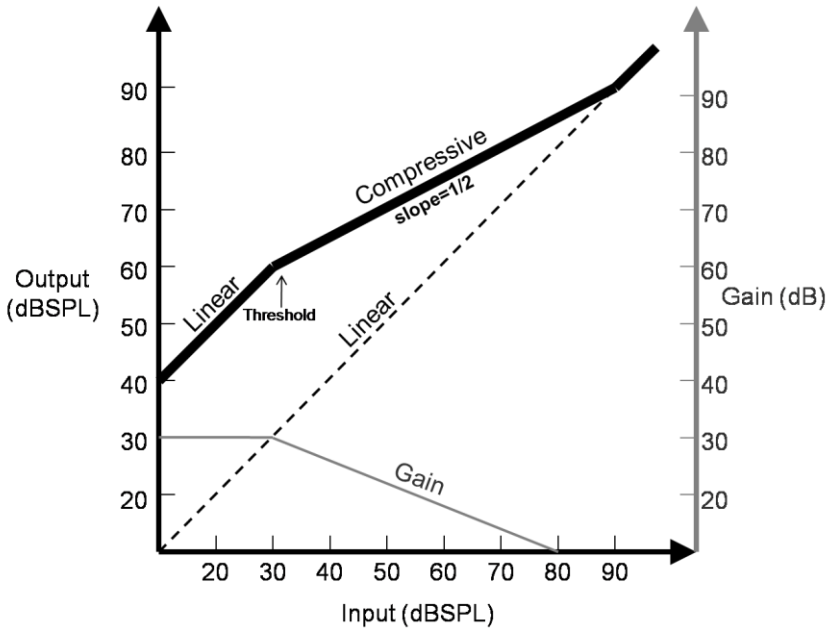

Figure 1.2. Static gain curve showing wide dynamic-range compression
Low level inputs are amplified by a constant gain, but above a given threshold the gain is reduced, resulting in a compressed range of output levels (Gain is shown in gray; input-output function is shown in black)

To automatically control the gain electronically, the incoming sound level is detected and the gain changed accordingly (Kates, 2005). As shown in Figure 1.3, a hearing aid typically splits the signal into at least two frequency bands, detects the incoming level, and applies the appropriate amount of gain.

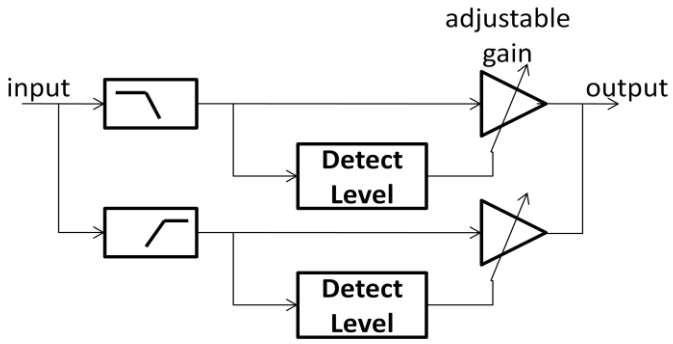

Figure 1.3. Simplified hearing aid block diagram
The input is filtered into 2 or more frequency bands, the input level is determined, and the gain for each frequency band is adjusted

However, this gain adjustment does not occur instantaneously. Because a single sample does not accurately represent the intensity of the signal, the level must be detected over some time interval. Additionally, the gain is often controlled to change somewhat slowly over time and thus to minimize distortion (Souza, 2002). For example, fast amplitude modulation can result in spectral components that may not otherwise exist. Typically, the change in amplitude is described by the exponential function, $y = y_0 e^{-\frac{t}{\tau}}$, where y is the amplitude with initial condition y_0, t is the post-onset time, and τ is a time constant that is chosen by the designer (ANSI, 2003; Moore, 2008a).

Figure 1.4 shows some examples of fast and slow compression. When the signal level rises above the threshold, the gain is reduced as a function of time. For fast time

constants, the gain is reduced over a short period of time. As implemented in many hearing aids, this time constant is often on the order of 100 milliseconds, which corresponds to the approximate length of a syllable. For slow time constants, the gain is

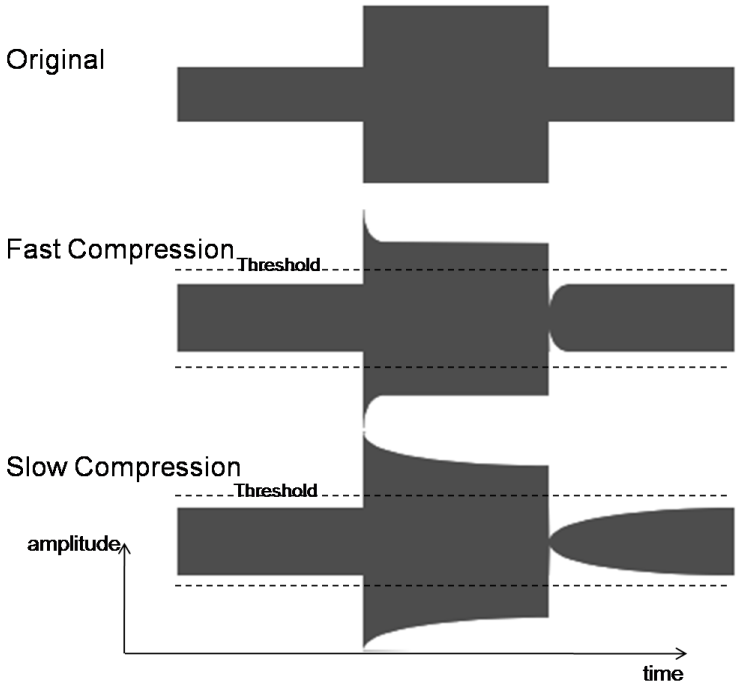

Figure 1.4. Dynamic gain curves
Sounds above the threshold are compressed over time. Fast compression results in faster gain adjustments than does slow compression

changed slowly, taking many seconds in some designs (Moore, 2008b). One advantage of a fast compression system is that short, quiet syllables will be boosted to levels near those of nearby syllables. However, in addition to the desired signal, noise is also boosted, often resulting in an objectionable *pumping* or *breathing* sound (Moore and Glasberg, 1988). Slow compression systems do not suffer from noise pumping, but any quiet sounds that follow a period of intense sounds may not be returned to audible levels (Moore, 2008a, 2008b). It has also been pointed out that relatively slow time

constants may be desired to preserve the slowly varying envelope of the signal (Plomp, 1988). However, an 'optimal' time constant will balance the needs for both audibility and minimal distortion.

In a 'cocktail party' situation, in which many people are speaking at once, hearing impaired listeners often have trouble segregating the voice of the person directly in front of them from all the other voices. Unfortunately, this is a very difficult problem to solve. Hearing aid manufacturers have tried to minimize background sounds with directional microphones (Dillon, 2001), adaptive digital noise reduction (Bentler and Chiou, 2006), and use of binaural hearing aids to improve localization (Bruce, 2006). Many of these techniques are used in modern digital hearing aids, but, unfortunately, only about 50% of patients are satisfied with the performance of their hearing aids in noisy situations (Edwards, 2007). Perhaps one reason for this dissatisfaction is that computational algorithms are not currently able to decide what information is important to preserve and what is background noise. The only technology that has been shown to improve speech intelligibility is microphone directionality, though this often assumes the sound source of interest is directly in front of the patient (Dillon, 2001).

One reason for the difficulty with noise reduction is the fact that the physiological mechanisms (and neural coding) of hearing impairment and subsequent hearing aid amplification are not well understood. Although a vast amount of research has gone into the behavioral results of hearing aid design and fitting strategies (Moore et al., 1999b; Peters et al., 1998; Souza and Tremblay, 2006; Souza, 2002), little is known about the underlying neurophysiology of hearing aid use. Knowledge of such

neurophysiology may be beneficial and may offer new insights into the design of auditory prostheses.

1.1.3.2 Biologically-Inspired Designs

Biondi (1978) suggested that, by comparing neural coding in normal ears and impaired ears (with amplification), a hearing aid might be designed that minimizes the difference. Several scientists have used computational models of auditory physiology to implement such a system (Bondy et al., 2004; Chen et al., 2005; Kates, 1993; Shi et al., 2006).

One set of algorithms designed to restore cochlear patterns may collectively be called *spectral contrast enhancement* algorithms (Baer et al., 1993; Kates, 1994; Simpson et al., 1990; Yang et al., 2003). The motivating theory is that broadened auditory tuning degrades the signal-to-noise ratio within any particular spectral channel (Henry and Heinz, 2012). These algorithms attempt to increase the contrast between spectral peaks and valleys. Miller and colleagues (1999b) showed that contrast enhancement can improve the neural representation of vowel formants.

Bondy and colleagues (Bondy et al., 2004; Haykin et al., 2006) developed a system they called a 'neurocompensator', which attempted to restore the instantaneous firing rate of auditory nerve (AN) fibers. As the authors noted, however, the accuracy of the model is unknown for important auditory features like transients or phenomena like forward masking. It is also important to note that any imperfect restoration of the neural code could potentially result in audible artifacts. A system like this would benefit

from knowledge about what aspects of the neural code are most important to restore. The authors have since published work, developing a prediction of speech intelligibility based on neural information (Bondy et al., 2003; Zilany and Bruce, 2007a) that might be useful in combination with the neurocompensator algorithm.

Carney and colleagues developed an algorithm that was designed to introduce delay into the auditory signal where the phase response was predicted to be abnormal (Calandruccio et al., 2007; Carney, 2008; Shi et al., 2006). The motivating theory is that broadened auditory filters have a shallower phase response, and thus less group delay, than normal. The algorithm uses two parallel paths to estimate delay and to add frequency-dependent delay to the auditory signal. In the control path, an auditory model is used to estimate the group delay introduced by healthy nonlinear filters. In the main path, the signal is decomposed into frequency channels, a delay is added, and the channels are re-combined after passing through a synthesis filterbank. The merits and limitations of this approach are discussed further in Chapter 6.

1.2 Research Approach

To better understand how hearing aids affect the ability to listen in complex situations, we can learn from neurophysiology. The auditory nerve offers an excellent source of information about the peripheral auditory system. All information from the ear travels to the brain through the auditory nerve, and any peripheral hearing impairment will result in changes to the signals within the auditory nerve.

By modeling and measuring physiological responses to complex auditory stimuli, we can gain insights into the effects of impairment and subsequent amplification. The

research presented here uses both a computational model and an animal model to investigate how speech is encoded in the mammalian auditory nerve. Computer models allow us to examine how neurons are likely to respond, but they are limited in their ability to replicate all the details of a biological system. An animal model can give us a more accurate picture of the biological system, but requires much more time and effort (and often animals' lives) to collect that data.

1.2.1 Computational Model

Several models of the mammalian peripheral auditory system have been developed over recent years (reviewed by Heinz, 2010; Lopez-Poveda, 2005). We used a phenomenological model of cochlear physiology (Zilany and Bruce, 2006, 2007b; Zilany et al., 2009) that is an extension of several previous models (Bruce et al., 2003; Carney, 1994; Heinz et al., 2001b; Zhang et al., 2001). This particular model was chosen because

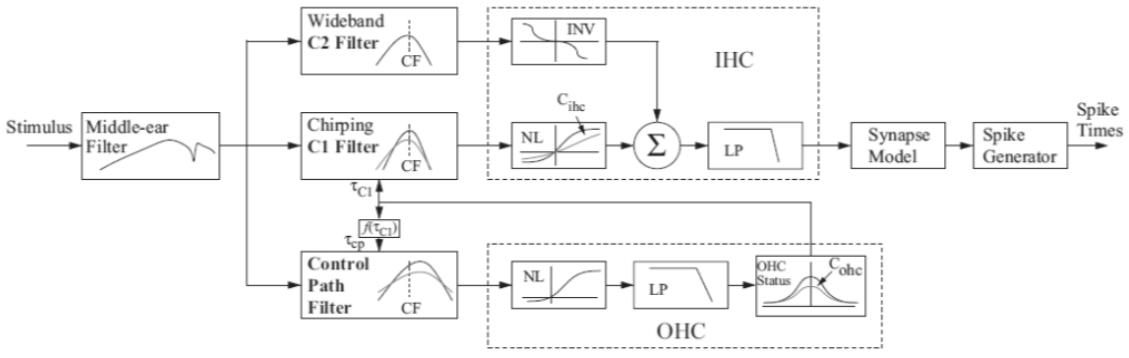

Figure 1.5. Auditory nerve model
The input is any acoustic waveform, and the output is a set of times at which auditory nerve spikes are predicted to occur. Reprinted with permission from (Zilany and Bruce, 2006). Copyright 2006, Acoustical Society of America.

it has been compared with auditory nerve data obtained with cats (Miller et al., 1997; Wong et al., 1998) and was found to match the physiological data for vowel responses quite well across a wide range of sound levels (Zilany and Bruce, 2007b). The model also provides control of both inner and outer hair cell functionality.

A schematic diagram of the auditory nerve model is shown in Figure 1.5. It is important to note that this model is not designed to accurately represent the biophysics of cochlear mechanisms, but is a phenomenological model that produces outputs similar to what can be measured in animals. The input to the model is an arbitrary acoustic waveform and the output is a series of times indicating when auditory nerve spikes are predicted to occur. Hearing loss is controlled by adjusting the values of C_{OHC} and C_{IHC}, which control the amount of dysfunction associated with the outer and inner hair cells, respectively. Total threshold shift for a single fiber is modeled as a combination of contributions from both types of hair cells. The desired audiometric hearing loss can thus be set for each characteristic frequency such that the total hearing loss (HL), in dB, at a specific frequency is represented by the equation, $HL_{total} = HL_{OHC} + HL_{IHC}$. We generally set the model such that two-thirds of the threshold shift is due to outer hair cell dysfunction and one-third due to inner hair cell dysfunction, which is consistent with average results for both human perception (Lopez-Poveda and Johannesen, 2012; Moore and Glasberg, 1997; Plack et al., 2004) and animal physiology (Bruce et al., 2003; Harding and Bohne, 2007, 2009).

Note that the synapse model for the Zilany and Bruce auditory nerve model was updated in 2009. Each section of this dissertation that utilizes the model also specifies

which version was used. Generally, work started after the 2009 update used the newer version. When using the updated synapse model (with longer adaptation time constants), we increased the inter-stimulus interval from 50 ms (of silence) to 1 sec.

Although this model was fit to cat data, the primary difference between different mammals is the size. The position along the length of the cochlea can be compared across species using the function $x = \frac{1}{a}\log_{10}\left(\frac{F}{A} + k\right)$, where *x* is the distance from the base of the cochlea, F is the frequency, and the constants *a*, *A*, and *k* depend on the species (Greenwood, 1990). Although the frequency tuning in humans may be sharper than in many laboratory mammals (Shera et al., 2002), this remains a topic of some controversy (Joris et al., 2011; Lopez-Poveda and Eustaquio-Martin, 2013; Ruggero and Temchin, 2005). Additionally, the consistent trend in all species tested (including humans) is that tuning gets sharper at high frequencies. Any within-species comparisons of normal versus impaired hearing would then be expected to show similar trends in other species.

1.2.2 Animal Model

The present work also involves acute surgical and experimental procedures to record responses directly from individual auditory nerve fibers in chinchillas. Chinchillas were chosen for several reasons. A large body of anatomical, physiological, and behavioral data exists on the auditory system for chinchillas (Morest et al., 1990; Ruggero et al., 1997; Ruggero and Rich, 1987; Shofner, 1999). Chinchillas have low-frequency hearing, similar to humans, and are thus a good model for studies such as

these that focus on clinically relevant issues related to the neural correlates of human auditory perception (especially percepts that are thought to require phase locking). The long experiments (18-36 hours) that can be performed with chinchillas also produce high yields of data, thereby reducing the total number of animals required.

All procedures were approved by the Purdue Animal Care and Use Committee. Male animals were usually obtained at around 6 months of age, weighing approximately 400-500g. Some animals were used as models of normal hearing, while others were exposed to noise to induce sensorineural hearing loss.

The acoustic trauma procedure for inducing sensorineural hearing loss was similar to the one used previously with cats (Heinz and Young, 2004; Heinz et al., 2005) and chinchillas (e.g., Kale and Heinz, 2010). Noise over-exposure typically results in mixed inner and outer hair cell dysfunction, which is likely to be common in many hearing impaired patients (Liberman and Dodds, 1984a). The animal was anesthetized using a combination of xylazine (1-1.5mg/kg im) and ketamine (50-65mg/kg im) and its head was restrained. Atropine (0.1mg/kg im) was given to control mucus secretions and eye ointment was used to prevent drying of the eyes. Prior to exposure, auditory brainstem response (ABR) thresholds and distortion product otoacoustic emissions (DPOAE) were measured to establish a baseline. ABR thresholds were measured with tone bursts at 0.5, 1, 2, 4, and 8 kHz using insert earphones.

Noise was presented from a loudspeaker approximately 30cm above the animal's head in a sound-attenuating chamber. The noise used for over-exposure was one octave wide, centered at 500 Hz, and was presented at 116dB SPL for 2 continuous

hours. After the exposure, animals were kept warm and monitored until recovery from anesthesia was complete. The animal was then allowed to recover for at least 4 weeks to allow temporary threshold shifts to dissipate (Nordmann et al., 2000). Prior to acute experiments on noise-exposed animals, hearing loss was confirmed (by verifying that ABR thresholds had shifted by at least 20dB at 2 kHz; Ngan and May, 2001) while under anesthesia.

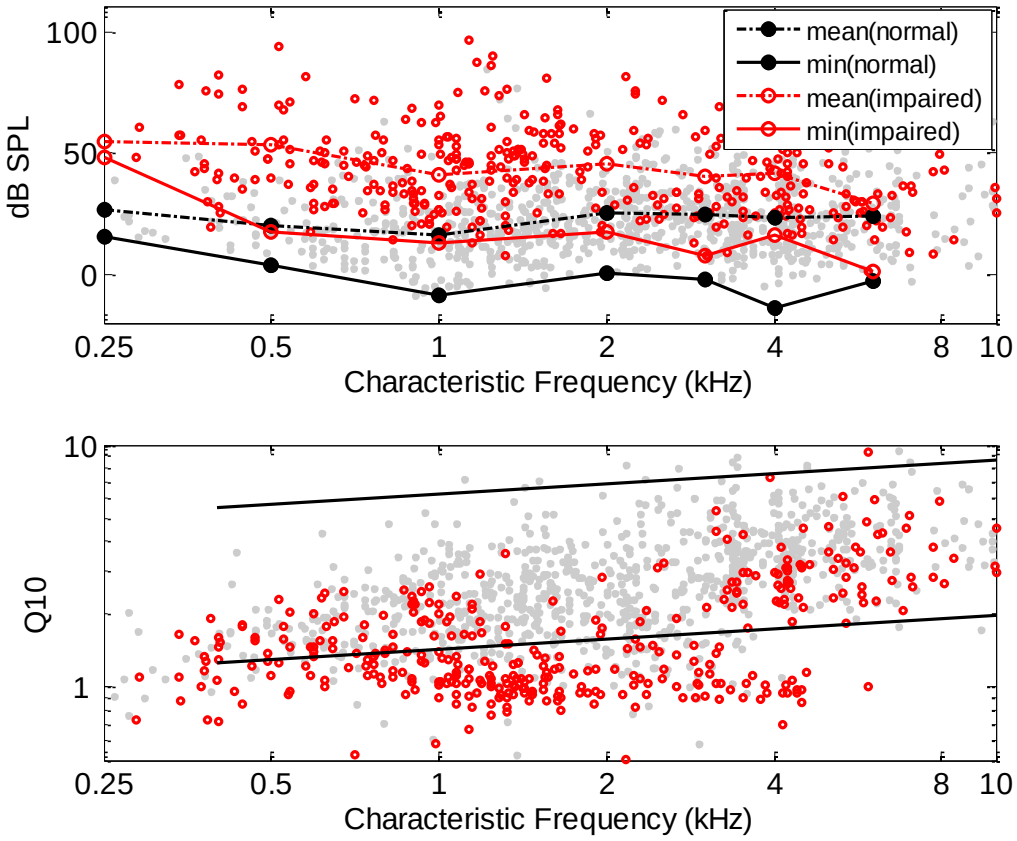

Figure 1.6. Auditory nerve thresholds and tuning sharpness
Fibers from normal-hearing animals are indicated by a gray dot; those from noise-exposed animals are indicated by an open red circle. Lines in the upper panel show population thresholds. Lines in the lower panel indicate 5th and 95th percentiles from Kale and Heinz (2010).

If more than 25% of the tuning curves for any non-exposed animal were broader than the 95th percentile of the normal chinchilla data from (Kale and Heinz, 2010), the data from that animal were discarded. As shown in Figure 1.6, this exposure results in a mild flat hearing loss (approximately 15-20dB HL) over a broad frequency range, consistent with previous studies (e.g., Harding and Bohne, 2009). Population thresholds in the auditory nerve (solid lines in upper panel) were calculated by evaluating the minimum threshold in one-half octave bands. Average thresholds (dashed lines) are also shown for comparison. (Data shown here are from all animals in our lab that were over-exposed to this noise. Vowel data presented in the following chapters were recorded from a subset of these units.)

Standard neurophysiological procedures were used to record from the auditory nerve (Heinz and Young, 2004; Liberman and Dodds, 1984b). Chinchillas were initially anesthetized with xylazine (1-1.5mg/kg im) followed by ketamine (50-65mg/kg im). Atropine (0.1mg/kg im) was given every 24 hours to control mucus, and eye ointment applied to prevent drying of the eyes. A catheter was placed in the cephalic vein to administer intravenous fluids. Barbiturate anesthesia (sodium pentobarbital, 7.5mg/kg/hour iv) was used to maintain an areflexic state throughout the duration of the experiment, typically every 90-120 minutes. (For a few animals, the intravenous catheter could not be properly inserted, so supplemental doses of sodium pentobarbital were administered into the intra-peritoneal cavity.) Saline and lactated Ringer's solution were administered at a rate of 2.5mL per hour to prevent dehydration. A tracheotomy was performed to create a low-resistance airway. Rectal temperature was

maintained near 37°C with a heating pad. With the head in a stereotaxic headholder, a craniotomy was performed to create an opening in the posterior fossa and the cerebellum was minimally aspirated to expose the auditory nerve.

The bulla was vented to equalize the middle ear pressure, and recordings were made in a sound-attenuating chamber. A glass micropipette (10-30 MΩ, filled with 3M NaCl) was inserted into the auditory nerve under visual control. Computer controlled stimuli were presented via a calibrated closed-field acoustic system using a hollow ear bar. (Calibrations were performed for each animal with a probe microphone placed within 3mm of the tympanic membrane.)

Single AN fibers were isolated by searching with a broadband noise. Each fiber was characterized by using an automated tuning curve algorithm (Chintanpalli and Heinz, 2007; Liberman, 1978), and the fiber CF, threshold and Q10 (ratio of CF to bandwidth 10dB above threshold) are estimated. As suggested by Liberman (1984), CFs for impaired fibers were chosen by hand near the high-frequency slope to estimate the original CF prior to impairment. This is based on Liberman's labeling study, which showed that the high-frequency slope of the tuning curve can be used as an indicator of where that fiber innervates along the length of the cochlea. Fibers were also classified as high spontaneous rate (≥ 18 spikes/sec), medium spontaneous rate (0.5 < SR ≤ 18 spikes/sec), and low spontaneous rate (≤ 0.5 spikes/sec) as suggested by Liberman (1978). Spontaneous rate was estimated from a 20sec period of silence, then a peri-stimulus time histogram was measured to verify AN (rather than cochlear nucleus) responses based on the histogram shape, latency, and a monopolar spike waveform.

1.2.3 Data Analysis Techniques

1.2.3.1 Classical Neural Metrics

The simplest way to characterize a single-fiber neural response is to quantify the firing rate, or the number of spikes in response to a sound. However, because rate-based measures often fail to account for basic perceptual phenomena, such as speech perception (Sachs and Young, 1979) or pitch coding (Cedolin and Delgutte, 2005), temporal-based measures have been explored in depth as well (e.g., Larsen et al., 2008; Young and Sachs, 1979). Characterization of the temporal properties of single-fiber neural responses has historically been based on simple periodic stimuli such as pure tones or other periodic stimuli. Although simple metrics such as vector strength or synchronization index (Goldberg and Brown, 1969; Johnson, 1980) provide useful information, they do not apply to complex stimuli such as running speech. A metric based on autocorrelation of actual nerve spikes, such as those based on the interspike interval histogram (Cariani and Delgutte, 1996), is likely more physiologically realistic and generalizable in that it can be used across a variety of auditory stimuli.

1.2.3.2 Shuffled Correlation Metrics

Joris and colleagues (2003; 2006a) recently extended the neural analysis work of Perkel (1967) by using a "shuffled" autocorrelation (SAC) function[1] to characterize temporal coding of single-fiber responses in the auditory system. The SAC is calculated by building a histogram of interspike interval durations across (but not within) several repetitions of a stimulus, such that the SAC is the set of time intervals from each spike to

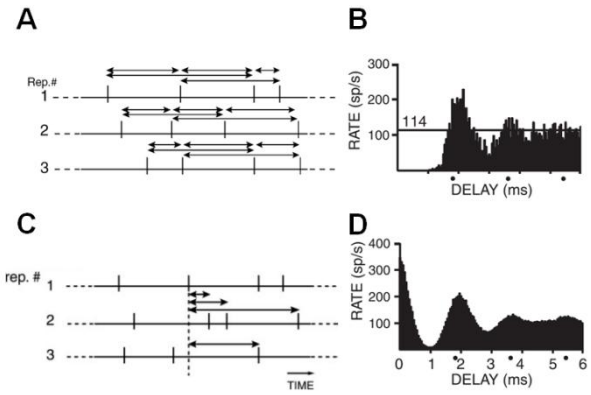

Figure 1.7 All-Order Interval Histogram (A,B)
& Shuffled Autocorrelation Function (C,D)
[graphs A,C reprinted with permission from Joris et al (2006a); graphs B&D reprinted with permission from Louage et al (2004)]

all subsequent spikes in all the other repetitions. Unlike the often used all-order interspike interval histogram which simply measures timing between spikes within each

[1] Joris and others have referred to the correlation function as a "correlogram", but this is a misnomer. In this dissertation, any reference to a correlogram will refer to a collection of correlation functions comprising a three dimensional plot of correlation as a function of both lag and frequency. (see, for example, Figure 4.5)

repetition (as shown in Figure 1.7, A-B), the SAC is not limited by the refractory period of the neural responses within a single repetition (Figure 1.7, C-D). This can be seen by comparing panels B&D at delays less than 1 ms. Also note that, because the histogram is based on roughly N^2 comparisons for N spikes, the SAC is a smoother function than the all-order interval histogram.

Shuffled autocorrelation functions can also be used to study how the envelope (slowly varying time structure) and the temporal fine structure (faster oscillations) of the signal are coded. The envelope (ENV) and temporal fine structure (TFS) are thought to contribute to perception differently. The details of this are discussed in more detail in Chapter 2.

1.2.3.3 Spectro-Temporal Manipulation Procedure

To study the relative timing across multiple fibers, we would like to study the responses of several closely-spaced fibers in each animal. However, this is very difficult in practice because the spacing of CFs found during an experiment can be quite sparse, and even estimates of CF are somewhat variable (Chintanpalli and Heinz, 2007).

The spectro-temporal manipulation procedure (STMP; Heinz, 2007; Larsen et al., 2008) was used in the present work to study predicted spatiotemporal patterns based on the responses of a single fiber to several stimuli. In a manner similar to the *spectrum manipulation procedure* for predicting rate responses (LePrell et al., 1996; May et al., 1996), the sampling rate of the stimulus was modified to shift the spectral content and predict the response of a nearby CF to the same stimulus. The STMP, however, uses a

subsequent step of scaling the recorded spike times to accurately predict both rate and *temporal responses* of nearby CFs (i.e., to correct for the temporal scaling that results from changing the sample rate).

As an illustrative example, using the STMP, we can predict the response of two fibers, A and B, to a single stimulus using the response of a single fiber (at CF_0) to two stimuli, A and B (as illustrated in Figure 1.8). For example, to predict the response of the fiber corresponding to CF_A in Figure 1.8 (blue curve, upper panel), we would play the vowel at CF_0/CF_A times the original sample rate, thus shifting the spectrum up (as shown in the lower panel, blue curve). The recorded spike times would then also be scaled up by a

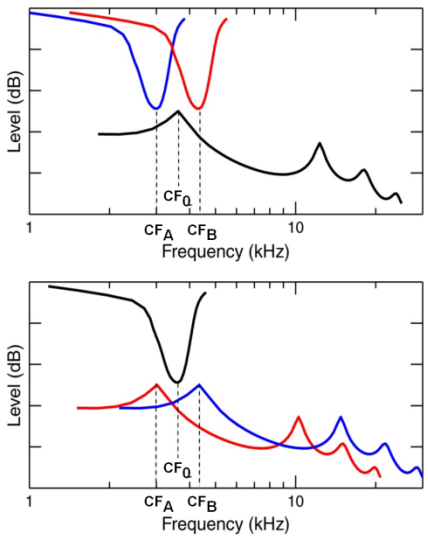

Figure 1.8 Spectro-Temporal Manipulation Procedure (STMP)
The STMP can be used to predict responses of multiple fibers to a single stimulus (A) from a single fiber responding to multiple frequency-shifted stimuli (B). Figure reproduced with permission from (Heinz, 2005).

factor of CF_0/CF_A to accurately predict the temporal response to the original stimulus (illustrated here as a vowel with a formant centered at CF_0). Similarly, we could predict

the response of a fiber at CF_B by shifting the playing the stimulus at a low rate (thus shifting the spectrum down), then scaling the spike times down.

A computational model has been used to test this technique by comparing the spatiotemporal patterns obtained with the STMP (i.e., temporal pattern across predicted CFs) and those obtained through the modeled CFs directly (i.e., temporal pattern across actual CFs) (Larsen et al., 2008). The benefit of this technique is that it allows analysis of neurophysiological data of closely and accurately spaced CFs, which is quite difficult to do with conventional population studies.

1.3 Overview of This Dissertation

Chapter 2 presents a modeling study that evaluates the optimal hearing aid gain for different configurations of hearing loss. This work demonstrates a quantitative approach to factoring in the physiological effects of SNHL to hearing aid fittings. It builds upon the work of Bruce and colleagues (Bruce et al., 2007), but quantifies the strength of envelope and temporal fine structure coding with a set of neural metrics that have been used in other studies (Heinz and Swaminathan, 2009). This study also explores the idea that the optimal gain may in fact differ from individual to individual, depending on the proportion of outer and inner hair cell dysfunction. However, while Chapter 2 focuses on neural coding within individual auditory nerve fibers, the rest of this dissertation focuses on across-fiber coding.

Chapter 3 presents a small initial study in which we evaluated spatiotemporal coding in the auditory nerve. We presented broadband noise and a speech sentence to one

chinchilla with normal hearing and another with noise-induced hearing loss. Using the STMP approach combined with shuffled correlation metrics, we calculated estimates of the cross-fiber correlation and delay, and demonstrated increased cross-CF correlation and decreased delays in noise-exposed fibers compared with normal-hearing fibers. This work provides a foundation to build upon in the next chapter, which narrows the focus to the study of vowel coding.

Chapter 4 extends the work presented in Chapter 3 by evaluating spatiotemporal patterns for vowels in noise. This allows us to investigate questions about spatiotemporal coding, while also keeping us grounded by comparing our results with previously published data on vowel coding. Consistent with our earlier results, the data indicate that impairment reduces cross-fiber delays. The data also indicate that the spatiotemporal code is robust in the presence of noise, consistent with previous research. These data can be used as a baseline for evaluating the ability of hearing aids to restore neural coding, which is the topic of the next chapter.

Chapter 5 presents our investigation into the effects of hearing aid amplification on spatiotemporal coding of vowels in noise. We evaluated the neural responses when the stimuli were amplified with a linear gain prescription (NAL-R) and a wide dynamic-range prescription (DSL[i/o]). As we hypothesized, neither hearing aid prescription improved the spatiotemporal coding of the signal. Although this was not necessarily surprising, before evaluating proposed hearing aid algorithms to improve spatiotemporal coding (in Chapter 6), it was critical for us to document the effect of existing hearing aids on spatiotemporal coding.

Chapter 6 includes our investigation of the general approach used in the spatiotemporal pattern correction scheme proposed by Carney and colleagues (Carney, 2008; Shi et al., 2006). Rather than directly evaluate that one algorithm, we investigated the more general issue of the underlying assumption that across-fiber delays can be controlled by introducing frequency-dependent delays into the stimulus. Although we can indeed detect these time delays in modeled single-fiber auditory nerve responses, we found that neither our correlation metrics nor a simple model of a coincidence detector neuron were affected by these delays. This general result suggests that correcting the spatiotemporal code may not be as theoretically simple as adding acoustic frequency-dependent delays, and future hearing aid technologies are likely to require more complexity and/or ingenuity than the simple approaches proposed to date.

The final chapter discusses some of the limitations of the work presented here, as well as what the present work suggests about the potential roles of modeling, physiology, and psychophysics for the future of hearing aid design. Some potential opportunities for future research are also presented.

CHAPTER 2. PHYSIOLOGY-BASED HEARING AID FITTING

This work was presented as the following conference poster:

Boley, J. *and M. Heinz, Quantifying the Effects of Hearing Aid Dynamics on Temporal Coding in the Auditory Nerve, First International Symposium on Audible Acoustics in Medicine and Physiology, September 2008.*

A hearing aid often attempts to restore the impaired ear's missing gain, compressing the dynamic range to make soft sounds audible while keeping loud sounds comfortable. Although hearing aids have been tremendously successful in many situations, patients still have an abnormal degree of difficulty in acoustically complex environments (Gatehouse et al., 2003).

People with normal hearing have a remarkable ability, commonly known as the 'cocktail party effect' (Cherry, 1953), to understand a single person in a room full of other people speaking simultaneously. Hearing impaired listeners often complain of an inability to perform such tasks, even when all the sounds are individually audible. Duquesnoy (1983) pointed out that people may in fact "listen in the dips" of the background noise to extract information about important sounds from a complex mixture. Normal hearing listeners seem to be able to use the small amount of auditory information in short, relatively quiet intervals, but hearing impaired listeners have trouble hearing in these situations. Moore (2003) suggested that the temporal fine

structure (TFS) of the acoustic waveform is important for understanding speech in complex acoustic environments and Lorenzi and colleagues (2006) showed that the ability to listen in the presence of modulated noise is correlated with the suprathreshold ability to utilize TFS.

2.1 Background

Bondy and colleagues (Bondy et al., 2004; Haykin et al., 2006) attempted to minimize the difference between normal and impaired coding by optimizing parameters of an amplification algorithm. However, the authors considered only the rate-place encoding of the auditory signals and did not calculate any measure of phase locking, where phase locking refers to the fact that auditory neurons tend to fire in sync with a particular phase of the stimulus waveform. More recently, Bruce and colleagues (Bruce et al., 2007) claim to have calculated neural information based on both average discharge rate and spike timing. Their results suggested that more than the prescribed gain was generally needed to optimize the slowly varying rate, whereas less gain (especially at high levels) was needed to optimize timing information. However, the only difference between these two measures was the length of the averaging window; the authors used a very short window size to evaluate temporal coding. Bruce averaged spike counts using a Hamming window length of 256μs, which has the effect of attenuating fluctuations faster than approximately 2.5 kHz. This metric might therefore measure timing (e.g. phase locking) in response to low frequencies, but it may not be sufficient because synchronous timing can be measured up to at least 5 kHz in

the auditory nerve of some mammals (Johnson, 1980; Weiss and Rose, 1988). Because precise timing may be important for hearing in complex situations (Gilbert and Lorenzi, 2006; Hopkins and Moore, 2007; Hopkins et al., 2008; Lorenzi et al., 2006), future physiologically-based designs should consider metrics that include both long-term rate and precise temporal coding.

2.2 Experimental Methods

To evaluate the effect of a hearing aid on neural coding, we compared the predicted neural signals from three systems, as illustrated in Figure 2.1. Using an auditory nerve model (Zilany and Bruce, 2006, 2007b), we predicted spiking patterns from a normal-hearing auditory system and an impaired auditory system. For the impaired system, we also calculated the output of the model preceded by a hearing aid amplification algorithm.

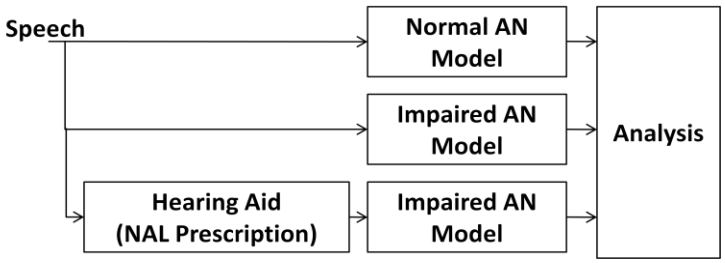

Figure 2.1 Schematic diagram of experimental conditions

2.2.1 Computational Model of Hearing Impairment

A population of 30 auditory nerve fibers was modeled to represent center frequencies ranging from 250 Hz to 8 kHz. For impaired hearing simulations, the

coefficients C_{OHC} and C_{IHC} were chosen to result in a mild hearing loss, as shown in Figure 2.2. The coefficients were adjusted to achieve the desired hearing loss using three scenarios: only IHC dysfunction, 2/3 of threshold shift (in dB) due to OHC dysfunction (1/3 due to IHC dysfunction), and nearly all OHC dysfunction. A hearing aid gain profile (Figure 2.3) was fit to this audiogram, based on the NAL-R prescription (Byrne et al., 1990). A speech stimulus was then run through the AN model for three separate scenarios: a normal-hearing case, an impaired case, and an impaired case with a hearing aid. The speech stimulus was a single word in quiet, and the level was adjusted from 60dB SPL to 100dB SPL in 10dB steps. The resulting neural spike patterns were then analyzed for comparisons across these three cases to quantify the ability of the hearing aid to restore normal temporal coding. The shape of the frequency-gain curve was set to the NAL-R prescription, then the overall level was adjusted (-40 to +40dB) to determine if the optimal overall gain differed from the prescribed overall gain.

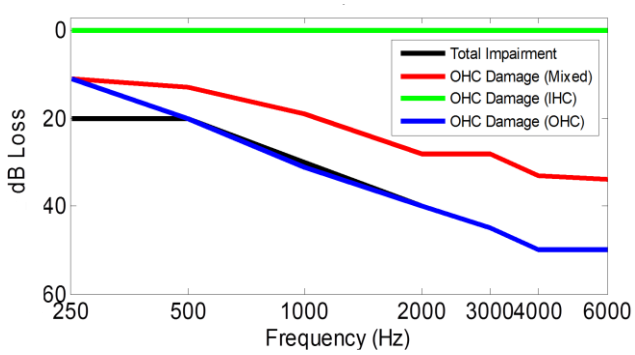

Figure 2.2 Audiogram showing mixtures of OHC and IHC dysfunction
Total modeled threshold shift (black) was accomplished via three configurations: nearly all OHC dyfunction (blue), all IHC dysfunction (green) and a mixture (red). Blue, green and red lines indicate the threshold shift due to outer hair cells.

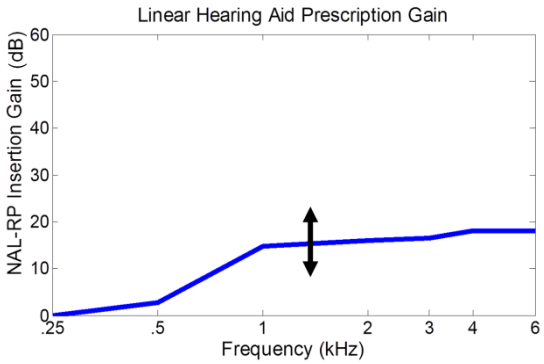

Figure 2.3 NAL-R prescriptive gain for the audiogram shown in Figure 2.2. The overall level of this curve was adjusted to determine optimal gains.

2.2.2 Measuring Envelope and Fine Structure

Joris and colleagues (2006a) applied the shuffled autocorrelation function (SAC, see section 1.2.3.2) to a strategic set of signals in order to determine what part of the temporal code was responding to the envelope and which part was due to the temporal fine structure. They presented a stimulus, A+, recorded the resulting neural pattern, and calculated the SAC (Figure 2.4A). They then presented an inverted polarity version of the same stimulus, A-, again recorded the spikes, and calculated the SAC (Figure 2.4B) which matches the first SAC (except for some scaling due to adaptation). By analyzing the spikes from A+ in reference to A-, a cross-stimulus autocorrelation (XAC) function was then calculated (Figure 2.4C). Heinz and Swaminathan (2009) have referred to this correlation between A- and A+ as a shuffled cross-polarity correlation, or SCC(A+,A-) because the two signals are not identical and the function is therefore not an autocorrelation. The SCC(A+,A-) was calculated in a manner similar to the SAC (as described in section 1.2.3.2) but the intervals are based on the times between each

spike in A+ and subsequent spikes in A-. The envelope of the signal is the same for both A+ and A-, and anything that is common between the SAC and SCC(A+,A-) is taken as a measure of the envelope. The average of the SAC and SCC(A+,A-) (labeled SUMCOR in Figure 2.4D) therefore estimates the autocorrelation function based on the neural envelope (ENV) response. The difference between the SAC and SCC(A+,A-) (DIFCOR; Figure 2.4E) estimates the autocorrelation function based on the neural temporal fine structure (TFS) response. The peak heights of the SUMCOR and DIFCOR functions can then be used to represent the amount of envelope and fine-structure temporal encoding, respectively.

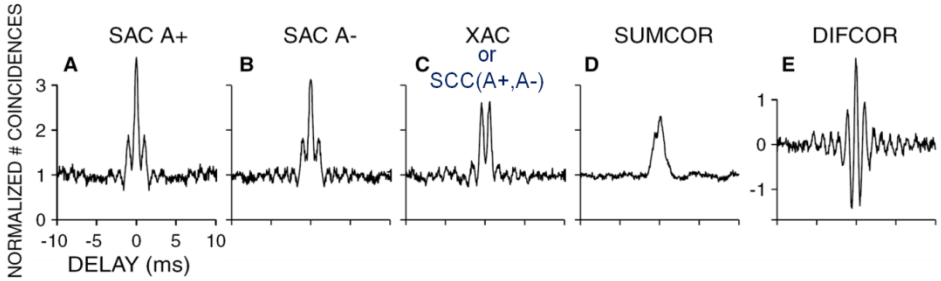

Figure 2.4. Shuffled correlation functions
A&B) SACs of polarity-inverted versions of the stimulus; C) shuffled cross-polarity correlation (SCC); D) SUMCOR is the average of SAC and SCC, and represents enveleope coding; E) DIFCOR is the difference between SAC and SCC, and represents temporal fine structure coding [reprinted with permission from Joris et al (2006a), copyright 2006, Acoustical Society of America]

Unfortunately, because these metrics measure responses at the level of the auditory nerve, they measure the ear's response to a stimulus, which may have a different ENV & TFS than the original acoustic signal. As Ghitza (2001) pointed out, the narrow-band

filtering of the cochlea can change the relative amount of ENV and TFS available to the auditory nerve. The narrow filters of the cochlea are in fact able to generate a "recovered envelope" from the broadband fine-structure (as depicted in Figure 2.5).

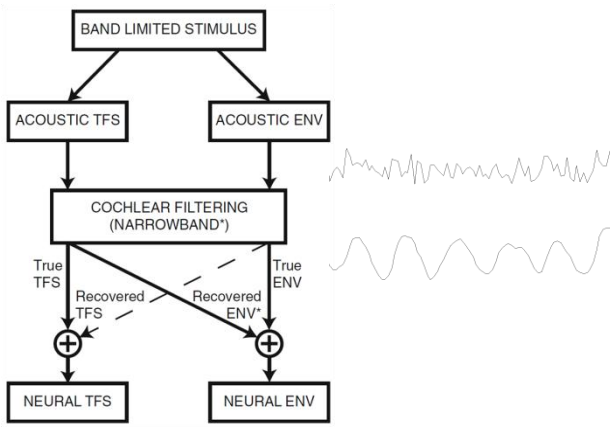

Figure 2.5. Theoretical framework illustrating recovered envelopes (Waveforms illustrate a signal before and after narrowband cochlear filtering). Reprinted, with permission, from Heinz and Swaminathan (2009).

Heinz and Swaminathan (2009) extended the SAC and SCC metrics by calculating neural ENV and TFS correlation coefficients based on two separate measurements of SUMCOR or DIFCOR (ρ_{env} and ρ_{tfs}, respectively). These metrics, as calculated below, are used to evaluate the similarity in ENV or TFS coding between two different sets of neural spike trains. ρ_{tfs} and ρ_{env} are defined as:

$$\rho_{tfs} = \frac{difcor_{AB}}{\sqrt{difcor_A \times difcor_B}} \qquad \text{Equation 1}$$

$$\rho_{env} = \frac{(sumcor_{AB} - 1)}{\sqrt{(sumcor_A - 1) \times (sumcor_B - 1)}} \quad \text{Equation 2}$$

These metrics are similar to the well-known Pearson correlation coefficient, which is simply the covariance divided by the standard deviations of the two random variables. Similarly, these neural cross-correlation metrics range from 0 to 1, where 0 indicates no correlation and a value of 1 indicates excellent correlation. These neural cross-correlation metrics have general applicability, because conditions A and B can be responses to two different stimuli measured from the same neuron (e.g., to quantify recovered envelopes, as in Heinz and Swaminathan, 2009), or the same stimulus applied to two different neurons. For example, these metrics could be applied to one normal-hearing and one aided+impaired at the same characteristic frequency, as in the present chapter, or two different CFs, and in Chapters 3-6.

Figure 2.6 illustrates how the correlation metrics can be applied to compare two conditions. Panel A shows the shuffled autocorrelation function for condition A [SAC(A+); thick line] and the shuffled cross-polarity correlation function [SCC(A+/A-); thin line], whereas panel B shows the same for condition B. Panel C shows the shuffled cross-condition correlation function [SCC(A+,B+); thick line] and the shuffled cross-polarity/cross-condition correlation function [SCC(A+,B-); thin line]. DifCors for A-C are shown in panels D-F, and SumCors for A-C are shown in panels G-I. The correlation coefficients, ρ_{env} and ρ_{tfs}, are calculated based on these correlation functions according to equations 1-2.

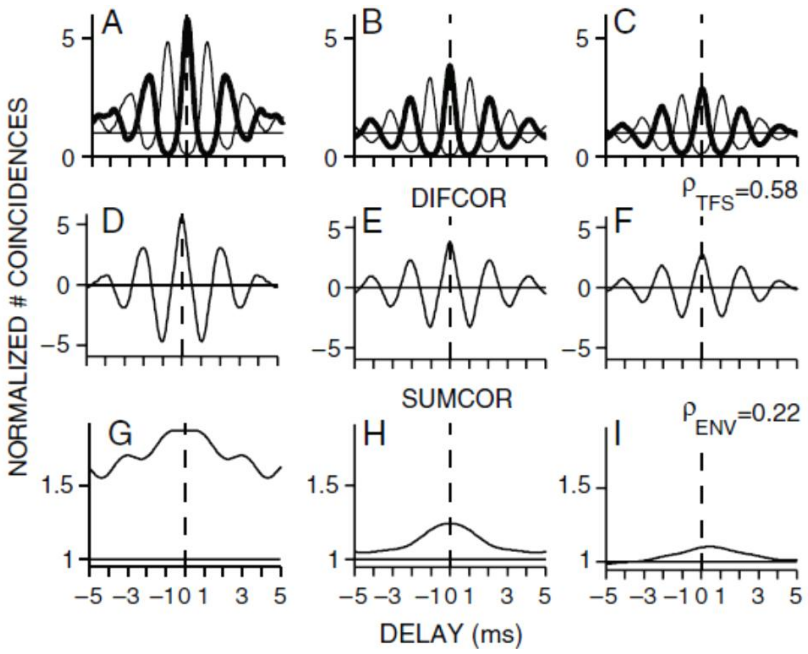

Figure 2.6 Example of comparison between normal and impaired envelope coding The first column represents correlation analyses of one condition; the second column represents correlation analyses of another condition. The third column represents cross-condition correlation analyses. Reprinted, with permission, from Heinz and Swaminathan (2009).

The neural firing rate, ρ_{tfs}, and ρ_{env} were compared across the normal and aided+impaired cases (where the hearing aid applied a simple linear gain to a word in quiet). Each of these metrics was averaged across the 30 model AN fibers. The optimal gain for each input level was the one which most closely restored the coding to normal, as indicated by the smallest difference in rate or by a cross-correlation coefficient (either ρ_{tfs}, or ρ_{env}) closest to a value of 1. (Although rate was a function of time, the gain which optimized rate was defined here as the gain which minimized the average difference between normal and aided+impaired conditions.)

2.3 Results

We calculated the average (long-term) and short-term firing rates as in the study by Bruce and colleagues (2007) and found very similar results - gains above NAL-R were typically required to optimize average discharge rates (Figure 2.7A) while lower, compressive gains (less gain for higher input levels) were required to optimize short-term discharge rates (Figure 2.7B). When we optimized the gain for envelope coding, however, the results differed from those for average discharge rates (i.e., with an 8 ms time window, as used by Bruce and colleagues). The general trend for envelope optimization appeared to be a gain above NAL-R at moderate levels that decreased for higher input levels (Figure 2.7C). Optimization of temporal fine structure required less gain than any other metric used, as shown in Figure 2.7D. So, overall we see a consistent trend similar to the results of Bruce and colleagues (2007), where slow temporal information requires higher gain and faster temporal information lower gain. However, there are some specific differences that highlight the importance of explicitly evaluating ENV and TFS information.

This indicates that a prescription lower than NAL-R may be preferred for encoding fine structure information. However, more gain is needed for encoding envelope information than is provided by the prescription. In fact, one gain setting could work well for speech in quiet, where envelope information is important, and a lower gain setting might work better in noisy conditions where temporal fine structure is thought to be important. Compression, or gain that decreases as level increases, may be preferred for encoding both envelope and fine structure information. This is consistent

with what we know about the physiology of a normal-functioning auditory system in which the cochlear amplifier provides less gain for high-intensity sounds. It also appears that the NAL-R prescription balances the needs for both envelope and fine-structure coding when there is a mixture of both inner and outer hair cell dysfunction.

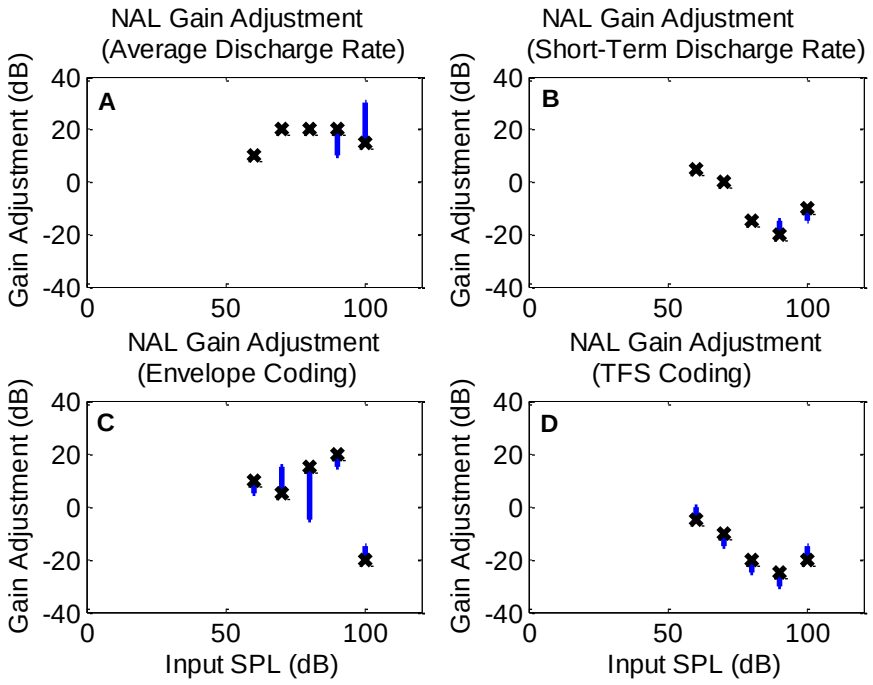

Figure 2.7 Optimal gains for mixed hair cell dysfunction.
A-B: optimization of average and short-term discharge rate, as in Bruce et al (2007). C-D: optimization of envelope and fine structure coding (Results were similar for OHC dysfunction.) Lines indicate the top 10% of optimal gains.

We used a similar approach to evaluate the effects of selective inner hair cell (IHC) dysfunction. As discussed in section 1.1.2, IHC dysfunction is known to result in elevated thresholds without a loss of tuning, whereas OHC dysfunction results in both elevated thresholds and degraded tuning (Liberman and Dodds, 1984a).

Using the frequency dependent gain of the NAL-R prescription as a baseline for a mild hearing loss, the optimal gains were very similar to the case of mixed hair cell dysfunction when optimizing for average discharge rate, short-term rate, or fine structure coding (Figure 2.8A,B,D). However, when optimizing to achieve near-normal envelope coding (Figure 2.8C), we found that the optimal gains differed substantially from those for mixed OHC and IHC damage.

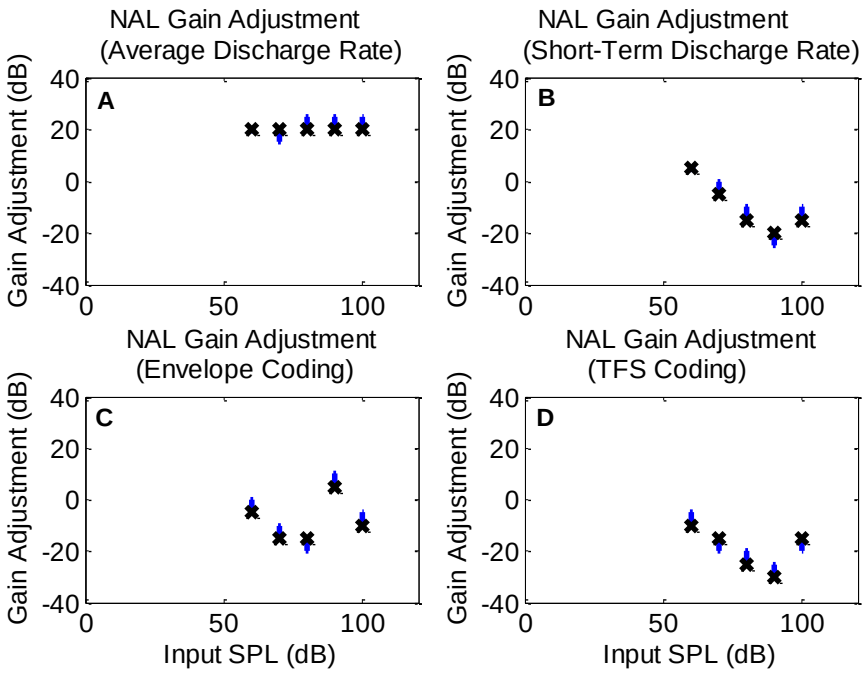

Figure 2.8 Optimal gains for *inner* hair cell dysfunction.
A-B: optimization of average and short-term discharge rate, as in Bruce et al (2007). C-D: optimization of envelope and fine structure coding. Lines indicate the top 10% of optimal gains.

Bruce (2010) recently showed that when the gain was optimized for restoring the average discharge rate, the result was a large spread of synchrony to vowel formants.

He showed that this spread of synchrony at optimal gain only occurred for simulations of IHC dysfunction, and not for simulations of OHC dysfunction. This suggests that it is important to consider responses to the envelope and fine structure separately, especially when considering IHC dysfunction.

To better understand why the optimal gains for envelope coding were so different, we looked at the neural metrics in detail for a 70dB SPL stimulus (Figure 2.9). The top row shows SumCors for both IHC and mixed dysfunction when +20dB of gain (re NAL-R) is applied; the bottom row shows the same plots when -20dB of gain is applied. The first column (A,D) shows the SumCor for within-fiber envelope coding with normal hearing. The second column (B,E) shows the SumCors for within-fiber envelope coding for both the IHC dysfunction and the mixed dysfunction models. The third column (C,F) shows the across-condition envelope coding, comparing envelope coding in normal to either IHC dysfunction or mixed dysfunction.

The SumCors for the impaired systems with gain 20dB below NAL are shown in Figure 2.9E. Notice that the peak of the IHC curve has lowered substantially, as compared to Figure 2.9B, to be much closer to normal envelope coding (see Figure 2.9A,D). Also notice that the SumCor of the cross-correlation (Figure 2.9C,F) is reduced as the gain is reduced for the mixed dysfunction case (lower SumCor peak, ρ_{env}, indicating envelope coding further from normal), but the SumCor has increased as the

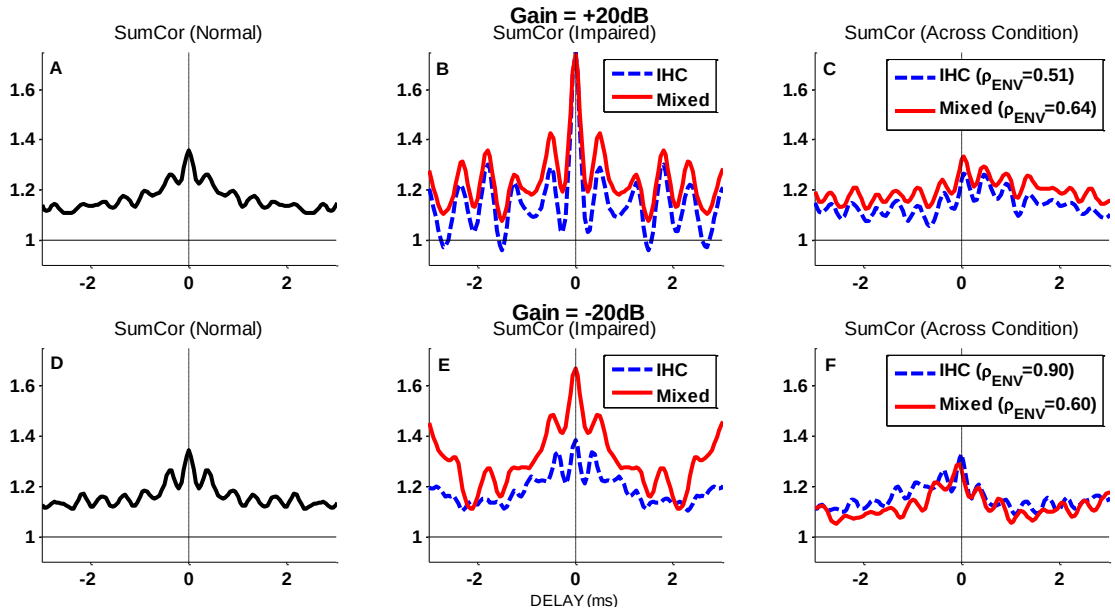

Figure 2.9 SumCors for normal (A,D) and impaired (B,E) conditions 20dB gain (A-C) or -20dB gain (D-F) relative to NAL-R. C&F show the SumCor for the cross-correlation of normal and each impaired condition. Whereas the first two columns (A,D,B,E) show within-fiber envelope coding, the third column (C&F) shows the similarity between normal envelope coding and aided-impaired envelope coding. Note that NAL-R+20dB was optimal for mixed hair cell dysfunction, whereas NAL-R-20dB was optimal for inner hair cell dysfunction.

gain is reduced for the case with only IHC dysfunction (higher SumCor peak, ρ_{env}, indicating envelope coding closer to normal). Therefore, we see here that the optimal gain setting is not necessarily the one that enhances envelope coding, but the one which best restores envelope coding to normal.

At the optimal gain for each stimulus level, the average correlation coefficients (across CF) were higher for envelope than TFS, as shown in Figure 2.10. This suggests that amplification is able to restore envelope coding more than TFS coding. Similarly, the average correlation at the optimal gain was higher (more similar to normal) when the hearing loss was due to inner, rather than mixed, hair cell dysfunction. This

indicates that hearing impairment due to inner hair cell dysfunction can be corrected more than impairment that involves outer hair cells.

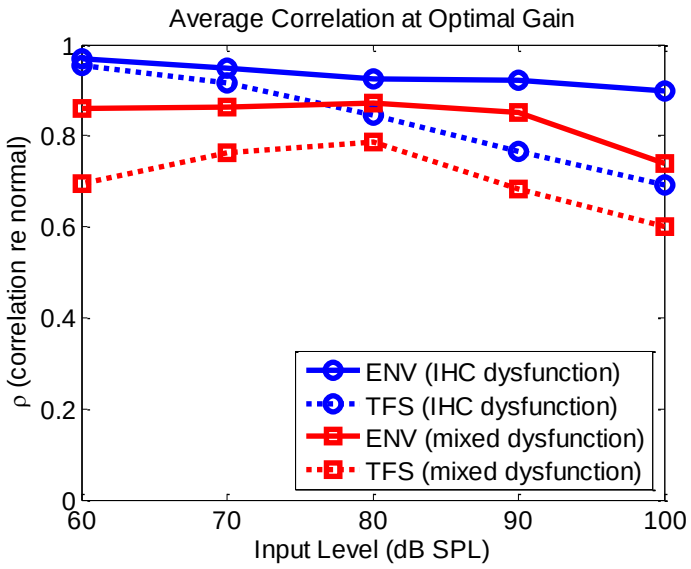

Figure 2.10 Average correlation at optimal gain

2.4 Discussion

These results suggest that hearing aid users may benefit from prescriptive gain settings that take into consideration their underlying physiology. If scientists are able to predict the degree of outer and inner hair cell impairment (Lopez-Poveda and Johannesen, 2012; Moore et al., 1999c), it may be beneficial to adjust the gain accordingly. In fact, this may reduce the inter-subject variability in performance that often occurs even among patients with very similar audiograms.

It is interesting to note that compression is needed to preserve timing information, in terms of both short-term rate and temporal fine structure (as calculated using the

correlation metrics). Little or no compression seems to be necessary for preserving rate information <100 Hz (given an 8 ms Hamming window), but it appears that some compression may in fact be necessary for envelope coding (which was only limited by the bandwidth of each nerve fiber).

Of particular interest is the result suggesting that the optimal gain for preserving envelope in a patient with primarily OHC dysfunction is drastically higher than the gain for a patient with primarily IHC dysfunction. If this is in fact true, then it would be beneficial to clinically assess a patient's OHC/IHC dysfunction before fitting a hearing aid. Given some information about the underlying physiology, a hearing aid could be better fit for the individual patient. For example, a computational model could be used to match the patient's behavioral performance by adjusting the relative OHC vs IHC dysfunction, then the hearing aid parameters could be adjusted to improve performance of the model. These optimized parameters could then be tested on the patient, thus minimizing the patient's time in the clinic but potentially maximizing performance.

The work presented in this chapter assumes that improving the neural coding within each auditory nerve fiber will translate directly to a perceptual improvement. If we could succeed at making all of the auditory nerve responses within an impaired system look exactly like the responses of a normal system, and if we can safely assume that all the impairment is peripheral, perhaps perception would return to normal. However, it is theoretically impossible to fully restore the response of every neural fiber to normal if there is any OHC dysfunction, even if we were reasonably sure of the model accuracy

(Giguère and Smoorenburg, 1999; Heinz, 2010). This limitation is illustrated in Figure 2.10, where the largest correlation coefficients (i.e., the degree to which coding can be restored to normal) were lower for mixed OHC/IHC loss than for IHC dysfunction alone.

Although restoring the temporal coding within each fiber may be beneficial, research suggests that within-fiber coding does not account for performance on some psychoacoustic tasks (e.g., Cedolin and Delgutte, 2005). Rather than focusing on the restoration of neural codes within each nerve fiber, it may be beneficial to use knowledge of how the brain uses this information and focus on restoring population codes as well. For example, evidence has been mounting that relative temporal coding (e.g., across multiple auditory nerve fibers) may be important perceptually (e.g., Carney, 1994; Cedolin and Delgutte, 2007; Shamma, 1985a), and decoding this information may be one of the first things the brain does when it receives information from the ear (Carney and Friedman, 1998; Carney, 1990; Wang and Delgutte, 2012). The remainder of this dissertation focuses on this relative temporal coding across CFs, commonly known as spatiotemporal coding.

CHAPTER 3. SPATIOTEMPORAL CODING IN THE AUDITORY NERVE

The work presented in this chapter was also published as the following book chapter (reprinted at the end of this document):

*Heinz, M., Swaminathan, J., **Boley, J.,** & Kale, S. (2010). Across-Fiber Coding of Temporal Fine-Structure: Effects of Noise-Induced Hearing Loss on Auditory-Nerve Responses. In E. A. Lopez-Poveda, R. Meddis, & A. R. Palmer (Eds.), The Neurophysiological Bases of Auditory Perception (pp. 621–630). New York: Springer.*

3.1 Background

Listening in a "cocktail-party situation" (i.e., with multiple competing sounds) is a complex task (Bregman, 1990), and peripheral hearing impairment hinders our ability to organize these auditory scenes (for a review, see Shinn-Cunningham and Best, 2008). The perceptual cues used to segregate sounds in a complex mixture have been well studied, but the neural codes are not as well understood. We know that temporal codes are more robust to noise than rate-based codes (Delgutte and Kiang, 1984; Sachs et al., 1983; Young and Sachs, 1979) and several researchers have proposed mechanisms for decoding temporal codes as a function of cochlear place (Carney, 1990, 1992; Carney et al., 2002; Deng and Geisler, 1987; Shamma, 1985b; Wang and Delgutte, 2012). In fact, these temporal-place (or 'spatiotemporal') cues are thought to be

important for several psychoacoustic phenomena, including speech perception, pitch, and intensity coding as well as tone-in-noise masking and interaural timing differences (Carney et al., 2002; Cedolin and Delgutte, 2007; Heinz, 2007; Joris et al., 2006b; Larsen et al., 2008; Shamma, 1985a; Shamma and Klein, 2000).

Shamma (1985a, 1985b) proposed that rapid phase shifts across fibers with different CFs may encode important information about auditory stimuli (for example, vowel formants). Evidence suggests that neurons in the brainstem (specifically, globular bushy cells in the cochlear nucleus) are sensitive to the phase slope of certain stimuli (Carney, 1990; Wang and Delgutte, 2012).

If the phase of the neural signals are at least partially determined by the phase of the auditory filters, we expect that impairment (with broad auditory filters and shallow phase responses) will decrease the slope of the neural phase shifts across CF. The goal of this study was to quantify the relative timing of auditory nerve responses across nearby CFs, and determine if spatiotemporal coding changes with hearing impairment.

3.2 Methods

In this study, we measured auditory nerve responses to broadband noise and sentence-level speech. The recordings were performed in two anesthetized chinchillas, using standard procedures (Heinz and Young, 2004; Kale and Heinz, 2010). Hearing impairment was induced in one animal by presenting a 50 Hz-wide noise band centered at 2 kHz for 4 hours at 115 dB SPL, after which the animal was allowed to recover for 6 weeks. Consistent with previous studies (Heinz and Young, 2004; Liberman, 1984), this resulted in increased thresholds, by approximately 30-50dB, and broadened tuning for

all fibers. For each fiber, the characteristic frequency (CF) was chosen by hand near the high-frequency slope of the tuning curve to approximate the CF prior to impairment (Liberman, 1984).

For each AN fiber, the STMP (presented in section 1.2.3.3) was used to predict the responses of nearby fibers with CFs within ± 0.5 octaves of the actual CF. Spikes within the first 1 ms of each response were assumed to be spontaneous activity, and these spike times were not scaled.

The methodological techniques used for across-CF correlations are described in section 1.2.3.2 and in a number of related publications (Heinz and Swaminathan, 2009; Joris, 2003; Joris et al., 2006a, 2006b; Louage et al., 2004). Briefly, a shuffled cross-correlation function (SCC) between the responses of two fibers with different CFs is calculated. The peak height of the SCC, relative to the geometric mean of the SAC peak heights for each CF (see Equations 1-2), determines the correlation coefficient (ρ), a metric that represents the similarity (normalized from 0 to 1) between temporal responses of the two fibers. Based on the difference between the SAC and SCC (derived from the responses to positive and negative polarity stimuli, see Section 2.2.2), we can calculate a correlation coefficient for the temporal fine structure (ρ_{tfs}). Because of the traveling wave, the peak of the cross-correlation function (SCC) between the two CFs will occur at a non-zero delay, which we call the characteristic delay (CD). The CD represents the traveling wave propagation time between these two CFs, which is expected to increase as ΔCF increases. Figure 3.1 illustrates how these cross-correlation metrics can be calculated from the within- and across-fiber correlation functions.

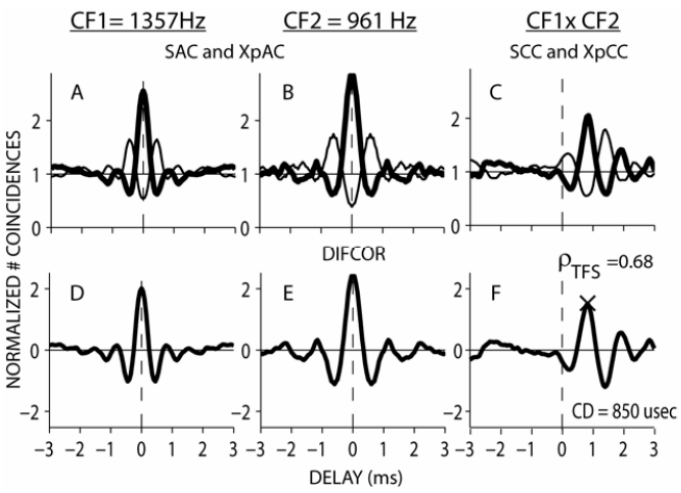

Figure 3.1 Temporal coding based on shuffled correlation functions. A&B: Shuffled AutoCorrelation (SAC; thick line) function and Cross-Polarity AutoCorrelation (X_PAC; thin line) function for two CFs separated by 0.5 octaves. C: Shuffled Cross-CF-Correlation (SCC; thick line) function and Cross-Polarity Cross-Correlation (X_PCC; thin line) function. D-F: Difcors calculated as the difference between auto- and cross-correlation functions shown in A-C, respectively. F: The relative peak height of the cross-fiber difcor indicates the correlation coefficient (ρ_{tfs}). The SCCS peak is shifted by the characteristic delay (CD). [Figure reproduced, with permission, from Heinz et al (2010)]

3.3 Results

Figure 3.2 shows the cross-CF analysis for a normal and an impaired auditory nerve fiber with similar CF, responding to a broadband noise stimulus. This is representative of the data collected from 17 normal and 19 impaired fibers. Panel A shows the tuning curves for a normal-hearing animal (dashed line) and a hearing-impaired animal (solid line). Based on similar high-frequency edges of the tuning curves (Liberman, 1984), the CF of the impaired fiber is approximated to be the same as the CF of the normal fiber, or 1.3 kHz. The correlation (ρ_{TFS}) and characteristic delay (Figure 3.2 B&C, respectively) are plotted as a function of ΔCF for every combination of effective CFs tested (based on the

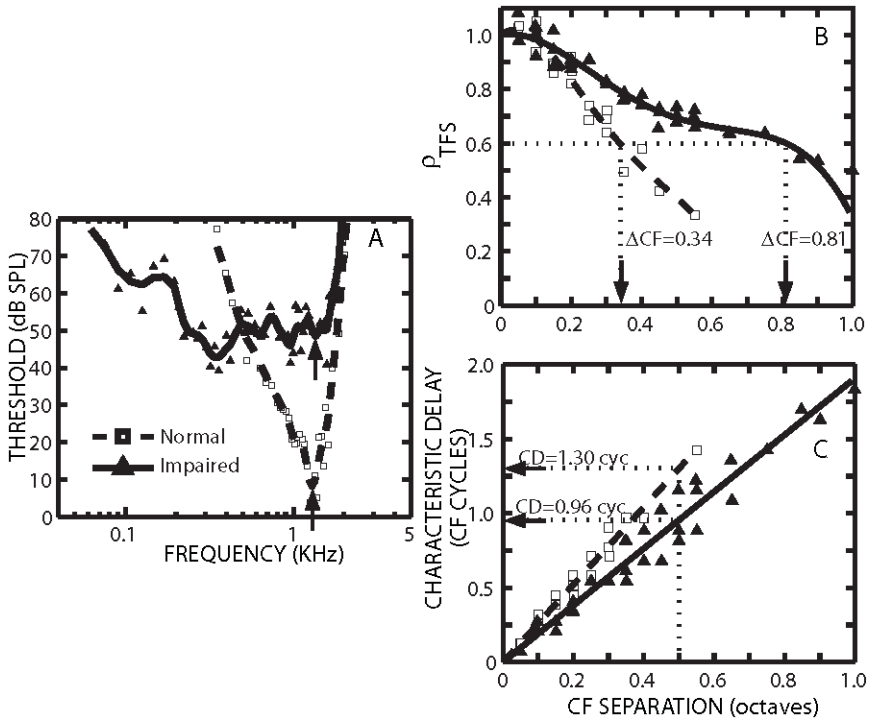

Figure 3.2 Normal-vs-impaired spatiotemporal coding based on individual fibers The STMP was used to predict a population of effective CFs, and spatiotemporal coding was quantified using shuffled across-CF correlation analyses. Open squares represent data from a normal-hearing animal; closed triangles represent data from a hearing-impaired animal. A) Tuning Curves for a normal and an impaired auditory nerve fiber; B) Predicted across-CF correlation strength as a function of effective CF separation; C) Predicted across-CF characteristic delay as a function of effective CF separation. [Figure reproduced, with permission, from Heinz et al (2010)]

STMP). The variation in ρ_{TFS} was fit with a fourth order polynomial constrained such that the value at a CF separation of 0 octaves was equal to 1.0. The variation in characteristic delay was fit with a line constrained such that the value at a CF separation of 0 octaves was equal to 0. The *spread of correlated activity* ($\rho_{0.6}$) is quantified by measuring the CF separation at which the correlation falls to a value of 0.6, and the characteristic delay ($CD_{0.5}$) was quantified by the time delay between the CFs separated by 0.5 octaves. For the normal fiber shown in Figure 3.2, $\rho_{0.6}$ was 0.34 octaves, whereas

for the impaired fiber, $\rho_{0.6}$ was 0.81 octaves (as shown in panel B). For the normal fiber, $CD_{0.5}$ was 0.96 CF cycles, whereas for the impaired fiber, $CD_{0.5}$ was 0.81 CF cycles (as shown in panel C). These results suggest that noise-induced hearing loss tends to reduce traveling wave delay and increase the spread of correlated activity across the cochlea.

Population responses to both broadband noise and a speech sentence are shown in Figure 3.3. The left column shows the strength of within fiber TFS coding (panel A), and across-CF coding (panels C,E) in response to broadband noise. The right column shows similar data for the speech stimulus. The CF region near 1 kHz contains data for both normal and impaired AN fibers. Although within-fiber coding of fine structure does not appear to be degraded by impairment (Figure 3.3A,B), cross-CF correlations were affected by impairment (Figure 3.3C,D). After impairment, the width of correlated activity (across a range of basilar membrane locations) was increased. The increased width of correlated activity appears to be greater for speech than for noise.

The characteristic delay between effective CFs that were half of an octave apart decreased by approximately 0.25 cycles after impairment (Figure 3.3E,F). The size of this phase shift is consistent with level-dependent phase shifts seen in guinea pigs (Palmer and Shackleton, 2009). For pure tones, the relative phase for a half-octave region can vary by approximately 0.25 to 0.5 cycles over a 40-50 dB range of levels. This effect is thought to be related to nonlinear cochlear tuning that is seen with normal outer hair cell function. If impairment alters the phase by 0.25 cycles (e.g., from in-

phase to uncorrelated), this could have a significant impact on any neural mechanism that depends on this delay, such as cross-fiber coincidence detection.

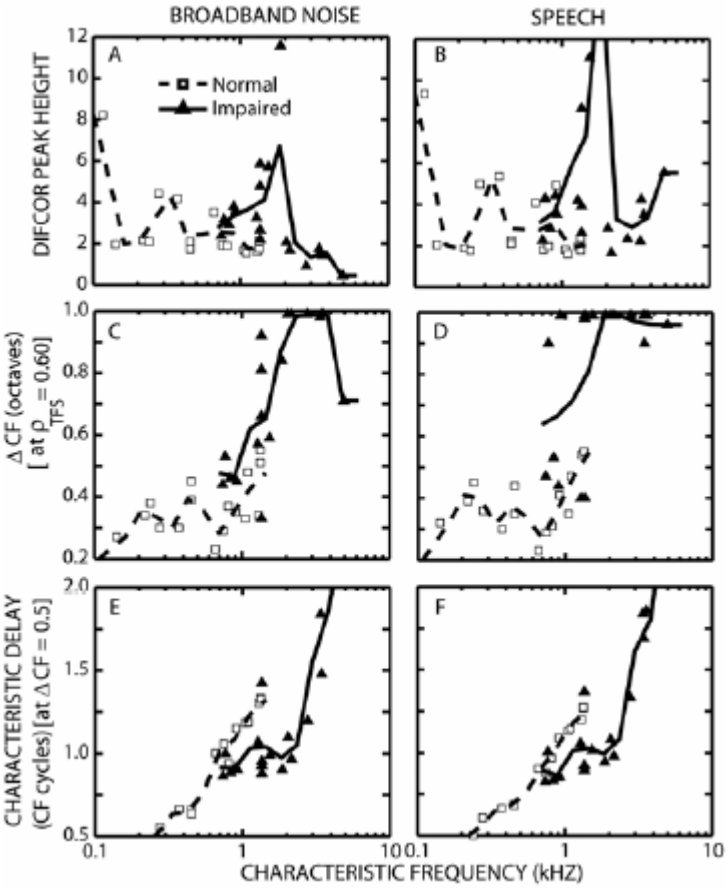

Figure 3.3 Comparison of normal and impaired spatiotemporal coding
Spatiotemporal coding of temporal fine structure based on the population of auditory nerve fibers responding to broadband noise (left column) and a speech sentence (right column). A-B: Within-fiber TFS coding strength is represented by difcor peak heights. C-D: The smallest CF separation at which ρ_{tfs} dropped to a value of 0.6 represents the width of correlated activity. E-F: The characteristic delay at 0.5 octaves of CF separation estimates the phase delay (in CF cycles) between the two locations on the basilar membrane. The lines represent moving averages using a 0.7 octave wide triangular window. Figure reproduced, with permission, from Heinz et al (2010).

3.4 Discussion

We have shown that across-CF coding of temporal fine structure is altered by sensorineural hearing loss. We found that impairment resulted in broader regions of correlated activity, which might be expected due to broadened tuning. This broadened tuning may be perceptually relevant for listening in complex conditions because the number of independent neural channels of information would be reduced, potentially making some listening tasks more difficult.

We also found a reduction in estimated traveling wave delay following SNHL. This increase in propagation speed would increase the coincidence of temporal information across a population of fibers with different CFs, thus altering the normal spatiotemporal patterns that have been hypothesized to include robust neural cues for pitch, speech, and intensity coding (Carney, 1994; Cedolin and Delgutte, 2007; Shamma, 1985a).

By better understanding the effects of SNHL on spatiotemporal coding, we may be able to indentify some ways to improve the design of auditory prostheses like hearing aids and cochlear implants. The next chapter investigates how SNHL affects spatiotemporal coding of vowels, both in quiet and in the presence of background noise.

CHAPTER 4. SPATIOTEMPORAL CODING OF VOWELS IN NOISE

Portions of this work were presented in the following conference posters:

Boley, J. *and M. Heinz, Predicted Effects of Amplification on Spatiotemporal Coding of Vowels in Noise, International Hearing Aid Research Conference, August 2010.*

Boley, J. *and M. Heinz, Impaired Spatiotemporal Coding of Vowels in Noise, International Hearing Aid Research Conference, August 2012.*

In the previous chapter (and Heinz et al., 2010; see appendix), we showed that across-fiber coding of temporal information is altered by sensorineural hearing loss. We found that impairment resulted in broader regions of correlated activity, which is expected due to broadened tuning. This broadened tuning may be perceptually relevant for listening in complex conditions because the number of independent neural channels of information would be reduced, potentially making some listening tasks more difficult. We also found a reduction in the estimated traveling wave delay between different places along the length of the cochlea. This increase in propagation speed should increase the coincidence of temporal information across a population of fibers, thus altering the normal spatiotemporal patterns.

4.1 Background

Miller and colleagues (1997) showed that NIHL degrades phase locking to vowel features - following NIHL, fibers with characteristic frequencies (CFs) near the formants

tend to phase-lock to many individual harmonics rather than the formant. They also showed that responses of fibers with CFs in the spectral trough of a vowel (between F1 and F2) are not suppressed as they are in normal-hearing animals, thus degrading the contrast between the formant peak and the trough.

A subsequent study (Schilling et al., 1998) showed that impairment caused an upward spread of F1 synchrony and that precisely aligned frequency-shaped amplification can limit this spread, thus improving the representation of higher formants. However, the researchers found strong phase-locking to harmonics in the trough, thus degrading the spectral contrast. They also demonstrated that the frequency-shaped amplification did not prevent upward spread of higher formant synchrony (i.e., to F2 and F3).

The work presented here extends previous studies by more thoroughly quantifying the effects of NIHL on vowel coding in noise. Specifically, this expands upon the work of Heinz (2007), which evaluated spatiotemporal coding of the first formant and trough of the vowel /ɛ/ in noise. This work adds to this previous research by quantifying impaired rate-place, temporal-place, and spatiotemporal coding of the first two formants in noise. The work presented here also expands upon the cross-CF coincidence model (Deng and Geisler, 1987; Heinz, 2007) by calculating the cross-correlation of adjacent CFs using novel neural metrics recently developed in our lab, as discussed in Heinz and Swaminathan (2009). These correlation metrics are used to calculate the characteristic delay (an estimate of the traveling wave delay) between two CFs. This will enable us to better understand how impairment affects cross-CF coding of vowels. We expected that

impaired spatiotemporal coding would be characterized by a decrease in traveling wave delay and a spread of correlation across CF (Heinz et al., 2010).

4.2 Methods

The noise exposure and surgical procedures are discussed in section 1.2.2. Nine chinchillas with normal hearing and eight noise-exposed chinchillas were presented with the vowel /ɛ/ in quiet and in noise. AN tuning thresholds and bandwidths for these animals are shown in Figure 4.1. Responses to the vowel were measured in conditions for which rate-place coding was expected to be poor – at moderately high levels and in the presence of background noise. Vowels were synthesized with a cascade formant synthesizer (Klatt, 1980). A fundamental frequency of 100 Hz was used, with formant frequencies of 500, 1700, 2500, and 3300 Hz (as in Miller et al., 1997; Schilling et al., 1998; Young, 2008). The presentation level was adjusted according to the rate-level function measured when the second formant was centered on the fiber CF. The measured rate-level function was fit with a model (Sachs et al., 1989) and the presentation level was chosen as the level which produced a rate two-thirds of the way from the spontaneous rate to the saturation rate.

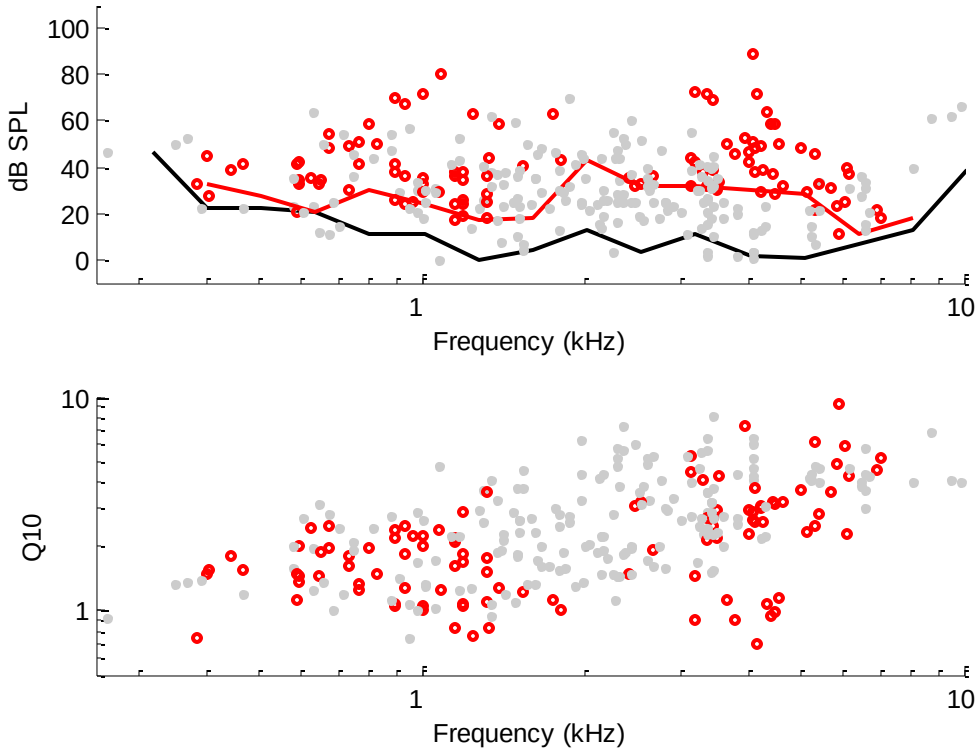

Figure 4.1 Auditory nerve tuning thresholds and sharpness for this experiment

Background noise was a frozen waveform of speech-shaped noise (Byrne et al., 1994). When noise was included in the stimulus, it was presented at two levels: one at the same sound pressure level as the vowel (0dB signal-to-noise ratio, or SNR) and one which elicits the same firing rate as the vowel (approximating equal sensation level). For each condition, both vowel and noise stimuli were 2sec in duration, with 10 ms rise/fall ramps. (After adjusting the sampling rate using the STMP (described in section 1.2.3.3), the stimuli durations ranged from 1.4sec to 2.8sec.) Stimuli were repeated once every 3sec until 2000 spikes were recorded, based on our observation that this number of spikes ensures consistent quantitative metrics (Heinz and Swaminathan, 2009).

4.2.1 Quantifying the Strength of Temporal Coding

The methodological techniques used for across-CF correlations are described in sections 1.2.3.2, 1.2.3.3, and 3.2. Briefly, the STMP was used to predict responses of nearby fibers with effective CFs placed at -0.75, -0.50, -0.25, -0.15, -0.05, 0, 0.05, 0.15, 0.25, 0.35, and 0.50 octaves from the actual CF. In contrast to the method used for data analysis in the previous chapter (which assumed 1 ms of spontaneous activity at the beginning of each spike train), we used a frequency-dependent time function fit to the response latency of a 0.1 ms condensation click at 50dB SPL (as in Wang and Delgutte, 2012). The following quadratic function was fit to the latency predicted by a computational model of normal hearing (Zilany and Bruce, 2006; Zilany et al., 2009):

$$\lambda = 0.005228*x^2 - 0.01203*x + 0.008404 \qquad \text{Equation 3}$$

where λ is the approximated neural conduction delay (in seconds) and x is the proportion of the cochlear length. The proportional position for a given frequency can be calculated according to the equation given by Greenwood (1990):

$$F = A(10^{ax} - k) \qquad \text{Equation 4}$$

where F is the frequency (in Hz) corresponding to that position, x is the proportional length of the cochlea, and constants A=163.5, k=0.85, and a=2.1 were used for chinchillas.

The frequency-dependent delay represents the travelling wave delay, and is approximately 1.5 ms at the base of the cochlea and 8.5 ms at the apex. By not scaling any spikes before this conduction delay, we expect to obtain a better approximation of the temporal firing patterns. If we were to use the constant 1 ms delay used for the

SMTP in Chapter 3, we would see a similar pattern of characteristic delay as a function of CF (shown in Figure 4.2). However, the results presented in Chapters 4 and 5 use the frequency-dependent neural conduction delay, so patterns look different.

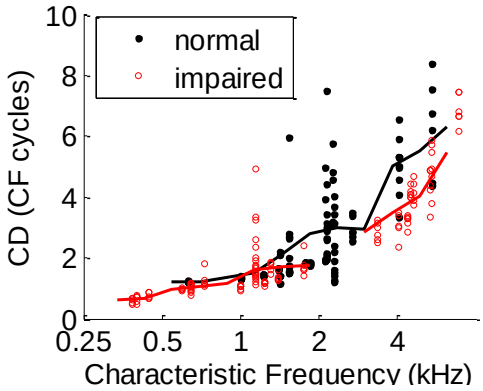

Figure 4.2 Characteristic Delay based on STMP with assumed 1 ms conduction delay (All conditions pooled.) Filled black symbols represent data from normal-hearing animals; open red symbols represent data from hearing-impaired animals. Lines indicate a moving average over 0.7 octaves. The values are qualitatively similar to those in Figure 3.3 (E,F).

A shuffled auto-correlation function (SAC) was calculated for each effective CF (based on the STMP) and a shuffled cross-correlation function (SCC) between pairs of effective CFs was calculated. The peak height of the SCC, relative to the geometrics mean of the SAC peak heights for each CF, indicates the correlation coefficient (ρ; see Equation 1 and Equation 2). This SCC peak occurs at a non-zero delay, which we refer to as the characteristic delay (CD). The CD represents the estimated traveling wave propagation time between these two CFs, the absolute value of which is expected to increase as ΔCF increases.

Figure 4.3 illustrates that the correlation patterns are qualitatively similar whether measuring a large population of fibers or predicting those responses using the STMP. These curves were generated by analyzing the output of a computational model of the auditory periphery (Zilany and Bruce, 2007b) in response to the vowel stimulus shifted such that the second formant (F2) was at six different frequencies (from 425 Hz to 2.4 kHz, in 0.5 octave steps). Each of these six stimuli were analyzed for actual model CFs spanning a range of ±1 octave (solid line in Figure 4.3), and also for a single model CF (centered on F2) and using the STMP to predict the responses of nearby effective CFs spanning the same ±1 octave range.

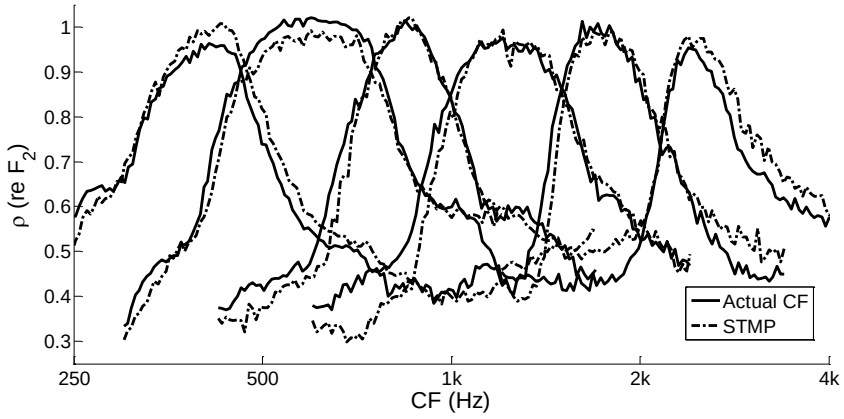

Figure 4.3. Model comparison of STMP vs. actual CFs
Correlation coefficients (ρ) for a range of CFs spanning a ±1 octave range relative to the CF corresponding to the vowel formant F2. Curves for six different F2 frequencies are shown.

4.2.2 Data Analysis

Figure 4.4 illustrates several ways in which we can quantify neural coding of vowels. The spectrum of the vowel was shifted to center either F1 or F2 (the first or second vowel formant) on the fiber CF, then the STMP was applied to obtain responses for nearby effective CFs. The first column of Figure 4.4 shows the F0-period histograms

(short-term rate versus time) for the range of effective CFs; the second column shows the average rate as a function of effective CF; the third column shows the degree of synchrony to the vowel fundamental frequency (F0); the fourth column shows the cross-CF correlation (ρ) relative to the CF centered on the vowel feature; and the last column shows the characteristic delay relative to the response for the CF at the vowel feature. The top row shows coding relative to the first formant (F1), and the bottom row shows coding relative to the second formant (F2).

The collected data was analyzed for rate-place, temporal-place, and spatiotemporal coding robustness (i.e., consistent spectral coding as the signal-to-noise ratio decreases). Rate-place coding was calculated by averaging the firing rate of each neuron over the duration of the stimulus. We calculated firing rate as a function of CF (actual or effective by STMP) for each stimulus condition.

Temporal-place coding was quantified with the *average localized synchronized rate* (ALSR; Young and Sachs, 1979). For each neuron, the synchronized rate to each harmonic was computed based on the Fourier transform of the period histogram. For each harmonic frequency, ALSR is the synchronized rate to the harmonic averaged across all CFs within ±0.5 octaves of the harmonic. ALSR was calculated as a function of harmonic number for all stimulus conditions.

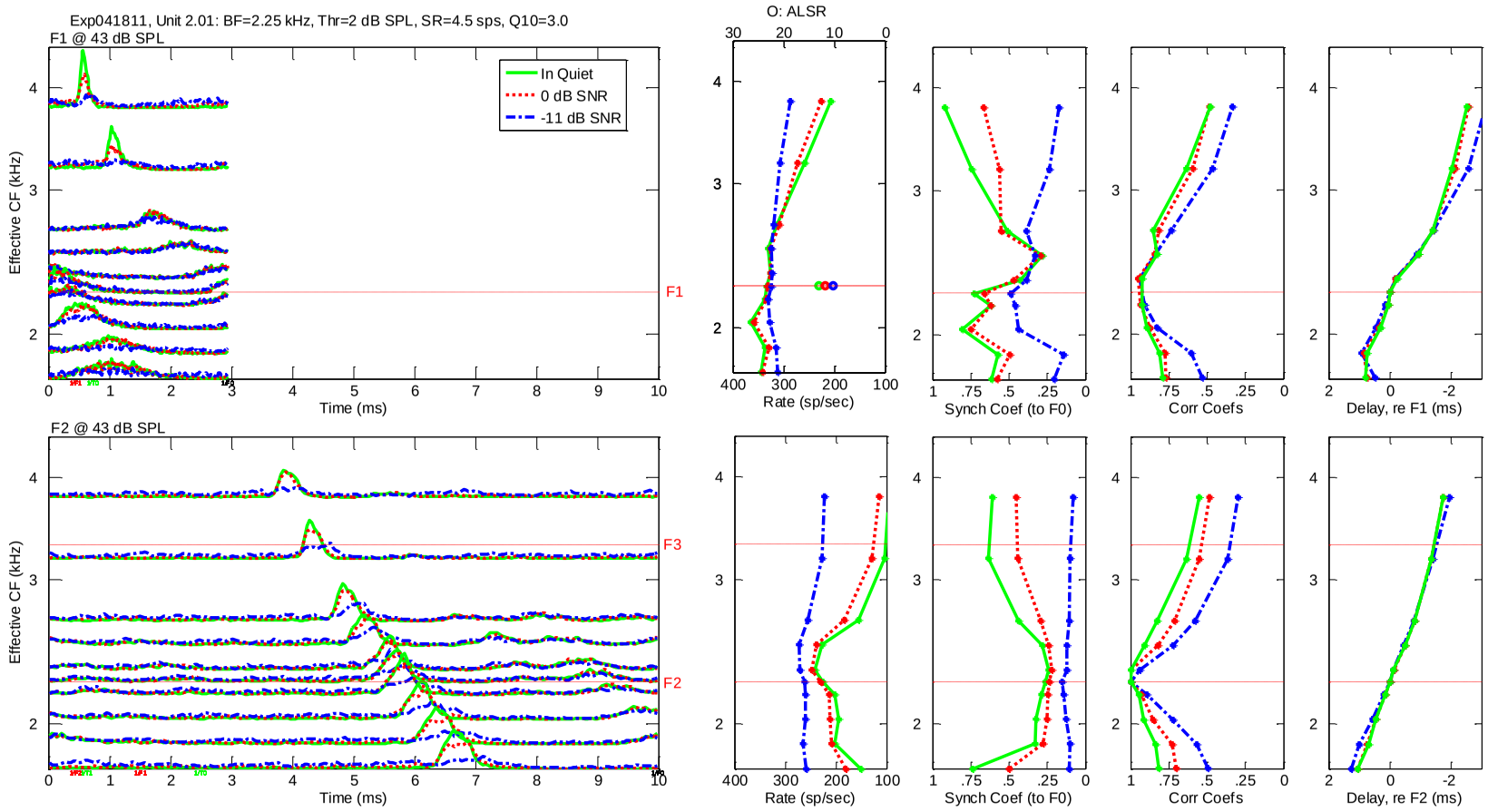

Figure 4.4 Neural response patterns to the vowel /ɛ/ at different signal-to-noise ratios. From left to right (each as a function of effective CF from the STMP): F0-period histograms, average rate, synchronization coefficient relative to the vowel F0, cross-CF correlation relative to the response for the vowel feature, and characteristic delay relative to the response for the vowel feature (note that 1 cycle at this fiber CF, 2.25 kHz, is approximately 0.4 ms). Top: relative to the first formant (F1). Bottom: relative to the second formant (F2). Vowel features (F1, F2, F3) are indicated by thin red dashed lines in each panel.

Spatiotemporal coding was characterized by the cross-CF correlation coefficient (ρ) and the characteristic delay (CD), as discussed in previous papers (Heinz and Swaminathan, 2009; Heinz et al., 2010). For each effective CF, correlations were calculated relative to the CF centered on each formant. As illustrated in Figure 4.5, the peak of each SCC was picked manually, starting at CF (no delay; thick curve) and moving outward, so as to minimize the time difference between adjacent channels.

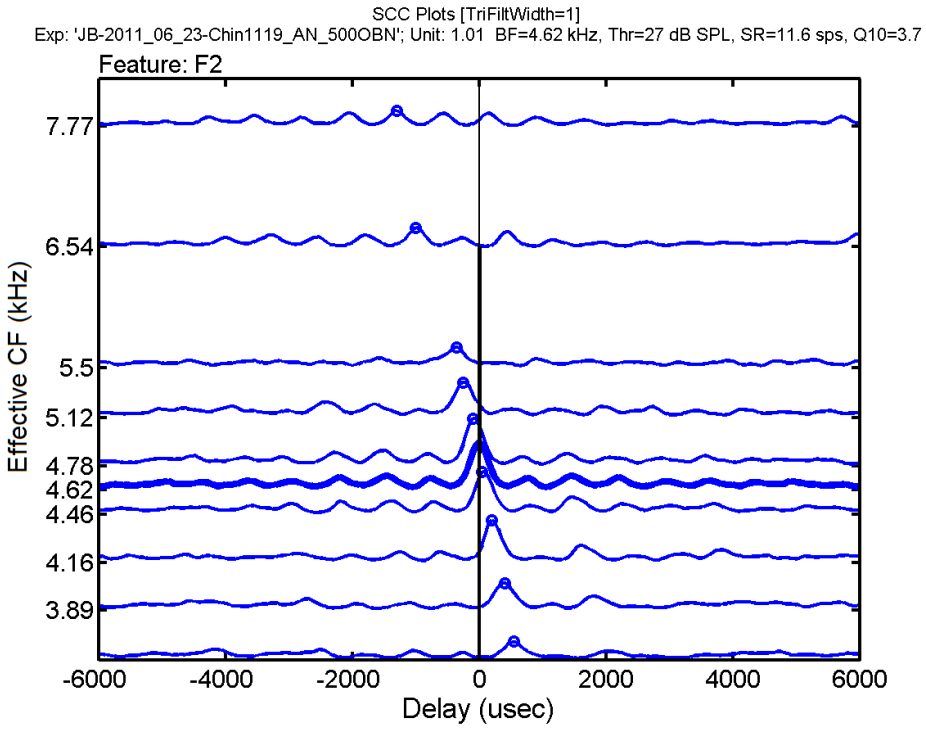

Figure 4.5. Example shuffled cross-correlogram.
Each SCC is calculated relative to the fiber's responses when the formant (F2) was centered at actual fiber CF (4.62 kHz in this example).

As in the previous chapter, the spread of correlated activity ($\rho_{0.6}$) was quantified by measuring the CF separation at which the correlation falls to a value of 0.6, and the

characteristic delay ($CD_{0.5}$) was characterized by the time delay between the CF centered on the formant peak and a CF 0.5 octaves lower (more apical).

4.3 Results

4.3.1 Synchrony

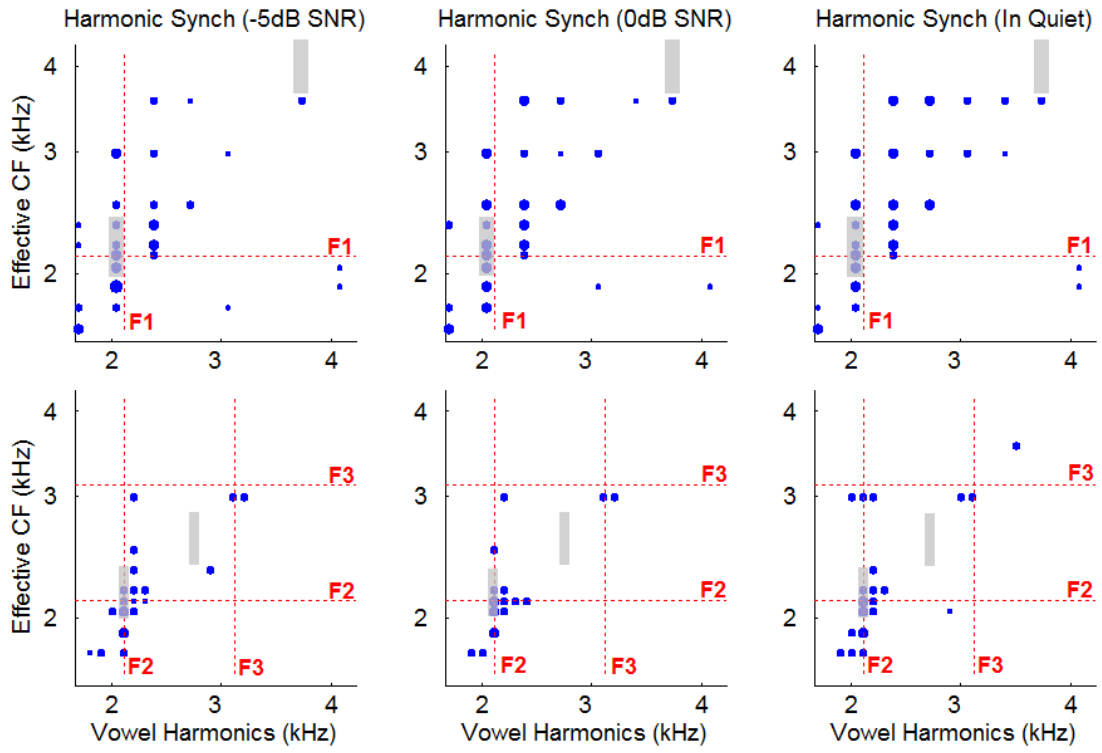

Figure 4.6. Synchrony to individual harmonics for an example unit (Normal hearing; CF=2.1 kHz) The size of each dot at a particular CF indicates the vector strength for the frequency of a vowel harmonic. Responses along the diagonal indicate tonotopic responses (synchronized to CF). The upper row represents the condition in which F1 was centerd on the fiber CF, and the bottom row represents F2 at CF. Lines indicate formant frequencies. Gray boxes indicate boundaries for calculating average localized synchronized rate (ALSR) for formant and trough freqeuncies.

Figure 4.6 illustrates how we quantified synchrony to individual harmonics of the stimulus. For each harmonic of the stimulus (plotted along the abscissa), we calculated the vector strength at that frequency for each effective CF (plotted along the ordinate).

The size of the dots indicate vector strength. In this example, we can see that a wide range of effective CFs are synchronized to the harmonic nearest to F1 at all noise levels (top row). (Note that the formant frequency does not necessarily correspond to the frequency of a harmonic.) In quiet, CFs above F_1 (in the spectral trough between F1 and F2) respond to individual harmonics, but the synchrony is reduced at higher noise levels (see upper left panel). In the bottom row, we can see that the response to F2 is localized to the fibers in that CF region (near 2.1 kHz).

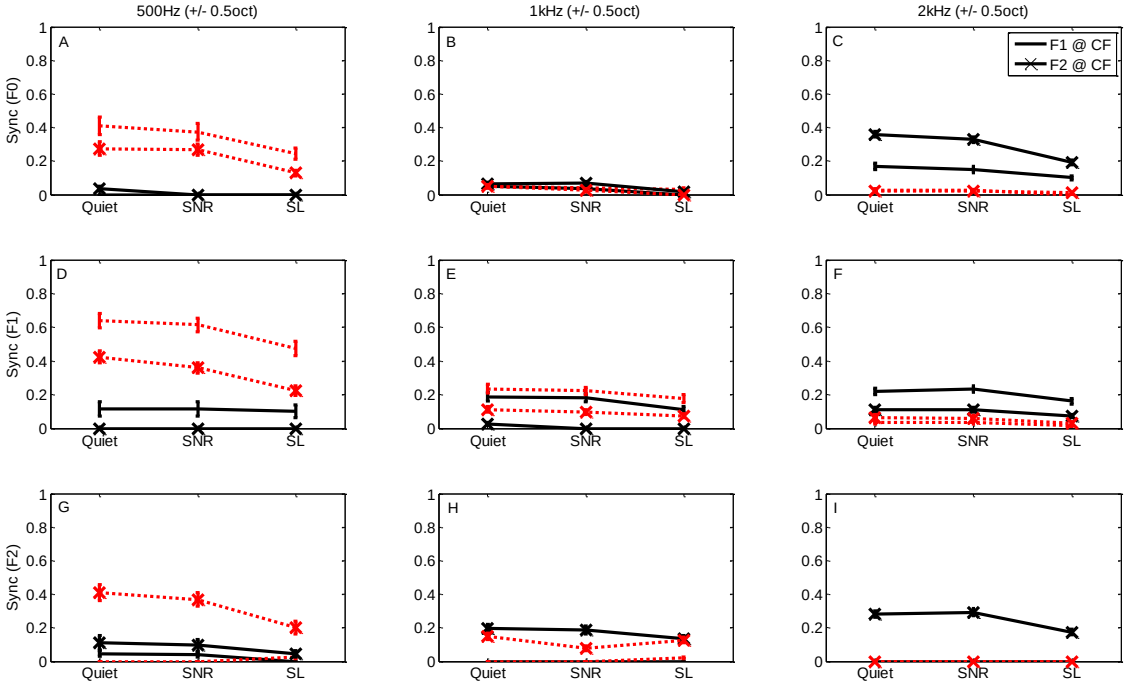

Figure 4.7 Synchrony to each formant vs. SNR for different CF regions. Solid black lines represent normal-hearing conditions, whereas dashed red lines indicate hearing impaired conditions. 'X' indicates conditions for which F_2 is centered at CF; otherwise F_1 is centered at CF. Noise levels are 'in quiet', 'equal SNR', and 'equal sensation level' (SL). Error bars indicate standard error.

Figure 4.7 shows the synchrony to F0, F1, and F2 as a function of SNR for different CF regions (including both actual and effective CFs, by STMP). Here we can see that, for

low frequency CFs, impairment (dashed lines) tends to increase the synchrony to both F0 and F1 (panels A,D). The synchrony to F_2 is only increased when F_2 is centered on CF (panel G). When F1 is centered at CF, the response of impaired fibers is dominated by synchrony to F0 and F1 (as shown in panels A,D) and not to F_2 (see panel G). At mid frequency CFs, we see that impairment slightly increases synchrony to F1 (panel E) but not to other features (panels B,H). This is consistent with the "synchrony capture" that Miller and colleagues (1997) discussed - that is, impairment causes an increased response to F1 (as seen in panel E) and individual harmonics, and a reduced response to F2 (as seen in panel H). At higher frequency CFs (panels C,F,I), impairment tends to reduce synchrony to all of these features.

4.3.2 Rate and ALSR

The strength of rate coding was quantified as the firing rate for a CF at the vowel formant relative to the rate for a CF in the spectral trough between that formant and the next highest (e.g., F1 relative to the trough between F1 and F2; LePrell et al., 1996). The strength of ALSR coding was quantified similarly – the average synchronized rate near (within 0.25 octaves) the formant relative to the average synchronized rate near the trough (regions illustrated by the gray boxes in Figure 4.6). The strength of ALSR coding was therefore the synchrony to the formant frequency relative to synchrony to the trough frequency (at their respective CF regions).

As expected and shown in Figure 4.8, we found that rate coding strength was degraded in noise (black circles with solid line), and also with impairment (red circles with dashed line). As shown in Figure 4.8, ALSR coding strength was much more robust

to both impairment (comparing black squares with solid line to red squares with dashed line) and noise (comparing squares at different noise levels).

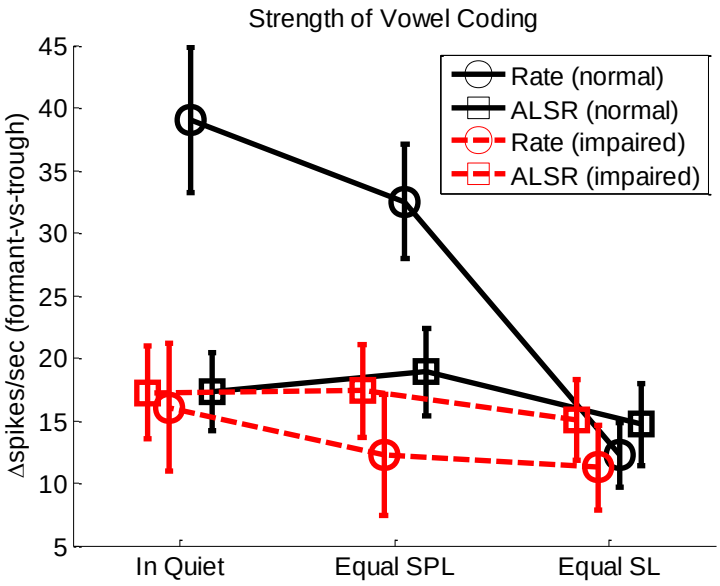

Figure 4.8 Rate and ALSR coding (mean ± standard error)
Data for the first two formants (F1 and F2) and all AN fibers are pooled.

4.3.3 Spatiotemporal Coding

The width of the correlated activity patterns are shown in Figure 4.9. The correlation width is quantified as the maximum difference (in octaves) between the points at which $\rho \geq 0.6$ (as illustrated in Figure 3.2). Here, we can see that the correlation width around F1 does not change much with impairment (upper panel), but we see wider areas of correlated activity around F2 with impairment (lower panel). For normal hearing, the correlated regions appear to become more narrow with noise, but we did not see this trend in the impaired data.

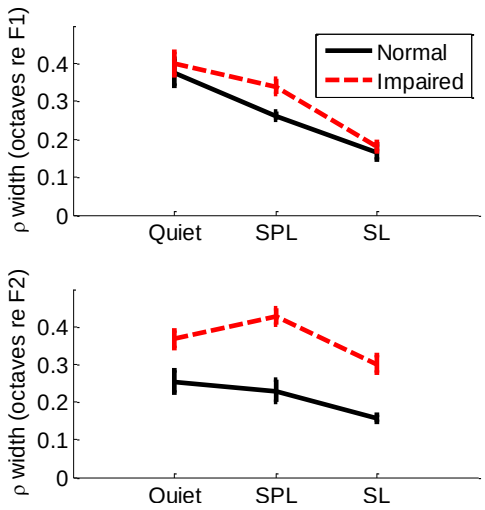

Figure 4.9. Width of correlated activity region (CF = 1 kHz ± 0.5 octaves)
Error bars indicate standard error.

Figure 4.10 shows characteristic delay as a function of CF. Here, the characteristic delay is quantified as the delay at a CF one-half octave away from the formant. The top row shows CD relative to the first formant; the bottom row shows CD relative to the second formant. The columns show CD at each of three noise levels. Lines indicate the median value within a 3/4-octave band. We can see that, near F2, characteristic delays are reduced (i.e., a faster traveling wave) for impaired conditions (shown in red), particularly for CFs in the 1-2 kHz region. A Wilcoxon rank sum test indicated a significant difference between normal and impaired, for CFs centered on F2 near 1 kHz (± 0.5 octaves; $p<0.05$).

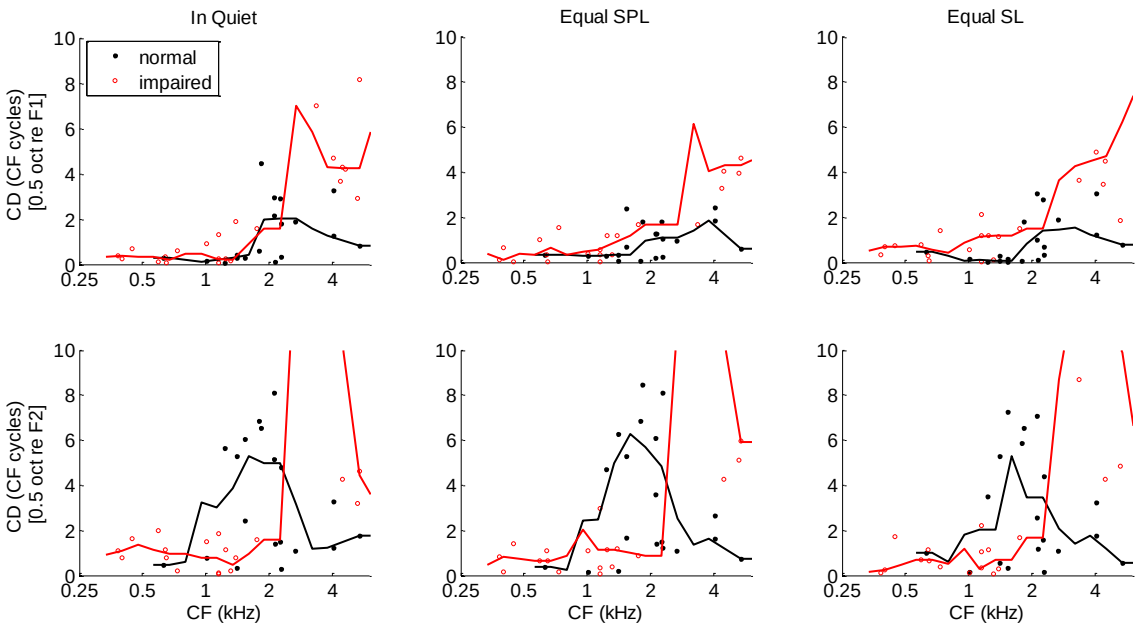

Figure 4.10. Characteristic delay (CD) as a function of characteristic frequency Top row: delay 0.5 octaves from CF for conditions in which F1 was centered at CF; Bottom row: delay when F2 was centered at CF. Lines indicate a median filter with a window of 0.75 octaves

4.4 Discussion

We have shown a measureable difference between normal and impaired spatiotemporal coding of a vowel. We did not observe any effect of noise on spatiotemporal coding, suggesting that spatiotemporal coding is robust to noise but not to impairment. This is consistent with previous studies that showed that the average localized synchronized rate is also fairly robust to noise (Heinz, 2007; Sachs et al., 1983).

4.4.1 Limitations of the Correlation Coefficient (ρ)

In the previous chapter, we showed that, for broadband signals, the correlated activity pattern gets wider with impairment. As shown in Figure 4.11, we found the

same widened pattern in this vowel data. This was previously interpreted as evidence of broadened auditory filters, with consequences in terms of the number of independent channels of information for the brain. However, the data presented in this chapter suggest that the interpretation may not be so simple for complex sounds with spectral structure in the presence of noise. We found that the correlation patterns often get narrower as the noise level is increased (and the vowel level is held constant). Because the overall level is actually increasing slightly, we expect the auditory filters to get broader, in contrast to the across-CF correlation patterns getting narrower. This effect is likely due to the widespread effects of synchrony capture (in quiet) creating wide correlation regions when there are specific spectral features (e.g., formants) that engage a wide region of the cochlea. The addition of noise, however, may act to decorrelate the signal, resulting in a narrower correlated region.

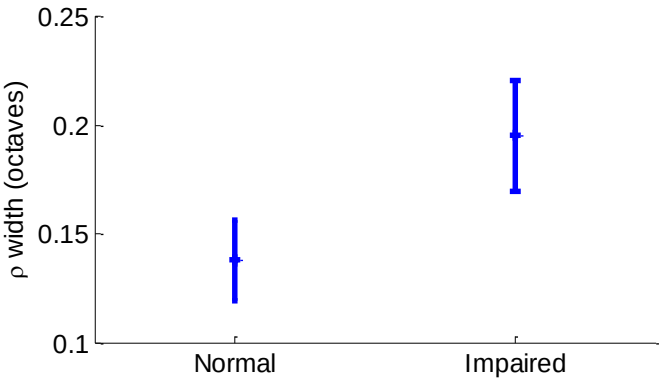

Figure 4.11 Correlation width of normal vs. impaired (μ ± σ)
Population data for 19 normal-hearing units and 22 hearing-impaired units. The difference was statistically significant (Wilcoxon rank sum; $p<0.05$)

4.4.2 Characteristic Delay

The characteristic delay measurements were performed manually because we were unable to design a peak-picking algorithm that would reliably pick realistic peaks in all conditions. It is a difficult problem because the shuffled cross-correlation functions are often nearly periodic, with multiple peaks of approximately equal amplitude. We had a rigorous rule for picking these peaks manually (as discussed in section 4.2.2), but this may underestimate the actual delay if the signal is periodic (e.g., high SNR). The STMP spacing was wider at shifts furthest away from CF, so errors are most likely near the edges. For a 1 kHz CF, sufficient ringing at this frequency would result in a periodicity of 1 ms, so any error in picking the SCC peak would be seen as an underestimated CD by 1 ms. However, the same peak-picking rule was applied for all data (i.e., for both normal hearing and hearing impaired data), so we do not believe this error affected our results in any significant way.

Characteristic delay (in units of cycles at CF) increases with frequency consistent with sharper tuning at higher frequencies (Shera et al., 2002), but decreases with impairment, consistent with Heinz et al (2010). Whereas Heinz et al (2010) only quantified the effect of impairment for noise and a speech sentence, we see reduced delays with impairment for simple vowel sounds as well. If spatiotemporal coding is important for pitch-based segregation (e.g., Larsen et al., 2008), and if pitch is important for concurrent vowel identification (e.g., Keilson et al., 1997; Summers and Leek, 1998), then any degradation in spatiotemporal coding for vowels (as shown here) might result in a reduced ability to identify concurrent vowels. This could be detrimental for

listening to speech in complex listening situations, such as when multiple people are speaking simultaneously. If this is indeed the case, then we would really like hearing aids to restore at least some of the lost spatiotemporal coding. The next chapter quantifies spatiotemporal coding when a hearing aid is added to the impaired system.

CHAPTER 5. EFFECTS OF AMPLIFICATION ON SPATIOTEMPORAL CODING

A portion of this work was presented as the following conference poster:

***Boley, J.** and M. Heinz, Impaired Spatiotemporal Coding of Vowels in Noise, International Hearing Aid Research Conference, August 2012.*

5.1 Rationale

Schilling et al (1998) analyzed auditory nerve responses to vowels that had been amplified with a frequency-dependent gain similar to that used in linear hearing aids. They quantified the temporal-place representation of the vowel /ɛ/ and showed that this amplification strategy may improve neural coding of F1, but it does not appear to improve coding of higher formants. (The spread of F1 synchrony was reduced, but not the spread of F2 or F3 synchrony.) However, this can be controlled by applying gain to enhance the spectral peaks relative to troughs. This technique, called *spectral contrast enhancement*, applies gain based on knowledge of the stimulus rather than just knowledge of the hearing loss. Although spectral contrast enhancement techniques have been shown to improve both rate and temporal-place representations of the second formant (Miller et al., 1999b) and may even benefit from multiband compression (Bruce, 2004), similar techniques have had mixed results in perceptual experiments

(Baer et al., 1993; DiGiovanni et al., 2005; Franck et al., 1999; Simpson et al., 1990). A thorough analysis of vowel coding with modern amplification algorithms in quiet and in noise would be beneficial because it would allow us to see the effects of these strategies on spatiotemporal coding.

There is very little data on the relationship between hearing aid gain and neural coding. One of the fundamental characteristics of typical hearing aid amplification (as described in section 1.1.3) is a frequency-dependent gain, often boosting high frequencies (where the hearing loss is often greater). Schilling and colleagues (1998) showed that this strategy can improve some aspects of neural coding, but can also cause undesired distortions (e.g., sharp spectral changes may look like formants). Although the authors noted that amplification did not limit spread of F2 synchrony to higher CFs, they only characterized the effect in quiet and did not evaluate the effect of background noise on coding in this region. The present study was designed to characterize the effects of both a simple gain filter and a modern multichannel wide dynamic range compression algorithm on neural coding of vowel formants in background noise and to specifically characterize the effects of each on spatiotemporal coding.

5.2 Modeling Study

We used a recent computational model of the auditory nerve (Zilany & Bruce, 2007). This model allows selective control over the health of both outer and inner hair cells. As reviewed in section 1.1.2, outer hair cells provide gain and sharp tuning, while inner hair

cells transduce the acoustical energy to electrical signals, and thus damage to outer versus inner hair cells can have different effects on neural coding.

For impaired hearing simulations, the coefficients C_{OHC} and C_{IHC} were chosen to result in a mild hearing loss, as shown here (where 1/3 of the threshold shift in dB is modeled as inner hair cell dysfunction, 2/3 as outer hair cell dysfunction).

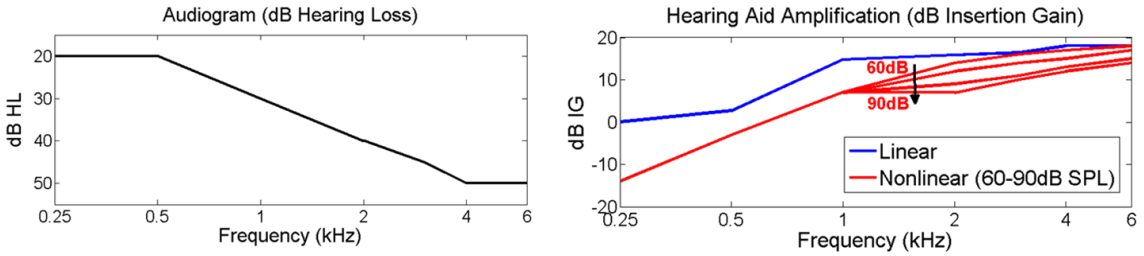

Figure 5.1 Model audiogram and hearing aid gain prescriptions
audiogram in left panel; linear (NAL-R; blue) and nonlinear (DSL [i/o]; red) gains (right panel)

A hearing aid gain profile was fit to this audiogram, based on the NAL-R linear prescription (Dillon, 2001) or the DSL [i/o] nonlinear (compressive) prescription (Scollie et al., 2005). Both prescriptions are used clinically, and are thus important to study. DSL[i/o] is a proprietary algorithm, but NAL-R is defined by a simple equation:

$$IG_i = 0.15\frac{H_{500} + H_{1k} + H_{2k}}{3} + 0.31H_i + k_i \qquad \text{Equation 5}$$

where IG_i is the insertion gain at a specific frequency, H_i is the audiometric hearing loss at a specific frequency (in dB), and k_i is a constant (in dB) defined in Table 1.

Table 1 NAL-R constants

Freq (Hz)	**250**	**500**	**1000**	**2000**	**3000**	**4000**	**6000**
k_i (dB)	-17	-8	1	-1	-2	-2	-2

DSL [i/o] uses different gains in the presence of background noise, so the algorithm was set to use a noise prescription with 2 frequency channels and a cross-over frequency of 2.1 kHz, such that F2 (1.7 kHz) and F3 (2.5 kHz) were in different frequency channels. Figure 5.2 shows the gains applied at each formant frequency for the various noise levels.

The vowel /ɛ/ was then run through the model with 4 separate model and hearing aid parameter sets: a normal-hearing case, an impaired case, and an impaired case with each hearing aid. The resulting neural spike patterns were than analyzed for comparisons across these four cases to quantify the ability of the hearing aid to restore normal temporal coding.

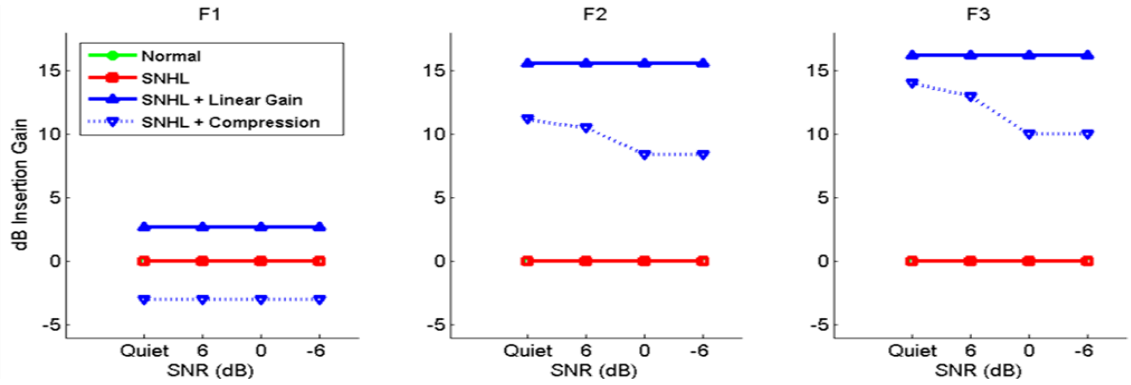

Figure 5.2 Gains applied at the first three formants (F1, F2, F3) of the vowel /ɛ/ [F1 = 0.5 kHz; F2 = 1.7 kHz; F3 = 2.5 kHz]

The same cross-CF correlation analyses were used here as in the previous chapter. Figure 5.3 shows the correlation patterns (A) and the characteristic delay (B) relative to each formant. For the normal hearing model (green circles), the correlations tend to drop rather quickly as we compare to characteristic frequencies (CFs) both above and

below the formant frequency. As expected for a cochlea with broadened tuning, these correlation patterns are much broader for the unaided impaired system (red circles). Neither linear (blue solid line) nor nonlinear (blue dashed line) hearing aid prescriptions appear to restore these patterns to normal for any of the formants.

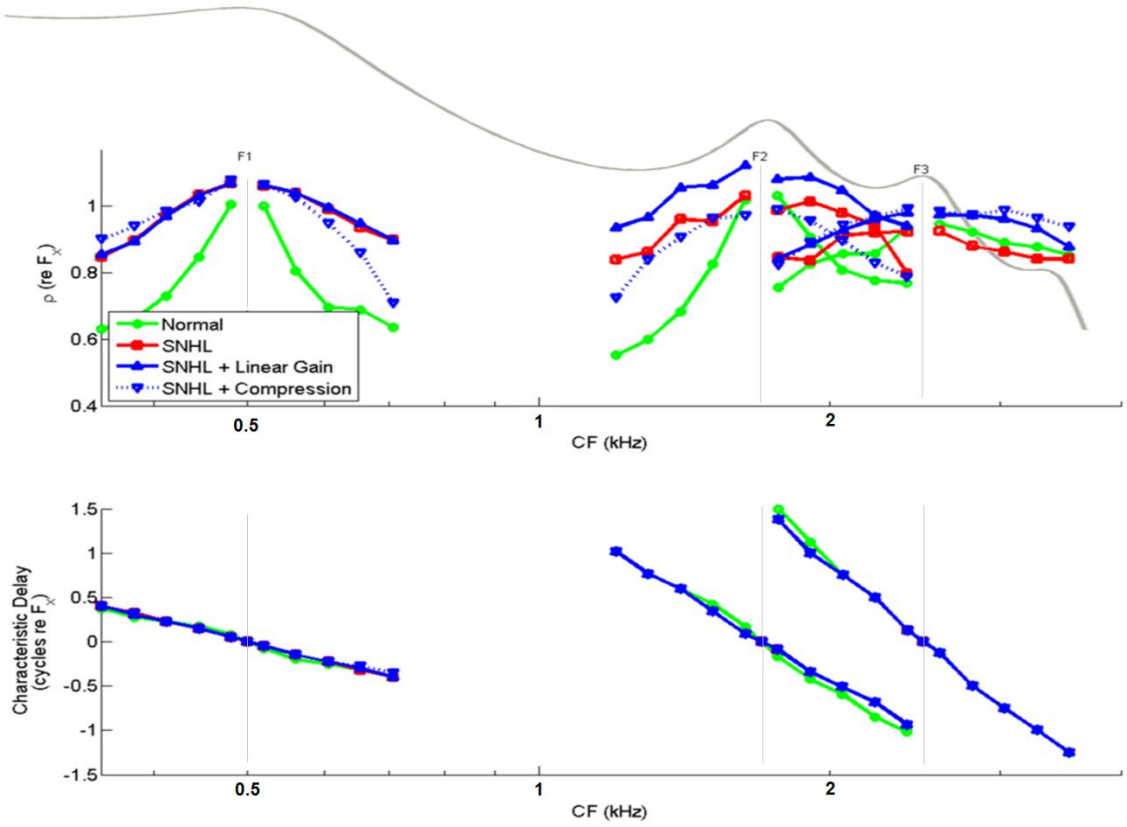

Figure 5.3 Spatiotemporal response patterns for the vowel /ɛ/ in quiet
Spectral envelope overlaid in gray for reference. Cross-fiber correlations (A) and characteristic delays (B) relative to each formant. Characteristic delays are plotted in cycles relative to the formant frequency of interest (F1, F2, or F3).

Figure 5.4 shows similar patterns for the vowel in noise. In this case, white noise was set to achieve a signal-to-noise ratio of -6dB at CF.

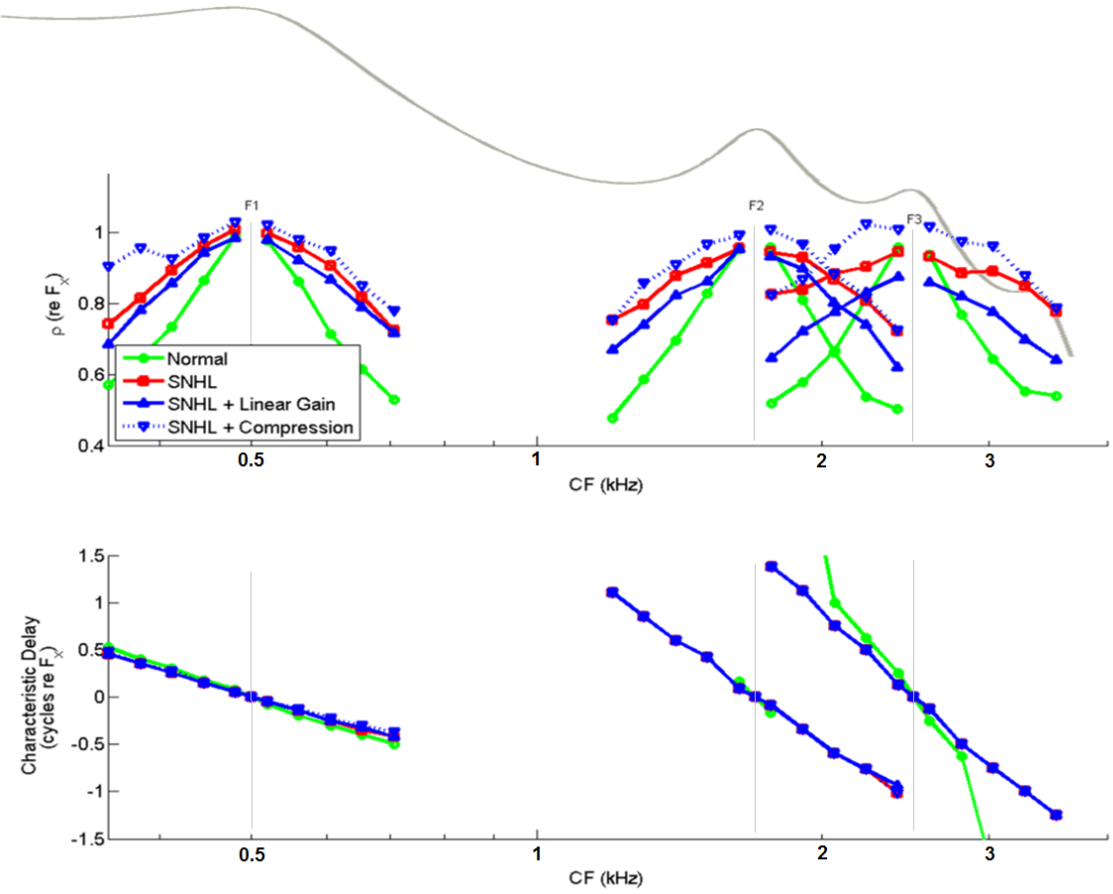

Figure 5.4 Spatiotemporal response patterns for the vowel /ɛ/ at -6dB signal-to-noise ratio (spectral envelope overlaid in gray for reference). Cross-fiber correlations (A) and characteristic delays (B) relative to each formant. Characteristic delays are plotted in cycles relative to the formant frequency of interest (F1, F2, or F3).

From these figures, we can see that impairment generally results in a wider spread of correlated activity near each formant. This is particularly apparent for F1 and F2 in quiet, and for all 3 formants at -6dB SNR. Neither linear nor compressive hearing aid prescriptions restore the cross-fiber correlations to normal, either in quiet or in noise.

Also note that the characteristic delay functions are generally shallower around the formant (vertical gray line) with impairment, suggesting that impairment speeds up the

traveling wave within the cochlea. This is a bit difficult to see in the figure, but a measurement of the slope indicates a difference (as shown in Figure 5.5). Neither linear nor compressive amplification restore these phase responses to normal.

Figure 5.5 shows a quantitative analysis of both the width of correlated activity and the slope of the characteristic delay function. The width was calculated as the number of octaves over which the correlation drops to 80% of the peak value. The slope was quantified over the center 0.1-octave range surrounding each formant. Here, we see that the width is greatly increased with impairment. Amplification does not appear to improve the spatiotemporal coding in quiet, but linear gain appears to help (at least somewhat) in noise. The spatiotemporal response (as quantified by both ρ and CD) to compression was equivalent to linear gain only for the 2^{nd} formant, implying that compression can improve F2 coding. The slope of the phase response is reduced with impairment, and neither amplification scheme restored the slope to normal for any of the tested conditions.

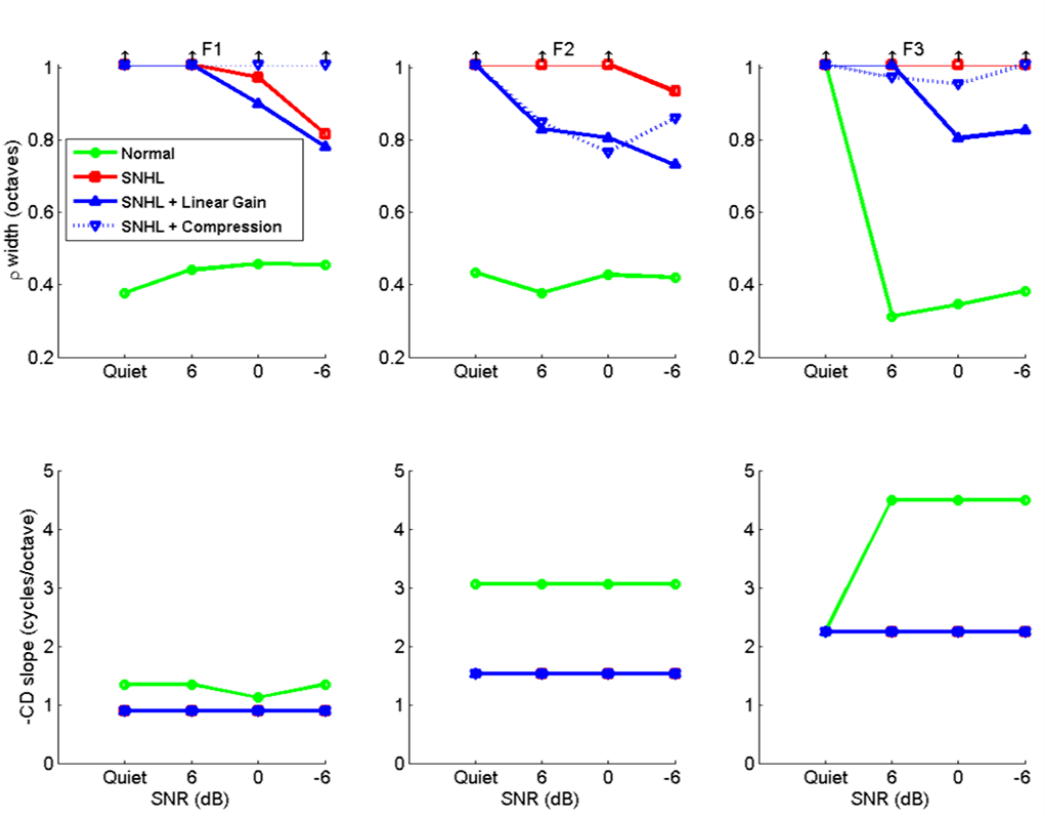

Figure 5.5 Width and slope of the spatiotemporal response patterns. Correlated activity width and characteristic delay slopes at 4 noise levels for normal hearing (green), SNHL (red), SNHL with linear gain (blue solid lines), and nonlinear gain (blue dashed lines). Width was quantified as the bandwidth over which the neural activity was correlated to the activity at the formant (ρ>0.8). Note that all SNHL data for CD lie on top of one another.

Of particular interest here is the shift in CD seen with impairment. Consistent with Heinz and colleagues (2010), we see an increase in CD shift with increasing CF for this simple vowel. At F1 the shift is 0.22 cycles per half octave (half of 0.44 cycles per octave); at F2 the shift is 0.76 cycles per half octave; at F3 the shift is 1.13 cycles per half octave.

5.3 Animal Study

5.3.1 Prescriptive Fitting of Hearing Aids

Using pooled data from 19 animals in our lab with noise-induced hearing loss (using the same noise exposure protocol, see section 1.2.2) and 67 animals with normal hearing, we calculated an average "audiogram" based on the minimum auditory-nerve thresholds of these animals. Some of the AN threshold data were discarded as outliers and not included in the estimate of threshold shift. Any data from an unexposed animal for which at least 25% of the tuning curves were abnormally broad were classified as outliers. That is, where Q10 was less than the 5th percentile for the normal data in (Kale and Heinz, 2010). None of the animals from which we collected vowel coding data for this experiment fit this criterion. Thresholds for normal AN fibers ranged from 0dB to 35dB SPL, and thresholds for impaired fibers ranged from 27dB to 49dB SPL.

This threshold shift serves as a model of the expected hearing impairment for the animals in this study, and was used to calculate hearing aid prescriptions. Figure 5.6 shows the individual auditory nerve tuning thresholds (upper panel) and tuning sharpness (Q10; lower panel). Data from normal-hearing animals are indicated by gray filled symbols, whereas data from noise-exposed animals are indicated by red open symbols. Solid lines connect the lowest threshold within each 1-octave band, and dashed lines connect the average threshold within each band.

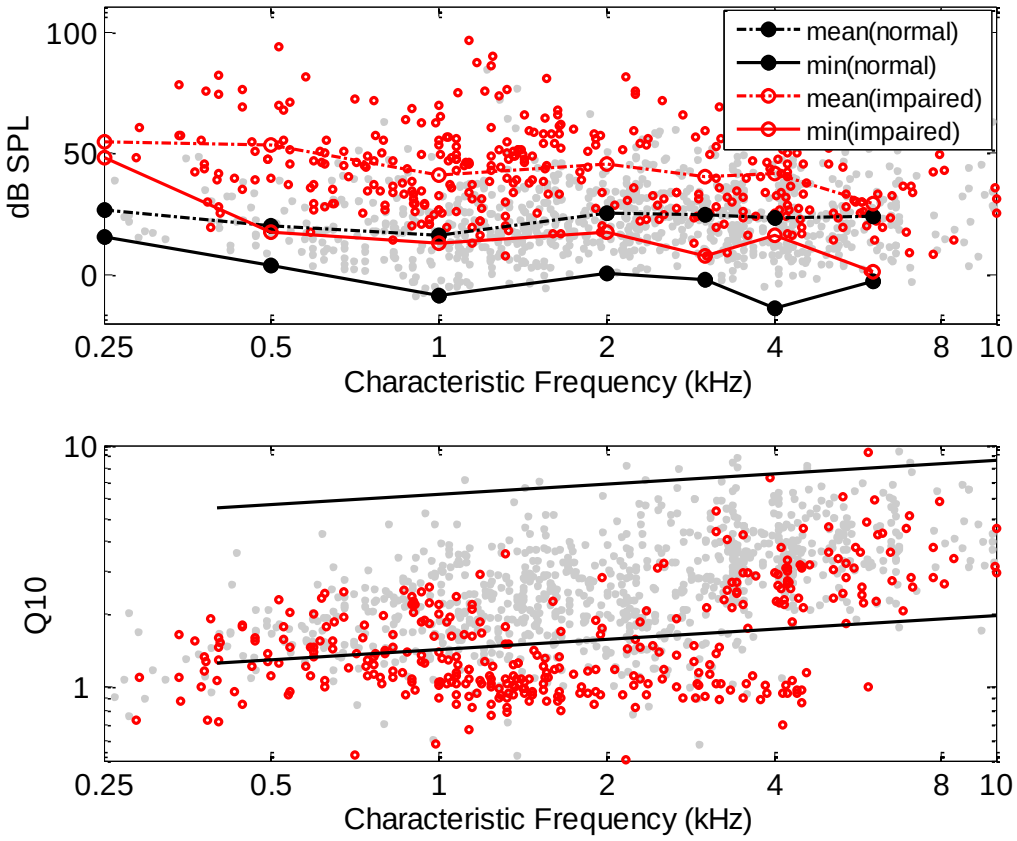

Figure 5.6. Auditory nerve thresholds (dB SPL) and tuning sharpness (Q10). Data from 45 non-exposed (filled gray symbols) & 19 noise-exposed animals (open red symbols). In upper panel, solid lines connect the lowest threshold within each 1-octave band, and dashed lines connect the average threshold within each band. In lower panel, lines indicate 5th and 95th percentiles of normal Q10 values found in Kale and Heinz (2010)

We also collected auditory brainstem responses from several of the noise-exposed animals, both before and several weeks after exposure. The ABR threshold shifts (see methods in Henry et al., 2011) and AN threshold shifts are shown in Figure 5.7.

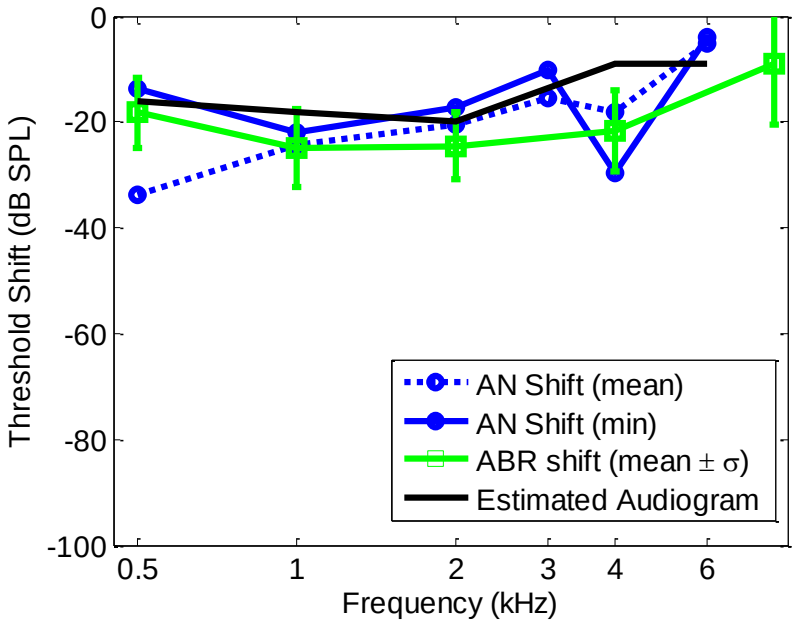

Figure 5.7. Threshold shifts of auditory nerve fibers and auditory brainstem responses (error bars indicate standard deviations for ABR threshold shifts)

Based on the combined AN and ABR data, we assumed an average audiogram of 16, 18, 20, 9, and 9dB at 0.5, 1, 2, 4, 6 kHz respectively, as shown in Figure 5.7 (solid black line). This audiogram represents the expected average audiogram for a noise-exposed chinchilla using our protocol, and was used for all animals due to the difficulty of estimating individual audiograms prior to completion of the acute AN experiment. This audiogram was used for determining hearing aid prescriptions. Note that this estimated audiogram is a conservative estimate, as the measured threshold shift is somewhat greater at some frequencies. The resulting hearing aid prescriptive gains will therefore be conservative as well (less gain than would otherwise be prescribed).

5.3.2 Experimental Procedures

The experimental procedures are identical to those described in chapter 4, but stimuli were amplified by one of two gain functions. Four chinchillas that has been exposed to noise were presented with amplified vowels in noise. The NAL-R formula (Dillon, 2001) was used to calculate a frequency-dependent insertion gain that was implemented as a symmetric (linear phase) 32 tap FIR filter in Matlab. (Note that linear phase filters have a constant delay as a function of frequency.) The DSL [i/o] fitting algorithm (Scollie et al., 2005) was used to calculate the prescriptive gain settings that would be used in a wide-dynamic-range compression (WDRC) hearing aid. This gain was also implemented as a symmetric FIR filter in Matlab, which is equivalent to assuming that we have an algorithm with fast time constants such that the exact target gain will be applied throughout the steady state vowel. Both algorithms were fit to the average threshold shift for chinchillas with noise-induced hearing loss (the audiogram shown in Figure 5.7). The DSL prescriptive gains were calculated for a 2-channel algorithm with a cutoff frequency of 922 Hz, such that the first two vowel formants would be in separate compression channels. (Note that this cutoff frequency was placed between F2 and F3 for the modeling study in section 5.2, but was placed between F1 and F2 for the animal study.) A complete set of DSL[i/o] filters was designed, corresponding to prescribed gains for input levels ranging from 0dB SPL to 100dB SPL, and the appropriate filter was used for each stimulus condition. The acoustic system was calibrated for each experiment, so the target levels are real-ear levels.

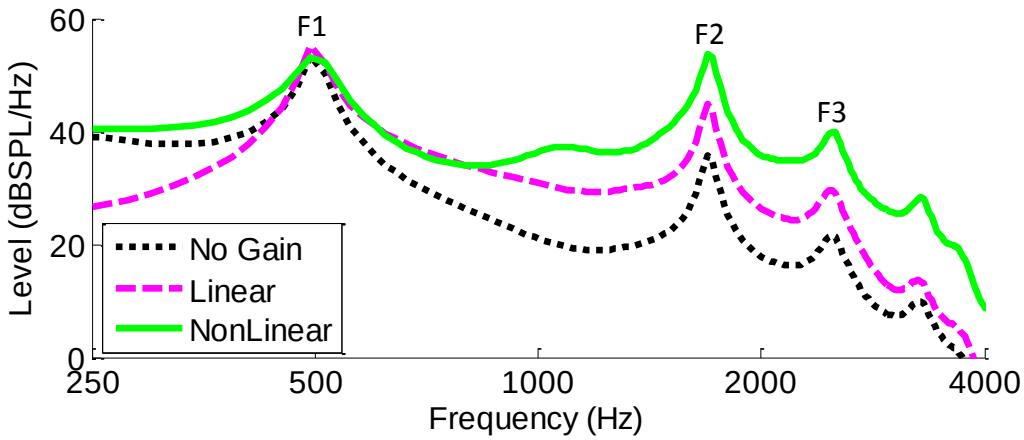

Figure 5.8 Example vowel spectra
(no gain, NAL linear gain, and DSL nonlinear gain)

Examples of the amplified vowel spectra are shown in Figure 5.8. The NAL (linear) gain prescription actually reduces the level below the first formant, but increases gain somewhat at higher frequencies. The DSL (nonlinear) gain prescription only increases gain above the first formant, applying more gain than the linear prescription for this particular input level. The amount of nonlinear gain depends on the presentation level and, as illustrated in Figure 5.9, the nonlinear prescription resulted in more gain at the second formant than the linear prescription for all conditions tested.

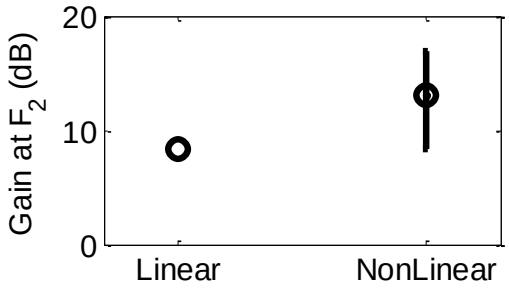

Figure 5.9 Hearing aid gains (at second formant)
Symbol indicates the mean; error bars indicate range [minimum,maximum]

5.4 Results

Tuning thresholds and sharpness for AN fibers used in this study are shown in Figure 5.10. The broad CF region around 1 kHz (±1.25 octaves) contains 15 normal-hearing AN fibers, 13 impaired, 5 impaired with linear amplification, and 6 impaired with nonlinear amplification.

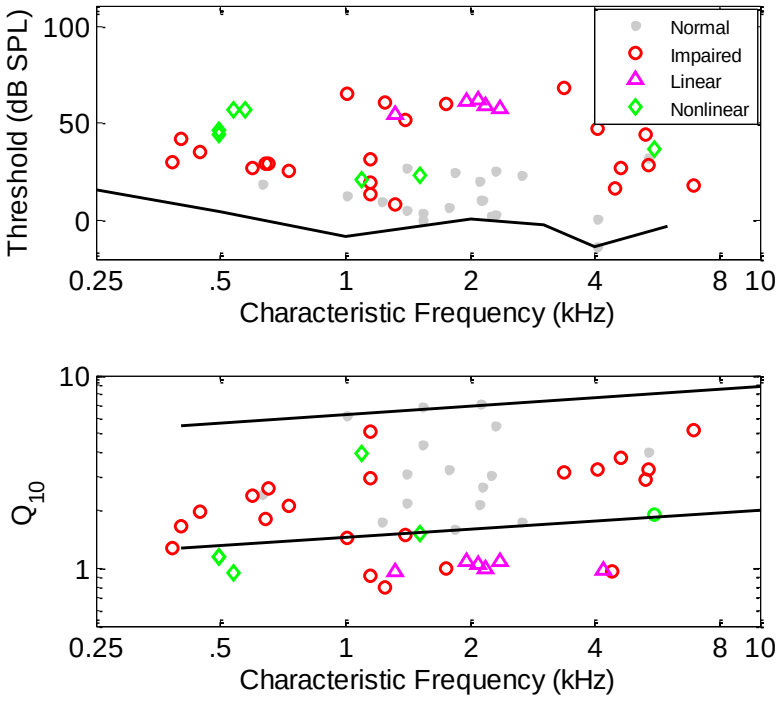

Figure 5.10 AN tuning thresholds and sharpness
AN fibers used for collecting vowel STMP data. Normal (gray dots); Impaired (red open circles); Impaired + linear amplification (magenta triangles); Impaired + nonlinear amplification (green diamonds)

As in the previous chapter, we used STMP with each AN fiber to predict a range of effective CFs, from which spatiotemporal coding could be quantified for each of the amplified conditions. We measured the width of correlated regions (at ρ=0.6) for the

two formant conditions and three noise conditions used in the previous chapter. Figure 5.12 shows the widths of correlated activity measured for normal, impaired, impaired with NAL (linear amplification), and impaired with DSL (nonlinear amplification).

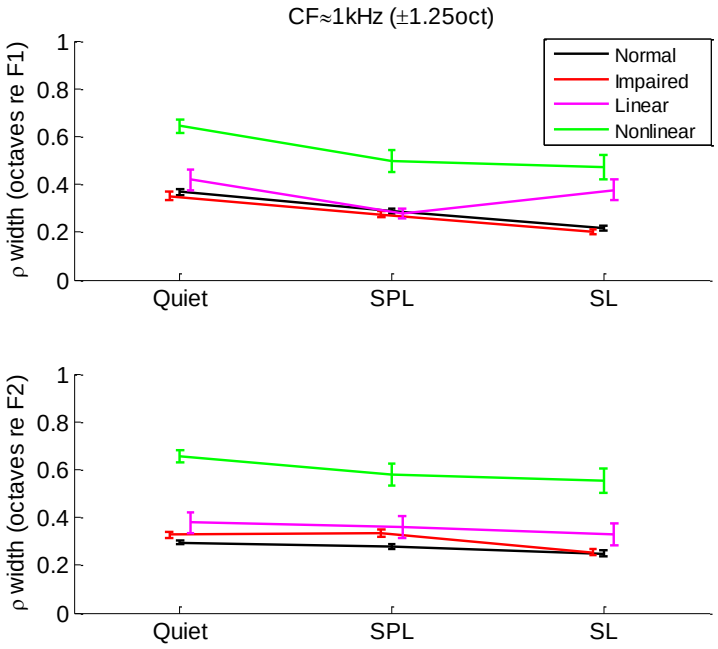

Figure 5.11 Width of correlated activity region (CF = 1 kHz ± 1.25 octaves) Upper panel: correlation width near F1; lower panel: correlation width near F2. Error bars indicate standard error.

Nonlinear amplification resulted in wider regions of correlated activity than any of the other conditions. This was not surprising, as the greater gain is expected to result in wider auditory filters.

We measured characteristic delays at 0.5 octaves from CF for the two formant conditions and three noise conditions used in the previous chapter. Figure 5.12 shows the characteristic delays measured for normal, impaired, impaired with NAL (linear amplification), and impaired with DSL (nonlinear amplification). Lines indicate trends

using an octave-wide median filter. CD in the 1-2 kHz region is higher for F2 than F1, suggesting sharper cochlear tuning for F2.

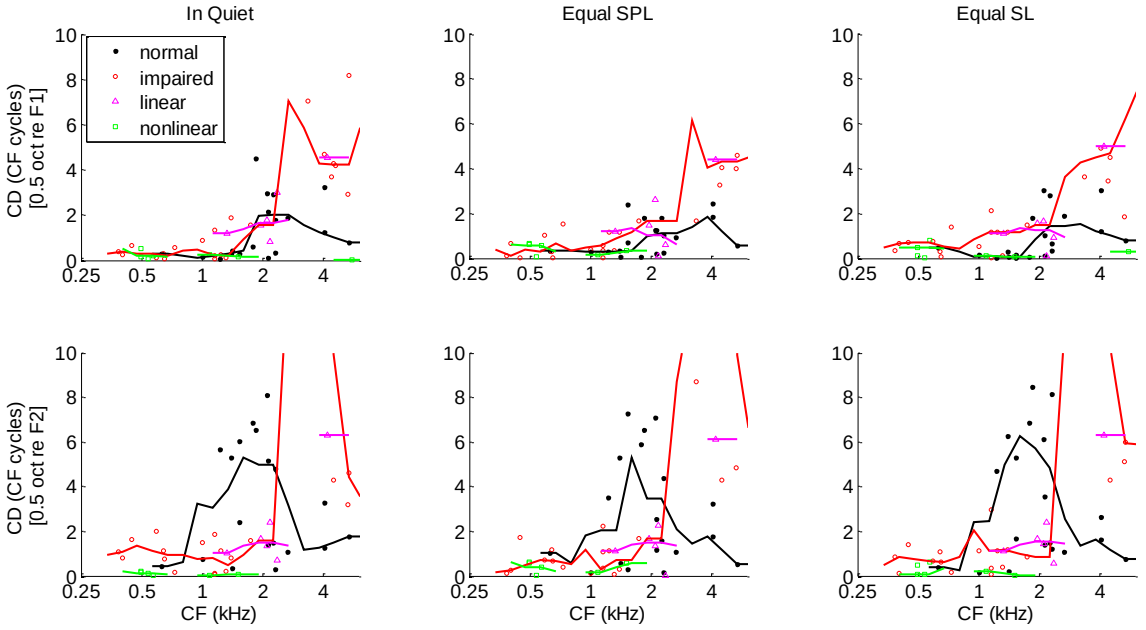

Figure 5.12. Characteristic delay (CD) as a function of CF for several conditions Top row: CD at 0.5 octaves from F1; bottom row: CD at 0.5 octaves from F2. Columns represent the vowel in quiet, with noise at equal SPL (0dB SNR), and noise at equal "sensation level" (SL).

When we calculate the average characteristic delay for the population of CFs in the 500Hz to 2 kHz range (for both F1 and F2 combined), we find the values shown in Figure 5.13. A Wilcoxon rank sum test indicated a significant difference between normal and impaired in noise (both noise conditions combined; $p<0.001$), but no significant differences between impaired and either aided condition. In quiet and in noise, nonlinear gain resulted in significantly less cross-CF delay than linear gain ($p<0.05$).

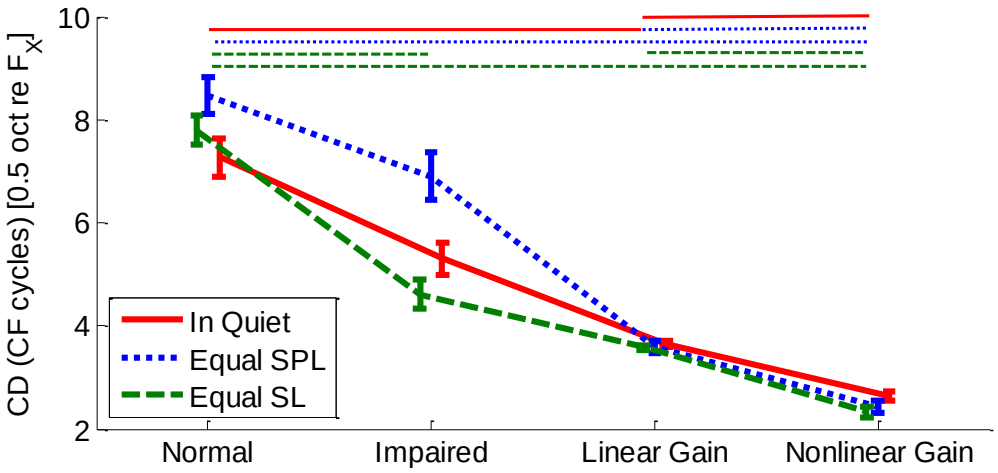

Figure 5.13. Mean characteristic delay (± standard error)
CD is quantified at 0.5 octaves from the formant for different conditions (CF=0.5-2 kHz). Upper lines indicate significant differences (Wilcoxon rank sum $p<0.05$)

5.5 Discussion

We have shown that hearing aid amplification does not improve spatiotemporal coding. In fact, the nonlinear prescription resulted in both more gain ($p<0.001$) and less characteristic delay than the linear prescription. This is consistent with the idea that increased sound levels result in broader auditory filters, which have less phase delay and group delay. It also implies that hearing aids may not improve any aspect of perception that depends on spatiotemporal coding. For example, people with hearing impairment have difficulty segregating multiple sources of speech even when amplification is used to ensure audibility (Rossi-Katz and Arehart, 2005; Summers and Leek, 1998).

5.5.1 Estimated Audiogram

Pickles (1988, p. 84) showed that normal behavioral hearing thresholds tend to be close to the lowest AN thresholds, but ABR thresholds have been shown to change less than AN thresholds with SNHL (Henry et al., 2011; Ngan and May, 2001). Additionally, although we chose a reasonable criterion for defining outliers in the AN threshold data, this procedure may have resulted in an overestimated average difference between normal and impaired thresholds. In other words, an audiogram based on average AN thresholds may also slightly underestimate hearing loss as measured by minimum AN thresholds. However, we did not depend on the average AN thresholds alone, but also the minimum thresholds across the population to quantify threshold shifts.

Because our estimated audiogram (used for determining target hearing aid gain) was based on a combination of both AN and ABR threshold shifts, we may not have used the exact same gains that would have been prescribed to these animals based on behavioral threshold shifts. It is possible that greater gain may have had a larger impact, but it would likely only degrade the spatiotemporal coding as the auditory filter would get broader with increased sound levels.

Interestingly, although we saw threshold shifts extending to near 6 kHz, some AN tuning curve bandwidths were in the normal range as low as 3 kHz. This suggests outer hair cell dysfunction occurred primarily in the range of 1.5-3 kHz (where Q10 was reduced, as shown in Figure 5.6), and perhaps not in surrounding regions, where the threshold shift appears to have occurred due to primarily inner hair cell dysfunction (where Q10 was normal). This inferred pattern of inner and outer hair cell dysfunction

is consistent with a previous study (Liberman and Kiang, 1984) which found a broader region of IHC damage than OHC damage following noise exposure.

5.5.2 Compression Speed

The stimuli used here were based on a static gain adjustment, not a time-varying gain as found in most modern hearing aids. This basically assumes that we have an algorithm with fast time constants such that the target gain will be applied effectively throughout the entire steady state vowel. In reality, research suggests that the optimal choice for time constants depends on several factors, and each hearing aid manufacturer may use different values (Moore, 2008a, 2008b).

5.6 Conclusion

Although there is growing evidence that spatiotemporal coding is important (Carney, 1994; Cedolin and Delgutte, 2007, 2010; Heinz, 2007; Larsen et al., 2008; Loeb et al., 1983; Shamma, 1985a), there has been no research into the effects of hearing aids on spatiotemporal coding of speech. Basic filter theory tells us that the group delay of a filter generally decreases as the transition bandwidth increases, so we should expect less delay in an impaired auditory system and perhaps even less with subsequent amplification. However, this had not been measured before.

We have shown that spatiotemporal coding (as measured by cross-CF delay) is indeed degraded with noise-induced hearing impairment, but is not improved with amplification. In fact, spatiotemporal pattern correction has not been a design goal for

most hearing aid research. However, a few researchers have attempted to design a hearing aid algorithm to do just this, and this is the topic of the next chapter.

CHAPTER 6. LIMITATIONS OF SPATIOTEMPORAL PATTERN CORRECTION

6.1 Introduction

According to basic filter theory, broad auditory filters (e.g., as a result of hearing impairment) are expected to have less associated delay than comparable narrow auditory filters. As illustrated in Figure 6.1, a broad filter generally has a shallower phase response and less delay than a narrow filter. Hearing impairment has therefore been expected to affect the spatiotemporal patterns by decreasing the relative delay

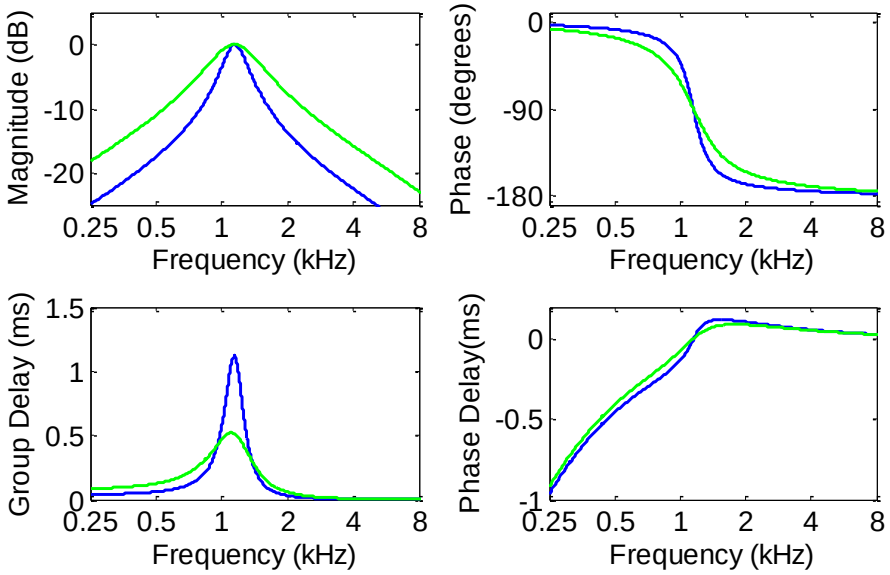

Figure 6.1 Example filter responses
Top-left: magnitude responses of a narrow (blue) and broad (green) filter; Top-right: phase responses of the same filters; Bottom-left: group delays; Bottom-right: phase delays

across CF. Empirical data presented in Chapters 3-4 confirm that cross-CF delay is reduced with impairment, and data presented in Chapter 5 show that hearing aids do not restore the cross-CF delay. The logical next step is to design a hearing aid that intentionally restores these delays. It has been hypothesized that by delaying specific frequencies more than others in the acoustic signal, the resulting neural patterns can be restored to near-normal.

Carney and colleagues developed an algorithm that was designed to introduce delay into the auditory signal where the phase was predicted to be abnormal (Calandruccio et al., 2007; Carney, 2008; Shi et al., 2006). The algorithm uses two parallel signal-processing paths to estimate the decrease in delay due to SNHL and to then add a compensatory frequency-dependent delay to the auditory signal prior to presentation to the ear (as shown in Figure 6.2). In the control path, an auditory model is used to estimate the difference in group delay introduced by healthy nonlinear filters and broader impaired filters. In the main path, the signal is decomposed into frequency channels, the missing delay is added, and the channels are re-combined after passing through a synthesis filterbank.

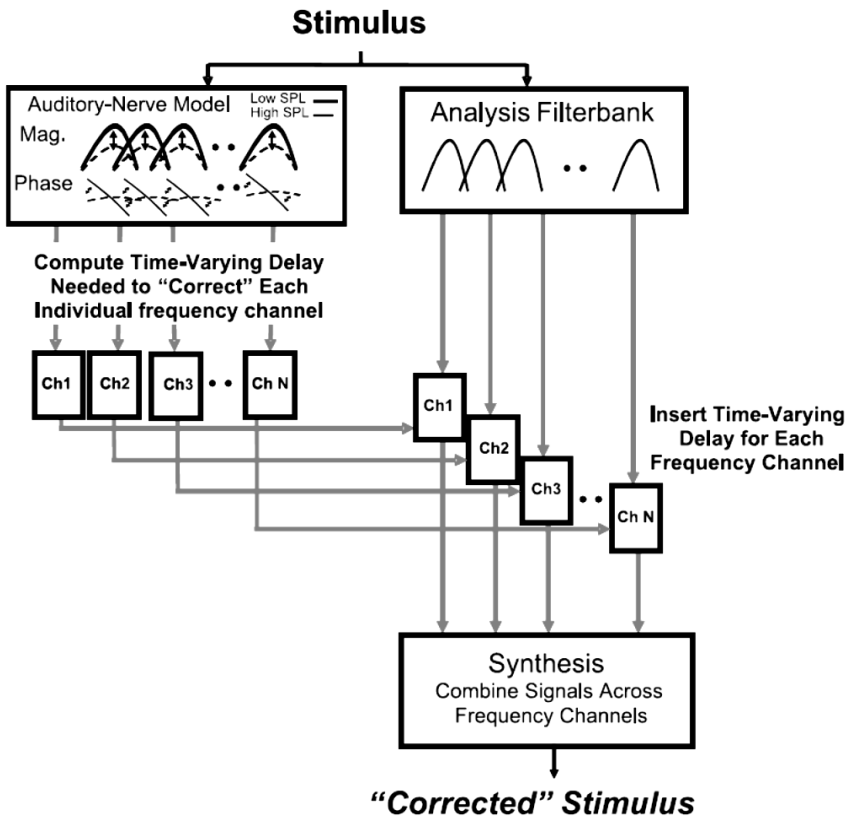

Figure 6.2 Schematic diagram of spatiotemporal pattern correction system The amount of compensatory delay is calculated in the control pathways (left) and applied within the analysis/synthesis filterbank (right) [reproduced, with permission, from Shi et al (2006)]

Although the fundamental idea seems reasonable and appeared promising at first (Shi et al., 2006), the processing strategy did not significantly improve speech perception (Calandruccio et al., 2007). The algorithm was designed to introduce an integer number of samples worth of delay in each analysis band, but this could potentially result in undesired artifacts like comb-filtering (where two analysis filters overlap) and abnormal phase transitions at the edges of a band. Unfortunately, the authors of these papers did not analyze the ability of their approach to restore normal spatiotemporal coding in the auditory nerve, either experimentally or with an AN model.

The goal of the present study was to fill this gap by quantifying the ability to modify the spatiotemporal coding of a stimulus with frequency-dependent delay.

6.2 Methods

Carney and colleagues (Calandruccio et al., 2007; Carney, 2008; Shi et al., 2006) used a filter bank to introduce frequency-dependent delays within the auditory signal, but the frequency resolution was restricted by the filter bank design. In fact, the phase transitions introduced by the Carney algorithm would be at the edge of each band, not the center. An alternative approach that would introduce a specific phase offset precisely at a particular frequency would be to design an all-pass filter. An all-pass filter applies equal gain at all frequencies, but the phase at each frequency can be controlled. Deshmukh et al (2007) used all-pass filters of this type to detect the harmonics of vowels; however, the goal of their work was not to modify the phase of the auditory signal.

Cho et al (1989) designed an adaptive notch filter that provides a foundation for an adaptive all-pass filter. The benefit of this design is that it can control a narrow frequency range and can be used with a variety of adaptation algorithms. (However, we will only use a static implementation here.) The notch filter is implemented with a lattice structure and each second-order section has the transfer function

$$H(z) = \frac{1 + k_0(1 + k_1)z^{-1} + k_1 z^{-2}}{1 + a_0(1 + a_1)z^{-1} + a_1 z^{-2}} \qquad \text{Equation 6}$$

where $a_1 = \alpha k_1$ and $a_0 = k_0$ (assuming α is close to a value of 1). The value of k_0 determines the frequency of the notch such that $\omega = \cos^{-1}(-k_0)$, k_1 determines the bandwidth, and α determines the depth of the notch.

Equation 6 can be generalized to the form:

$$H(z) = \frac{r_0 + r_1(1 + r_0 r_2)z^{-1} + r_2 z^{-2}}{v_2 + v_1(1 + v_0 v_2)z^{-1} + v_0 z^{-2}} \quad \text{Equation 7}$$

By building on this notch filter, we can modify it to have unity gain and a phase transition at a specified frequency. An all-pass filter with a non-zero phase response has the form

$$H(z) = \frac{B_0 + B_1 z^{-1} + B_2 z^{-2}}{A_2 + A_1 z^{-1} + A_0 z^{-2}} \quad \text{Equation 8}$$

where $A_i = B_i$. To follow convention, we can normalize the coefficients such that $A_2 = 1$. Therefore, $B_2 = 1$.

It can be shown that the maximum group delay (defined by the slope of the phase response) of a 1st order all-pass filter is defined by the equation

$$D_{max} = 1 - \frac{2}{1 - \frac{1}{p}} \quad \text{Equation 9}$$

where p is the filter pole[2] and D_{max} is given in samples. Solving for p, we get

[2] The pole of a 1st order filter, defined by the transfer function $H(z) = \frac{B_0 + B_1 z^{-1}}{A_1 + A_0 z^{-1}}$, is the value of z such that the denominator equals zero. Similarly, the (complex) poles of a second order filter are the values of z such that the denominator of the transfer function equals zero.

$$p = \frac{D_{max} - 1}{D_{max} + 1} \qquad \text{Equation 10}$$

For the 2nd order filter in Equation 7, we can set

$$r_0 = \frac{1}{1 - \frac{2}{1 - \frac{D_{max}}{2}}} = \frac{D_{max} - 2}{D_{max} + 2} \qquad \text{Equation 11}$$

A second-order filter stage can thus be designed to apply equal gain (e.g., 0dB) to all frequencies, but to alter the phase such that some frequencies are delayed more than others. The transfer function for such a filter is given in Equation 7, where the coefficients are set such that:

$$v_0 = r_0 = \frac{D_{max} - 2}{D_{max} + 2}$$

$$v_1 = r_1 = -\cos\frac{2\pi f}{F_S} \qquad \text{Equation 12}$$

$$v_2 = r_2 = 1$$

where f is the center frequency (in Hz) of the phase transition, and Fs is the sample rate. The center frequency, phase and group delay response can be arbitrarily controlled, as illustrated in Figure 6.3.

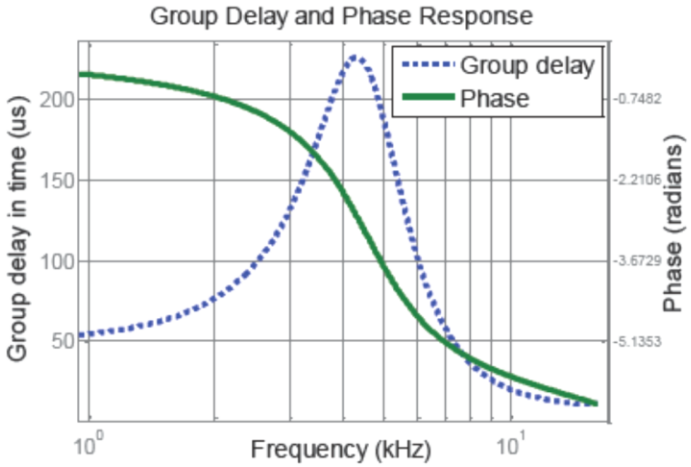

Figure 6.3 Example phase and group delay response of an all-pass filter

It is important to note the difference between phase delay and group delay. Phase delay is defined as $\frac{\varphi(\omega)}{\omega}$, where ω is the frequency (in radians/second) and $\varphi(\omega)$ is the phase (in radians) at that frequency. The phase delay is therefore the time delay (in seconds) for each Fourier component frequency. Group delay is defined as $-\frac{d}{d\omega}\varphi(\omega)$, or the negative derivative of the phase. For a linear phase system, the group delay is equivalent to the phase delay. For a nonlinear phase system (such as auditory filters), the group delay is often interpreted as the delay of the *envelope*. However, this interpretation is limited to a narrow range of frequencies over which the phase is approximately linear (Smith, 2007, p. 163). It is also interesting to note that when the phase response increases with frequency, the result is a negative group delay, which is certainly counterintuitive. Therefore, it can be difficult to interpret group delay, and phase delay may serve as a more accurate description of delay as a function of frequency.

We filtered the vowel /ɛ/ (in quiet, 65 dB SPL) with a series of second-order all pass filter stages, cascaded to increase the phase delay at 1 kHz, as shown in Figure 6.4, which was the frequency at which F2 was placed. These filtered stimuli were used as input (at 65dB SPL) to an auditory nerve model (Zilany et al., 2009) of normal hearing.

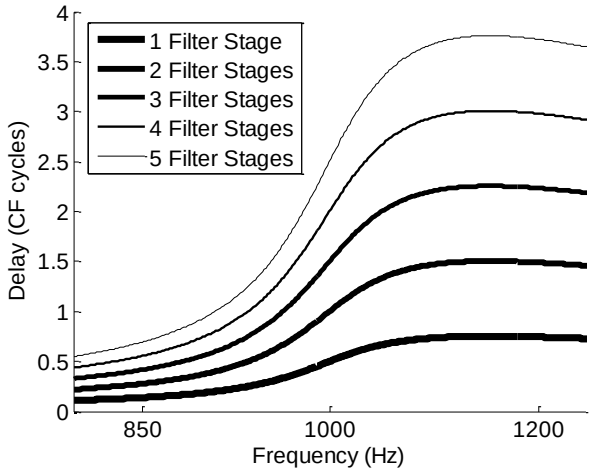

Figure 6.4 Phase delays for a series of all-pass filters

The resulting auditory nerve spikes were analyzed as in the previous chapters to quantify spatiotemporal coding. Specifically, we calculated the characteristic delay between AN fiber CFs surrounding the second formant of the vowel (which was placed at 1 kHz for this modeling study). We expected to see that, as the acoustic phase delay was increased by cascading additional all-pass filter stages, the characteristic delay would also increase.

6.3 Results

We used the reverse correlation (revcor) function (de Boer, 1978; Carney and Yin, 1988; Eggermont, 1983) to estimate the transfer function of the neural spikes relative to the unfiltered vowel, and the resulting spectra for an AN fiber with 1 kHz CF is shown in

Figure 6.5. Thinner lines indicate responses to stimuli with more all-pass filter stages. Here, we can see that this fiber is responding to energy near 1 kHz (as indicated by the magnitude peak near 1 kHz), and the revcor phase near 1 kHz is altered to a greater degree (as expected) with increasing numbers of all-pass filter stages.

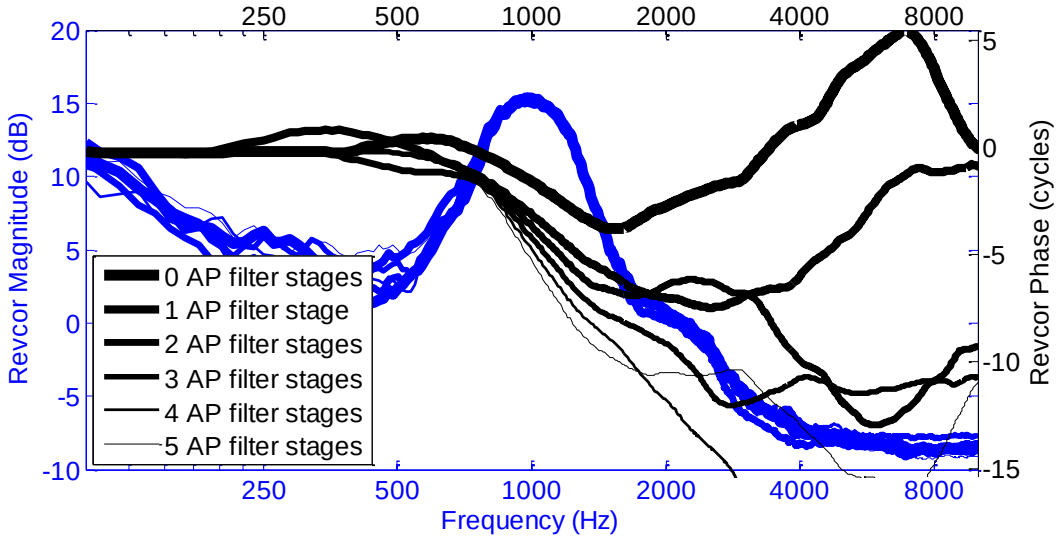

Figure 6.5 Revcor magnitude and phase for 1 kHz CF

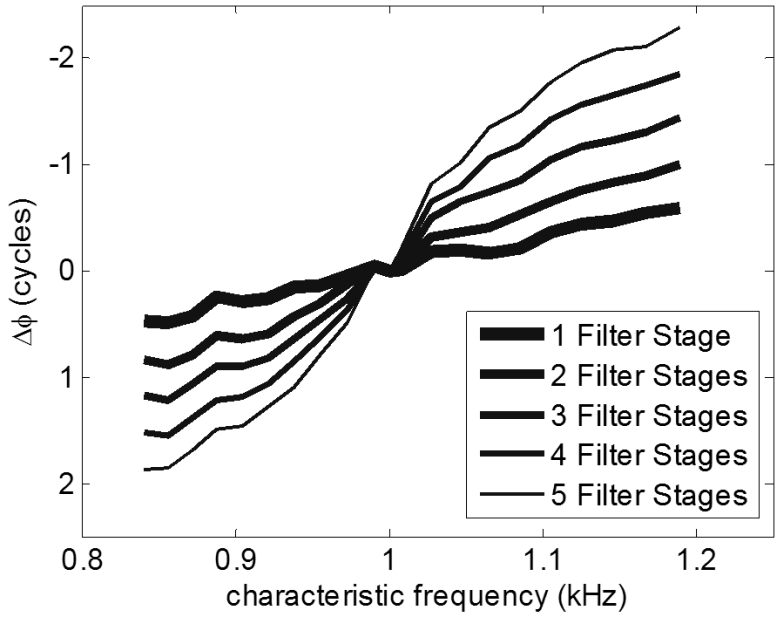

Figure 6.6 Reverse-correlation phase at CF (relative to phase at 1 kHz CF, no filter)

We evaluated the revcor phase at CF for 10 CFs covering a range of 1 kHz ± 0.25 octaves. Figure 6.6 shows the phase of the reverse-correlation function at each fiber's CF, relative to the phase measured from a 1 kHz CF with unfiltered input. We can therefore confirm that the all-pass filters are modifying the stimulus as expected, and that this is indeed represented in the neural signals.

However, any spatiotemporal decoding mechanism in the brain will depend on common excitation between two (or more) AN fibers, so a more relevant analysis is to evaluate the revcor phase at a single frequency for multiple CFs. If we look at the phase across CFs but at a common frequency (1 kHz), we see that there is a constant phase offset for each all-pass filter, but this does not vary across CF (as shown in Figure 6.7).

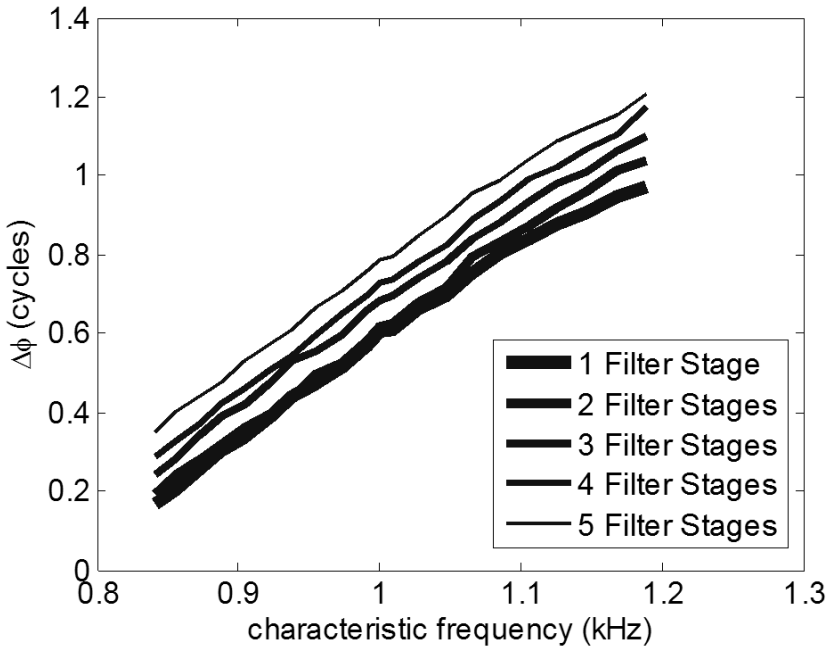

Figure 6.7 Reverse-correlation phase at 1 kHz (relative to no filter)

The characteristic delays relative to the vowel formant frequency are shown in Figure 6.8. We did not see any effect of the filters on spatiotemporal coding, as

assumed in the spatiotemporal correction algorithms. In other words, increasing the phase delay in the acoustic signal did not affect the relative timing across AN fibers with nearby CFs (Figure 6.7), although it did affect the phase response of individual fibers (Figure 6.6).

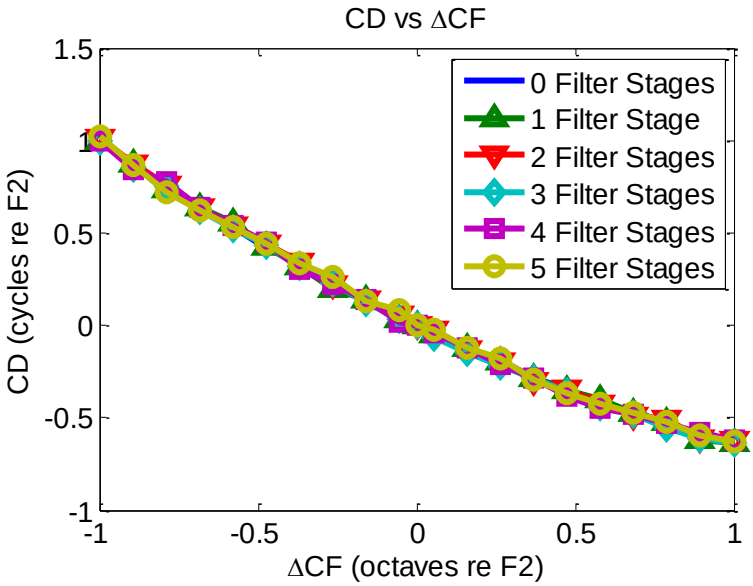

Figure 6.8 Characteristic delay relative to AN fiber with CF at F2 (1 kHz)

For completeness, we also implemented a simple model of a monaural across-CF coincidence detector neuron (Krips and Furst, 2009; Wang and Delgutte, 2012). This type of neuron has been shown to be sensitive to across-CF differences in phase that are relevant for intensity discrimination (Heinz et al., 2001a) and tone detection in noise (Carney et al., 2002). Furthermore, there is evidence that some cell types in the ventral cochlear nucleus act in a manner consistent with monaural across-CF coincidence detection (Carney, 1990; Wang and Delgutte, 2012). This modeled neuron had 10 auditory nerve inputs with CFs spanning a range of 0.5 octaves and fires whenever at

least 2 input fibers fire within a short time window of each other. If the all-pass filter were able to alter spatiotemporal coding systematically, then we would expect to see a systematic effect in the coincidence neuron firing rate as the number of all-pass filter stages were increased. However, we were unable to find a window size over which the firing rate of a coincidence detector would systematically vary as the phase of the stimulus is varied (as illustrated in Figure 6.9). (Note that this model of a coincidence detector did not include a refractory period, so the rates seen with a wide window are unrealistically high. However, the relationship between firing rate and all-pass filtering is not expected to depend on refractory periods.)

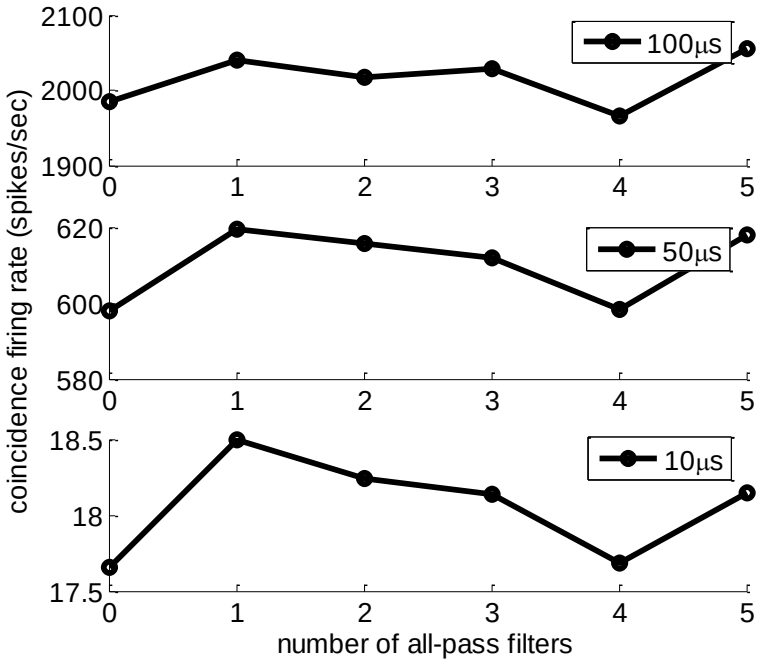

Figure 6.9 Firing rate of a model coincidence detector
(temporal coincidence windows of 100, 50 and 10μs)

Spatiotemporal pattern correction is expected to work if neighboring AN fibers are approximately independent (e.g., non-overlapping receptive fields). However, at levels

well above threshold, this is unlikely the case. Therefore, we might only expect spatiotemporal pattern correction to work when the receptive fields are quite narrow, such as with normal hearing and low intensity stimuli. This is consistent with a recent study (Zeyl and Bruce, In Press), which found that the algorithm by Carney and colleagues (Calandruccio et al., 2007; Carney, 2008; Shi et al., 2006) is most beneficial for low intensity stimuli. However, when we repeated our experiment at 10dB SPL, we did not see any evidence that all-pass filters affect spatiotemporal coding as expected. As expected (and as shown in Figure 6.10), the slope of the delay function is increased at low stimulus levels, but we did not see any systematic progression of the slope with increasing phase delay. Additional delay should also decrease the coincidence across CF, and although the predicted firing rate of a coincidence detector neuron is drastically reduced when we reduce the stimulus level, we do not see this effect from the all-pass filters (see Figure 6.11).

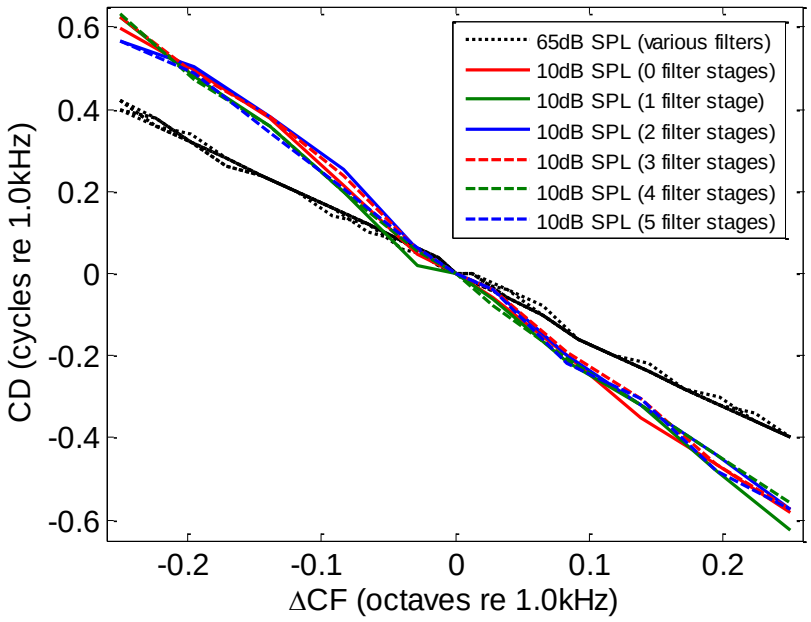

Figure 6.10 Characteristic delays, for stimuli at 10dB SPL
(data for 65dB stimuli are dotted lines, for reference)

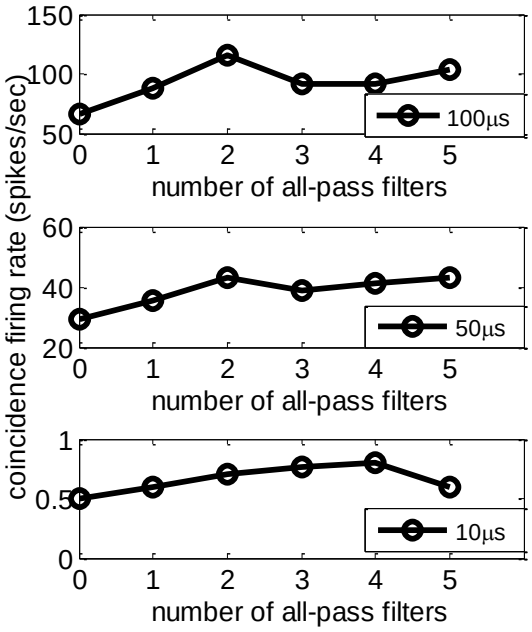

Figure 6.11 Firing rate of a model coincidence detector (stimuli at 10dB SPL)
(temporal windows of 100, 50 and 10μs)

6.4 Discussion

These results show that applying a frequency-dependent delay to a vowel does not affect the spatiotemporal coding of the original stimulus as expected (Shi et al., 2006). We confirmed that the neural responses (as measured in the revcor phase) were indeed altered by the all-pass filters, but the characteristic delays were not affected because each fiber is affected the same by the change in acoustic phase.

Although Wang & Delgutte (2012) found that coincidence detecting cells were sensitive to the phase transition sharpness of Huffman stimuli (impulse responses of all-pass filters), they found that the total firing rate is not a good measure of the coincidence detector's sensitivity to phase. They used metrics like the peak width and normalized duration to quantify the responses to click-like stimuli, but these metrics do not apply to the steady-state vowel stimulus used here.

6.4.1 Potential Binaural Artifacts from Spatiotemporal Pattern Correction

It is important to keep in mind that adding frequency-dependent delay will distort the auditory signal. Interaural timing differences (ITDs) are known to be important for localization, and Joris has suggested that the interaction between the acoustic signals and cochlear timing disparities are important for decoding ITDs (Joris, 2003; Joris et al., 2006b). Adding phase offsets to the acoustic signal may impact localization abilities. In fact, we also know that binaural phase differences can be perceived as pitch (Cramer and Huggins, 1958). Therefore, any algorithm that attempts to correct the phase for monaural speech stimuli should also be checked for binaural artifacts.

6.4.2 Use of Computational Models

In contrast to previous chapters, this chapter focused on the use of a computational model. This allowed us to analyze the expected neural patterns in detail before moving on to animal experiments, thus reducing the total number of required animal experiments. In fact, because the modeling work identified a fundamental limitation in this approach, we chose not to pursue these ideas with animals. Continued development of computational models may further reduce the number of animals needed in future experiments.

6.4.3 Conclusion

The spatiotemporal patterns that have been hypothesized to be perceptually relevant are the relative timing across CFs responding to the same stimulus feature (i.e., same frequency). Using a hearing aid algorithm to change the phase of that single frequency can have only one effect on all fibers and, although each fiber is affected by the acoustic phase change (Figure 6.6), the effect is the same for all CFs responding to that same frequency (Figure 6.7) and thus the relative timing across fibers is unchanged (Figure 6.8). This represents a fundamental limitation to the approach of spatiotemporal pattern correction. An ideal correction algorithm would alter the time delay (phase) for one CF more than for another CF, but this is not possible. This concept is similar to the inability to separately control cochlear regions with gain adjustment (Giguère and Smoorenburg, 1999; Heinz, 2010).

CHAPTER 7. DISCUSSION

This dissertation has used neurophysiology and computational modeling to evaluate the performance of hearing aids. By using both computational models and physiological experiments in animals, we have been able to explore several aspects of neural coding following hearing aid amplification, which to date has been a largely unexplored area. We have shown that within-fiber temporal coding can be improved with appropriate amplification, but that optimal gain may depend significantly on the ratio of inner and outer hair cell dysfunction for a given individual. We have also shown that sensorineural hearing loss produces degraded spatiotemporal coding, but that amplification from commonly used hearing aid algorithms do not appear to improve spatiotemporal coding. Unfortunately, we have also shown that proposed ideas for spatiotemporal pattern correction appear to be ill-fated, at least as proposed to date. Our results suggest that applying a frequency-dependent delay may not actually affect the spatiotemporal coding as expected.

The modeling studies presented here allowed us to study a variety a topics related to neural coding without the time & resource-consuming process of conducting a series of animal experiments (which typically last 24-36 hours with chinchillas). Of course, the model has not been designed/verified for every situation, and investigating new areas of

research can easily exceed the validated boundaries of any computational model. This situation is when it is helpful to have the resources and expertise to conduct animal experiments. The goal of incorporating physiological knowledge in hearing aid design will require a well-planned and coordinated combination of physiological modeling and experiments.

Computational and physiological approaches let us evaluate very detailed information about the sensory system, so they are tremendously valuable, but the ultimate goal of this knowledge is to improve human behavior (for example speech intelligibility). Therefore, it is important to learn from each of these three areas and combine domain-specific knowledge into one cohesive (and even synergetic) research track. The topic of this dissertation is one example where this is particularly true. Temporal and spatiotemporal coding have been receiving a lot of attention recently in the psychoacoustics literature. Spatiotemporal coding has been hypothesized to be very important perceptually, but there is still much that we do not understand about how temporal coding is used in the auditory system. By investigating the effects of SNHL and hearing aid amplification in spatiotemporal coding using a combined computational and experimental approach, we now have a better understanding of some of the critical factors that currently limit the ability of hearing aids to improve spatiotemporal coding, and perhaps perception. This knowledge will be useful in future efforts to develop novel strategies to improve human speech perception.

7.1 Limitations

7.1.1 Animal Species

One limitation in studying speech coding in animals is that species differences in basilar-membrane length and frequency range of hearing make direct comparison to speech coding in humans difficult (Kiefte et al., 2002; Recio et al., 2002). However, this limitation is not critical for the studies presented here because we focus on quantitative comparisons between temporal coding in normal and impaired cases within the same species. Any comparisons across species of the effects of SNHL must be made with full consideration of relevant species differences. Of particular relevance here is the recent finding that the chinchilla cochlea may be "more apical" than the human cochlea (Shera et al., 2007, 2008). The CF transition between basal and apical cochlear regions, estimated from OAEs, was 4 kHz for chinchilla and 1 kHz for humans. Although this discrepancy must be considered in quantitatively relating chinchilla responses to humans, it actually provides a benefit here because it makes it easier to study the effects of SNHL on the "apical" region of the cochlea, which is most important for the low frequencies in speech.

7.1.2 Cochlear Scaling Invariance

The STMP assumes cochlear scaling invariance, which is reasonable over our limited frequency shifts (≤±0.5 octave shifts). However, some properties are not scaling invariant, including roll-off in phase locking, refractoriness, adaptation, and increases in Q10 with CF. These effects are predicted to be negligible over our ±0.5 octave range

(Larsen et al., 2008); however, all results have been interpreted with these limitations in mind.

7.1.3 Model-Based Fitting Strategies

Any hearing aid fitting strategy that is based on minimizing the difference between a normal and impaired system must accurately model both systems. To be applicable to any particular patient, the model of hearing impairment must accurately reflect that person's impairments. Unfortunately, we do not currently have good noninvasive ways to measure many of the physiological aspects of hearing impairment; however people are actively working on this important issue (Lopez-Poveda and Barrios, 2013; Lopez-Poveda and Johannesen, 2012; Moore and Glasberg, 2004; Moore, 2004).

7.2 Relation to other research

7.2.1 Speech Coding

Evidence suggests that vowels are important for both sentence intelligibility (Cole et al., 1996; Kewley-Port et al., 2007) and talker identification (Owren and Cardillo, 2006). Furthermore, we know that pitch differences are a key feature used to group/segregate sounds (Bregman, 1990) and to identify concurrent vowels (Summers and Leek, 1998). Therefore, vowels are especially important as we try to understand speech coding in the presence of background sounds. The first two vowel formants are most important for speech recognition (Pols et al., 1969; Sakayori et al., 2002), so that is what we have focused on here. However, other features (such as pitch) are certainly also important for perception. We have extended previous research on vowel coding (Heinz, 2007; Miller et al., 1997; Sachs and Young, 1979; Sachs et al., 1983; Shamma, 1985a; Young

and Sachs, 1979), and shown that hearing aids are currently unable to improve the spatiotemporal coding of vowels.

7.2.2 Neural Degeneration

Recent research has shown that auditory nerves and their synaptic connections to hair cells are often impaired, even when thresholds look normal (Furman et al., 2013; Kujawa and Liberman, 2009; Lin et al., 2011; Maison et al., 2013; Makary et al., 2011; Sergeyenko et al., 2013). If information is missing among some proportion of the auditory nerve fibers, we should expect perception to be affected in some way. In particular, listening in noise is likely to be degraded if the brain relies on any type of spatial summation. For example, any mechanism which decodes spatiotemporal cues would receive fewer inputs. This could affect any of the percepts thought to be associated with spatiotemporal coding, including speech perception, pitch, loudness, localization, and masking (Carney, 1994; Cedolin and Delgutte, 2007, 2010; Heinz et al., 2001a; Joris et al., 2006b; Loeb et al., 1983; Shamma, 1985a). Future research to investigate the impact of neural degeneration on spatiotemporal coding and these percepts would likely be worthwhile.

7.3 Opportunities for Future Research

Chapter 2 presented some results indicating that optimal gain may depend on the relative health of inner and outer hair cells. However, the actual gain applied with a hearing aid is dependent on a combination of factors: the input level, the target gain for that input level, and the time constants for adapting the gain. The time constants can

have a significant impact on how often the target gain is reached for any particular phoneme. We have started an analysis to explore the effects of slow versus fast time constants on neural coding, but this is a very computationally expensive procedure. Although we can model many auditory nerve fibers in parallel, the state of the hearing aid at any particular time depends on the previous states. Therefore, to calculate the optimal gain over any meaningful length of speech, the optimal settings must be calculated for each small time segment before moving on to the next. This is a very computationally expensive process, but new advances in parallel processing may allow this important issue to be addressed in future studies.

It would also be interesting to explore the relationship between spatiotemporal coding and listening in cocktail-party types of situations. Unfortunately, we were unable to alter spatiotemporal coding as we hoped, but perhaps other types of stimuli could be used to explore the relationship between spatiotemporal coding and speech segregation. For example, it would be interesting to study the interaction between electronic filters (e.g., applied to the acoustic signal) and auditory filters. An approach similar to active noise cancellation could potentially be used to reduce undesired cochlear activity. If filters could be designed to alter the receptive field of auditory nerve fibers, such filters may be beneficial for speech (and perhaps even music) perception.

REFERENCES

REFERENCES

ANSI (**2003**). “ANSI S3 22-2003, Specification of hearing aid characteristics”.

Backus, B. C., and Guinan Jr., J. J. (**2007**). “Measurement of the distribution of medial olivocochlear acoustic reflex strengths across normal-hearing individuals via otoacoustic emissions,” Journal of the Association for Research in Otolaryngology, **8**, 484–496.

Baer, T., Moore, B. C., and Gatehouse, S. (**1993**). “Spectral contrast enhancement of speech in noise for listeners with sensorineural hearing impairment: effects on intelligibility, quality, and response times,” The Journal of Rehabilitation Research and Development, **30**, 49–72.

Békésy, G., and Bekesy, G. (**1952**). “Direct observation of the vibrations of the cochlear partition under a microscope,” Acta Oto-Laryngologica, **42**, 197–201.

Bentler, R., and Chiou, L. K. (**2006**). “Digital noise reduction: an overview,” Trends in Amplification, **10**, 67–82.

Biondi, E. (**1978**). “Auditory processing of speech and its implications with respect to prosthetic rehabilitation The bioengineering viewpoint,” Audiology, **17**, 43–50.

De Boer, E. (**1978**). “On cochlear encoding: Potentialities and limitations of the reverse-correlation technique,” Journal of the Acoustical Society of America, **63**, 115–135.

Bondy, J., Becker, S., Bruce, I., Trainor, L., and Haykin, S. (**2004**). “A novel signal-processing strategy for hearing-aid design: neurocompensation,” Signal Processing, **84**, 1239–1253.

Bondy, J., Bruce, I., Becker, S., and Haykin, S. (**2003**). “Predicting speech intelligibility from a population of neurons,” In S. Thrun, L. Saul, and B. Schölkopf (Eds.), NIPS 2003 Conference Proceedings: Advances in Neural Information Processing Systems, MIT Press, Cambridge, MA, pp. 1409–1416.

Bregman, A. S. (**1990**). *Auditory scene analysis : the perceptual organization of sound*, MIT Press, Cambridge, Mass.

Brownell, W. E., Bader, C. R., Bertrand, D., and de Ribaupierre, Y. (**1985**). "Evoked mechanical responses of isolated cochlear outer hair cells," Science, **227**, 194–196.

Bruce, I. (**2010**). "The Problems of Inner-Hair-Cell Impairment for Hearing Aid Gain Prescriptions," International Hearing Aid Research Conference, Lake Tahoe, CA.

Bruce, I. C. (**2004**). "Physiological assessment of contrast-enhancing frequency shaping and multiband compression in hearing aids," Physiol Meas, **25**, 945–956.

Bruce, I. C. (**2006**). "Hearing Aids," Encyclopedia of Biomaterials and Biomedical Engineering, Taylor & Francis, pp. 1–8.

Bruce, I. C., Dinath, F., and Zeyl, T. J. (**2007**). "Insights into optimal phonemic compression from a computational model of the auditory periphery," International Symposium on Auditory and Audiological Research, Helsingor, Denmark.

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," Journal of the Acoustical Society of America, **113**, 369–388.

Byrne, D., Dillon, H., Ching, T., Katsch, R., and Keidser, G. (**2001**). "NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures," Journal of the American Academy of Audiology, **12**, 37–51.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., et al. (**1994**). "An International Comparison of Long-Term Average Speech Spectra," Journal of the Acoustical Society of America, **96**, 2108–2120.

Byrne, D., Parkinson, A., and Newall, P. (**1990**). "Hearing aid gain and frequency response requirements for the severely/profoundly hearing impaired," Ear and Hearing, **11**, 40–49.

Calandruccio, L., Doherty, K. A., Carney, L. H., and Kikkeri, H. N. (**2007**). "Perception of temporally processed speech by listeners with hearing impairment," Ear and Hearing, **28**, 512–523.

Cariani, P. A., and Delgutte, B. (**1996**). "Neural correlates of the pitch of complex tones I Pitch and pitch salience," Journal of Neurophysiology, **76**, 1698–1716.

Carney, L. H. (**1990**). "Sensitivities of cells in anteroventral cochlear nucleus of cat to spatiotemporal discharge patterns across primary afferents," Journal of Neurophysiology, **64**, 437–456.

Carney, L. H. (**1992**). “Modelling the sensitivity of cells in the anteroventral cochlear nucleus to spatiotemporal discharge patterns,” Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences, **336**, 403–406.

Carney, L. H. (**1994**). “Spatiotemporal encoding of sound level: models for normal encoding and recruitment of loudness,” Hearing Research, **76**, 31–44.

Carney, L. H., and Friedman, M. (**1998**). “Spatiotemporal tuning of low-frequency cells in the anteroventral cochlear nucleus,” Journal of Neuroscience, **18**, 1096–1104.

Carney, L. H., Heinz, M. G., Evilsizer, M. E., Gilkeyz, R. H., and Colburn, H. S. (**2002**). “Auditory Phase Opponency: A Temporal Model for Masked Detection at Low Frequencies,” Acta Acustica United with Acustica, **88**, 334–347(14).

Carney, L. H. L. (**2008**). *Method for correcting sound for the hearing-impaired*, US Patent 7,428,313,United States patent.

Carney, L. H., and Yin, T. C. (**1988**). “Temporal coding of resonances by low-frequency auditory nerve fibers: single-fiber responses and a population model,” Journal of Neurophysiology, **60**, 1653–1677.

Cedolin, L., and Delgutte, B. (**2005**). “Pitch of complex tones: rate-place and interspike interval representations in the auditory nerve,” Journal of Neurophysiology, **94**, 347–362.

Cedolin, L., and Delgutte, B. (**2007**). “Spatio-temporal representation of the pitch of complex tones in the auditory nerve,” In B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Eds.), Springer-Verlag, Berlin, pp. 61–70.

Cedolin, L., and Delgutte, B. (**2010**). “Spatiotemporal representation of the pitch of harmonic complex tones in the auditory nerve,” Journal of Neuroscience, **30**, 12712–12724.

Chen, Z., Becker, S., Bondy, J., Bruce, I. C., and Haykin, S. (**2005**). “A novel model-based hearing compensation design using a gradient-free optimization method,” Neural Computation, **17**, 2648–2671.

Cherry, E. C. (**1953**). “Some Experiments on the Recognition of Speech, with One and with Two Ears,” Journal of the Acoustical Society of America, **25**, 975–979.

Chintanpalli, A., and Heinz, M. G. (**2007**). “Effect of auditory-nerve response variability on estimates of tuning curves,” Journal of the Acoustical Society of America, **122**, 203–209.

Cho, N. I. (**1989**). “Adaptive line enhancement by using an IIR lattice notch filter,” (C.-H. Choi and S. U. Lee, Eds.) IEEE Transactions on Acoustics, Speech, and Signal Processing, **37**, 585–589.

Cole, R. R. A., Mak, B., Fanty, M., Bailey, T., and Yan, Y. (**1996**). “The contribution of consonants versus vowels to word recognition in fluent speech,” 1996 IEEE International Conference on Acoustics, Speech, and Signal Processing Conference Proceedings, IEEE, 853–856.

Cooper, N. P., and Guinan Jr., J. J. (**2003**). “Separate mechanical processes underlie fast and slow effects of medial olivocochlear efferent activity,” Journal of Physiology, **548**, 307–312.

Cramer, E., and Huggins, W. (**1958**). “Creation of pitch through binaural interaction,” Journal of the Acoustical Society of America, **30**, 413–417.

Darrow, K. N., Maison, S. F., and Liberman, M. C. (**2006**). “Cochlear efferent feedback balances interaural sensitivity,” Nature Neuroscience, **9**, 1474–1476.

Delgutte, B., and Kiang, N. Y. (**1984**). “Speech coding in the auditory nerve: V Vowels in background noise,” Journal of the Acoustical Society of America, **75**, 908–918.

Deng, L., and Geisler, C. D. (**1987**). “A composite auditory model for processing speech sounds,” Journal of the Acoustical Society of America, **82**, 2001–2012.

Deshmukh, O. D., Espy-Wilson, C. Y., and Carney, L. H. (**2007**). “Speech enhancement using the modified phase-opponency model,” Journal of the Acoustical Society of America, **121**, 3886–3898.

DiGiovanni, J. J., Nelson, P. B., and Schlauch, R. S. (**2005**). “A psychophysical evaluation of spectral enhancement,” Journal of Speech, Language, and Hearing Research, **48**, 1121–1135.

Dillon, H. (**2001**). *Hearing Aids*, Thieme: Boomerang Press, New York, Sydney.

Duquesnoy, A. J. (**1983**). “Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons,” Journal of the Acoustical Society of America, **74**, 739–743.

Edwards, B. (**2007**). “The future of hearing aid technology,” Trends in Amplification, **11**, 31–45.

Eggermont, J. (**1983**). “Reverse-correlation methods in auditory research,” Quarterly Reviews of Biophysics, **16**, 341–414.

Franck, B. A., van Kreveld-Bos, C. S., Dreschler, W. A., and Verschuure, H. (**1999**). “Evaluation of spectral enhancement in hearing aids, combined with phonemic compression,” Journal of the Acoustical Society of America, **106**, 1452–1464.

Furman, A., Kujawa, S. G., and Liberman, M. C. (**2013**). “Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates,” Journal of Neurophysiology, **110**, 577–586.

Gatehouse, S., Naylor, G., and Elberling, C. (**2003**). “Benefits from hearing aids in relation to the interaction between the user and the environment,” International Journal of Audiology, **42 Suppl 1**, S77–S85.

Geisler, C. D. (**1989**). “The responses of models of ‘high-spontaneous’ auditory-nerve fibers in a damaged cochlea to speech syllables in noise,” Journal of the Acoustical Society of America, **86**, 2192–2205.

Geisler, C. D., and Gamble, T. (**1989**). “Responses of ‘high-spontaneous’ auditory-nerve fibers to consonant-vowel syllables in noise,” Journal of the Acoustical Society of America, **85**, 1639–1652.

Ghitza, O. (**2001**). “On the upper cutoff frequency of the auditory critical-band envelope detectors in the context of speech perception,” Journal of the Acoustical Society of America, **110**, 1628–1640.

Giguère, C., and Smoorenburg, G. (**1999**). “Computational modeling of outer hair cell damage: Implications for hearing aid signal processing,” Psychophysics, Physiology and Models of Hearing, World Scientific, River Edge, NJ, pp. 155–164.

Gilbert, G., and Lorenzi, C. (**2006**). “The ability of listeners to use recovered envelope cues from speech fine structure,” Journal of the Acoustical Society of America, **119**, 2438–2444.

Goldberg, J. M., and Brown, P. B. (**1969**). “Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization,” Journal of Neurophysiology, **32**, 613–636.

Greenwood, D. D. (**1990**). “A cochlear frequency-position function for several species--29 years later,” Journal of the Acoustical Society of America, **87**, 2592–2605.

Groff, J. A., and Liberman, M. C. (**2003**). “Modulation of cochlear afferent response by the lateral olivocochlear system: activation via electrical stimulation of the inferior colliculus,” Journal of Neurophysiology, **90**, 3178–3200.

Guinan, J. J. (**2006**). “Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans,” Ear and Hearing, **27**, 589–607.

Harding, G., and Bohne, B. (**2007**). “Distribution of focal lesions in the chinchilla organ of Corti following exposure to a 4-kHz or a 0.5-kHz octave band of noise,” Hearing Research, **225**, 50–59.

Harding, G., and Bohne, B. (**2009**). “Relation of focal hair-cell lesions to noise-exposure parameters from a 4- or a 0.5-kHz octave band of noise,” Hearing Research, **254**, 54–63.

Harrison, R. V, and Evans, E. F. (**1979**). “Some aspects of temporal coding by single cochlear fibres from regions of cochlear hair cell degeneration in the guinea pig,” Archives of oto-rhino-laryngology, **224**, 71–78.

Haykin, S., Becker, S., Bruce, I., Bondy, J., Trainor, L., and Racine, R. J. (**2006**). *Binaural adaptive hearing aid*, US Patent 7149320, United States.

Heinz, M. G. (**2005**). “Spectral coding based on cross-frequency coincidence detection of auditory-nerve responses,” Association for Research in Otolaryngology Abstracts, New Orleans, Louisiana.

Heinz, M. G. (**2007**). “Spatiotemporal encoding of vowels in noise studied with the responses of individual auditory nerve fibers,” In B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Eds.), Hearing–From Sensory Processing to Perception, Springer Verlag, Heidelberg, pp. 107–115.

Heinz, M. G. (**2010**). “Computational Modeling of Sensorineural Hearing Loss,” In R. Meddis, E. A. Lopez-Poveda, A. N. Popper, and R. R. Fay (Eds.), Computational Models of the Auditory System, Springer-Verlag.

Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2001**). “Rate and timing cues associated with the cochlear amplifier: level discrimination based on monaural cross-frequency coincidence detection,” Journal of the Acoustical Society of America, **110**, 2065–2084.

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," Journal of the Association for Research in Otolaryngology, **6**, 91–105.

Heinz, M. G., and Swaminathan, J. (**2009**). "Quantifying envelope and fine-structure coding in auditory-nerve responses to chimaeric speech," Journal of the Association for Research in Otolaryngology, **10**, 407–423.

Heinz, M. G., and Young, E. D. (**2004**). "Response growth with sound level in auditory-nerve fibers after noise-induced hearing loss," Journal of Neurophysiology, **91**, 784–795.

Heinz, M. G., Zhang, X., Bruce, I. C., and Carney, L. H. (**2001**). "Auditory nerve model for predicting performance limits of normal and impaired listeners," Acoustics Research Letters Online, **2**, 91–96.

Heinz, M., Swaminathan, J., Boley, J., and Kale, S. (**2010**). "Across-Fiber Coding of Temporal Fine-Structure: Effects of Noise-Induced Hearing Loss on Auditory-Nerve Responses," In E. A. Lopez-Poveda, R. Meddis, and A. R. Palmer (Eds.), The Neurophysiological Bases of Auditory Perception, Springer, New York, pp. 621–630.

Henry, K. S., and Heinz, M. G. (**2012**). "Diminished temporal coding with sensorineural hearing loss emerges in background noise," Nature Neuroscience, **15**, 1362–1364.

Henry, K. S., Kale, S., Scheidt, R. E., and Heinz, M. G. (**2011**). "Auditory brainstem responses predict auditory nerve fiber thresholds and frequency selectivity in hearing impaired chinchillas," Hearing Research, **280**, 236–244.

Hienz, R. D., Stiles, P., and May, B. J. (**1998**). "Effects of bilateral olivocochlear lesions on vowel formant discrimination in cats," Hearing Research, **116**, 10–20.

Hopkins, K., and Moore, B. C. (**2007**). "Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information," Journal of the Acoustical Society of America, **122**, 1055–1068.

Hopkins, K., Moore, B. C., and Stone, M. A. (**2008**). "Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech," Journal of the Acoustical Society of America, **123**, 1140–1153.

Johnson, D. H. (**1980**). "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," Journal of the Acoustical Society of America, **68**, 1115–1122.

Joris, P. X. (**2003**). “Interaural time sensitivity dominated by cochlea-induced envelope patterns,” Journal of Neuroscience, **23**, 6345–6350.

Joris, P. X., Bergevin, C., Kalluri, R., Mc Laughlin, M., Michelet, P., van der Heijden, M., and Shera, C. A. (**2011**). “Frequency selectivity in Old-World monkeys corroborates sharp cochlear tuning in humans,” Proceedings of the National Academy of Sciences, **108**, 17516–20.

Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. (**1994**). “Enhancement of neural synchronization in the anteroventral cochlear nucleus I Responses to tones at the characteristic frequency,” Journal of Neurophysiology, **71**, 1022–1036.

Joris, P. X., Louage, D. H., Cardoen, L., and van der Heijden, M. (**2006a**). “Correlation index: a new metric to quantify temporal coding,” Hearing Research, **216-217**, 19–30.

Joris, P. X., de Sande, B. Van, Louage, D. H., and van der Heijden, M. (**2006b**). “Binaural and cochlear disparities,” Proceedings of the National Academy of Sciences, **103**, 12917–12922.

Joris, P. X., Smith, P. H., and Yin, T. C. (**1994**). “Enhancement of neural synchronization in the anteroventral cochlear nucleus II Responses in the tuning curve tail,” Journal of Neurophysiology, **71**, 1037–1051.

Kale, S., and Heinz, M. G. (**2010**). “Envelope coding in auditory nerve fibers following noise-induced hearing loss,” Journal of the Association for Research in Otolaryngology, **11**, 657–673.

Kates, J. (**1994**). “Speech enhancement based on a sinusoidal model,” Journal of Speech, Language, and Hearing Research, **37**, 449.

Kates, J. M. (**1993**). “Toward a theory of optimal hearing aid processing,” The Journal of Rehabilitation Research and Development, **30**, 39–48.

Kates, J. M. (**2005**). “Principles of digital dynamic-range compression,” Trends in Amplification, **9**, 45–76.

Keilson, S. E., Richards, V. M., Wyman, B. T., and Young, E. D. (**1997**). “The representation of concurrent vowels in the cat anesthetized ventral cochlear nucleus: evidence for a periodicity-tagged spectral representation,” Journal of the Acoustical Society of America, **102**, 1056–1071.

Kewley-Port, D., Burkle, T., and Lee, J. (**2007**). "Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners," Journal of the Acoustical Society of America, **122**, 2365–2375.

Kiang, N. Y., Liberman, M. C., and Levine, R. A. (**1976**). "Auditory-nerve activity in cats exposed to ototoxic drugs and high-intensity sounds," The Annals of otology, rhinology, and laryngology, **85**, 752–768.

Kiefte, M., Kluender, K. R., and Rhode, W. S. (**2002**). "Synthetic speech stimuli spectrally normalized for nonhuman cochlear dimensions," Acoustics Research Letters Online, **3**, 41–46.

Klatt, D. H. (**1980**). "Software for a Cascade/Parallel Formant Synthesizer," Journal of the Acoustical Society of America, **67**, 971–995.

Krips, R., and Furst, M. (**2009**). "Stochastic properties of coincidence-detector neural cells," Neural Computation, **21**, 2524–2553.

Kujawa, S., and Liberman, M. (**2009**). "Adding insult to injury: cochlear nerve degeneration after 'temporary' noise-induced hearing loss," The Journal of Neuroscience, **29**, 14077–14085.

Larsen, E., Cedolin, L., and Delgutte, B. (**2008**). "Pitch representations in the auditory nerve: two concurrent complex tones," Journal of Neurophysiology, **100**, 1301–1319.

LePrell, G., Sachs, M. B., and May, B. J. (**1996**). "Representation of vowel-like spectra by discharge rate responses of individual auditory-nerve fibers," Auditory Neuroscience, **2**, 275–288.

Leshowitz, B. (**1975**). "Psychophysical tuning curves in normal and impaired ears," Journal of the Acoustical Society of America, **58**, S71.

Leshowitz, B. (**1976**). "Measurements of frequency selectivity in listeners with sensorineural hearing loss," Journal of the Acoustical Society of America, **59**, S2.

Liberman, M. C. (**1978**). "Auditory-nerve response from cats raised in a low-noise chamber," Journal of the Acoustical Society of America, **63**, 442–455.

Liberman, M. C. (**1984**). "Single-neuron labeling and chronic cochlear pathology I Threshold shift and characteristic-frequency shift," Hearing Research, **16**, 33–41.

Liberman, M. C., and Dodds, L. W. (**1984**). “Single-neuron labeling and chronic cochlear pathology III Stereocilia damage and alterations of threshold tuning curves,” Hearing Research, **16**, 55–74.

Liberman, M. C., and Dodds, L. W. (**1984**). “Single-neuron labeling and chronic cochlear pathology II Stereocilia damage and alterations of spontaneous discharge rates,” Hearing Research, **16**, 43–53.

Liberman, M. C., Gao, J., He, D. Z. Z., Wu, X., Jia, S., and Zuo, J. (**2002**). “Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier,” Nature, **419**, 300–304.

Liberman, M. C., and Kiang, N. Y. (**1984**). “Single-neuron labeling and chronic cochlear pathology IV Stereocilia damage and alterations in rate- and phase-level functions,” Hearing Research, **16**, 75–90.

Lin, H. W., Furman, A. C., Kujawa, S. G., and Liberman, M. C. (**2011**). “Primary neural degeneration in the Guinea pig cochlea after reversible noise-induced threshold shift,” Journal of the Association for Research in Otolaryngology, **12**, 605–16.

Loeb, G. E., White, M. W., and Merzenich, M. M. (**1983**). “Spatial cross-correlation A proposed mechanism for acoustic pitch perception,” Biological Cybernetics, **47**, 149–163.

Lopez-Poveda, E. A. (**2005**). “Spectral Processing by the Peripheral Auditory System: Facts and Models,” International Review of Neurobiology, **70**, 7–48.

Lopez-Poveda, E. A., and Barrios, P. (**2013**). “A perceptual model of auditory deafferentation,” Proceedings of Meetings on Acoustics, Acoustical Society of America.

Lopez-Poveda, E. A., and Eustaquio-Martin, A. (**2013**). “On the Controversy About the Sharpness of Human Cochlear Tuning,” Journal of the Association for Research in Otolaryngology, **14**, 673–686.

Lopez-Poveda, E. A., and Johannesen, P. T. (**2012**). “Behavioral estimates of the contribution of inner and outer hair cell dysfunction to individualized audiometric loss,” Journal of the Association for Research in Otolaryngology, **13**, 485–504.

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., and Moore, B. C. J. (**2006**). “Speech perception problems of the hearing impaired reflect inability to use temporal fine structure,” Proceedings of the National Academy of Sciences, **103**, 18866–18869.

Louage, D. H., van der Heijden, M., and Joris, P. X. (**2004**). "Temporal properties of responses to broadband noise in the auditory nerve," Journal of Neurophysiology, **91**, 2051–2065.

Maison, S. F., and Liberman, M. C. (**2000**). "Predicting vulnerability to acoustic injury with a noninvasive assay of olivocochlear reflex strength," Journal of Neuroscience, **20**, 4701–4707.

Maison, S. F., Usubuchi, H., and Liberman, M. C. (**2013**). "Efferent feedback minimizes cochlear neuropathy from moderate noise exposure," The Journal of Neuroscience, **33**, 5542–5552.

Makary, C. A., Shin, J., Kujawa, S. G., Liberman, M. C., and Merchant, S. N. (**2011**). "Age-Related Primary Cochlear Neuronal Degeneration in Human Temporal Bones," Journal of the Association for Research in Otolaryngology, **717**, 711–717.

May, B. J., Huang, A., Le Prell, G., and Hienz, R. D. (**1996**). "Vowel formant frequency discrimination in cats: Comparison of auditory nerve representations and psychophysical thresholds," Auditory Neuroscience, **3**, 135–162.

Miller, R. L., Calhoun, B. M., and Young, E. D. (**1999a**). "Discriminability of vowel representations in cat auditory-nerve fibers after acoustic trauma," Journal of the Acoustical Society of America, **105**, 311–325.

Miller, R. L., Calhoun, B. M., and Young, E. D. (**1999b**). "Contrast enhancement improves the representation of /ɛ/-like vowels in the hearing-impaired auditory nerve," Journal of the Acoustical Society of America, **106**, 2693–2708.

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," Journal of the Acoustical Society of America, **101**, 3602–3616.

Møller, A. (**1964**). "Effect of tympanic muscle activity on movement of the eardrum, acoustic impedance and cochlear microphonics," Acta Oto-Laryngologica, **58**, 525–534.

Moore, B. C. J. (**2000**). "Use of a loudness model for hearing aid fitting IV Fitting hearing aids with multi-channel compression so as to restore 'normal' loudness for speech at different levels," British Journal of Audiology, **34**, 165–177.

Moore, B. C. J. (**2004**). "Dead regions in the cochlea: conceptual foundations, diagnosis, and clinical applications," Ear and Hearing, **25**, 98–116.

Moore, B. C. J. (**2007**). *Cochlear Hearing Loss: Physiological, Psychological and Technical Issues*, Wiley, Chichester, 2nd ed.

Moore, B. C. J. (**2008**). "The choice of compression speed in hearing aids: theoretical and practical considerations and the role of individual differences," Trends in Amplification, **12**, 103–112.

Moore, B. C. J. (**2008**). "Advantages and disadvantages of fast and slow compression in hearing aids," Journal of the Acoustical Society of America, **123**, 3165.

Moore, B. C. J., and Glasberg, B. R. (**1988**). "A comparison of four methods of implementing automatic gain control (AGC) in hearing aids," British Journal of Audiology, **22**, 93–104.

Moore, B. C. J., and Glasberg, B. R. (**1997**). "A model of loudness perception applied to cochlear hearing loss," Auditory Neuroscience, **3**, 289–311.

Moore, B. C. J., and Glasberg, B. R. (**2004**). "A revised model of loudness perception applied to cochlear hearing loss," Hearing Research, **188**, 70–88.

Moore, B. C. J., Glasberg, B. R., and Stone, M. A. (**1999**). "Use of a loudness model for hearing aid fitting: III A general method for deriving initial fittings for hearing aids with multi-channel compression," British Journal of Audiology, **33**, 241–258.

Moore, B. C. J., and Moore, G. A. (**2003**). "Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelopes by normally hearing and hearing-impaired subjects," Hearing Research, **182**, 153–163.

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

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

Morest, D., Hutson, K., and Kwok, S. (**1990**). "Cytoarchitectonic atlas of the cochlear nucleus of the chinchilla, Chinchilla laniger," Journal of Comparative Neurology, **300**, 230–248.

Mulders, W. H., and Robertson, D. (**2002**). "Inputs from the cochlea and the inferior colliculus converge on olivocochlear neurones," Hearing Research, **167**, 206–213.

Mulrow, C. D., Aguilar, C., Endicott, J. E., Tuley, M. R., Velez, R., Charlip, W. S., Rhodes, M. C., et al. (**1990**). "Quality-of-life changes and hearing impairment A randomized trial," Annals of Internal Medicine, **113**, 188–194.

Mulrow, C. D., Tuley, M. R., and Aguilar, C. (**1992**). "Sustained benefits of hearing aids," Journal of Speech, Language, and Hearing Research, **35**, 1402–1405.

Ngan, E. M., and May, B. J. (**2001**). "Relationship between the auditory brainstem response and auditory nerve thresholds in cats with hearing loss," Hearing Research, **156**, 44–52.

Nordmann, A. S., Bohne, B. A., and Harding, G. W. (**2000**). "Histopathological differences between temporary and permanent threshold shift," Hearing Research, **139**, 13–30.

Owren, M., and Cardillo, G. (**2006**). "The relative roles of vowels and consonants in discriminating talker identity versus word meaning," Journal of the Acoustical Society of America, **119**, 1727–1739.

Palmer, A. R., and Moorjani, P. A. (**1993**). "Responses to speech signals in the normal and pathological peripheral auditory system," Progress in Brain Research, **97**, 107–115.

Palmer, A. R., and Shackleton, T. M. (**2009**). "Variation in the phase of response to low-frequency pure tones in the guinea pig auditory nerve as functions of stimulus level and frequency," Journal of the Association for Research in Otolaryngology, **10**, 233–250.

Perkel, D. H., Gerstein, G. L., and Moore, G. P. (**1967**). "Neuronal spike trains and stochastic point processes II Simultaneous spike trains," Biophysical Journal, **7**, 419–440.

Peters, R. W., Moore, B. C., and Baer, T. (**1998**). "Speech reception thresholds in noise with and without spectral and temporal dips for hearing-impaired and normally hearing people," Journal of the Acoustical Society of America, **103**, 577–587.

Pickles, J. (**1988**). *An Introduction to the Physiology of Hearing*, Academic Press, London, 2nd ed.

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

Plomp, R. (**1988**). “The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function,” Journal of the Acoustical Society of America, **83**, 2322–2327.

Pols, L., Kamp, L. Van Der, and Plomp, R. (**1969**). “Perceptual and physical space of vowel sounds,” Journal of the Acoustical Society of America, **46**, 458–467.

Recio, A., Rhode, W. S., Kiefte, M., and Kluender, K. R. (**2002**). “Responses to cochlear normalized speech stimuli in the auditory nerve of cat,” Journal of the Acoustical Society of America, **111**, 2213–2218.

Recio, A., Rich, N. C., Narayan, S. S., and Ruggero, M. A. (**1998**). “Basilar-membrane responses to clicks at the base of the chinchilla cochlea,” Journal of the Acoustical Society of America, **103**, 1972–1989.

Rhode, W. S. (**1971**). “Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique,” Journal of the Acoustical Society of America, **49**, 1218–1231.

Rossi-Katz, J. A., and Arehart, K. H. (**2005**). “Effects of cochlear hearing loss on perceptual grouping cues in competing-vowel perception,” Journal of the Acoustical Society of America, **118**, 2588.

Rothman, J. S., Young, E. D., and Manis, P. B. (**1993**). “Convergence of auditory nerve fibers onto bushy cells in the ventral cochlear nucleus: implications of a computational model,” Journal of Neurophysiology, **70**, 2562–2583.

Ruggero, M. A., and Rich, N. C. (**1991**). “Furosemide alters organ of corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane,” Journal of Neuroscience, **11**, 1057–1067.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., and Robles, L. (**1997**). “Basilar-membrane responses to tones at the base of the chinchilla cochlea,” Journal of the Acoustical Society of America, **101**, 2151–2163.

Ruggero, M. A., and Temchin, A. N. (**2005**). “Unexceptional sharpness of frequency tuning in the human cochlea,” Proceedings of the National Academy of Sciences, **102**, 18614–18619.

Ruggero, M., and Rich, N. (**1987**). “Timing of spike initiation in cochlear afferents: dependence on site of innervation,” Journal of Neurophysiology, **58**, 379–403.

Sachs, M. B., Voigt, H. F., and Young, E. D. (**1983**). "Auditory nerve representation of vowels in background noise," Journal of Neurophysiology, **50**, 27–45.

Sachs, M. B., Winslow, R. L., and Sokolowski, B. H. (**1989**). "A computational model for rate-level functions from cat auditory-nerve fibers," Hearing Research, **41**, 61–69.

Sachs, M. B., and Young, E. D. (**1979**). "Encoding of steady-state vowels in the auditory nerve: representation in terms of discharge rate," Journal of the Acoustical Society of America, **66**, 470–479.

Sakayori, S., Kitama, T., Chimoto, S., Qin, L., and Sato, Y. (**2002**). "Critical spectral regions for vowel identification," Neuroscience Research, **43**, 155–162.

Schilling, J. R., Miller, R. L., Sachs, M. B., and Young, E. D. (**1998**). "Frequency-shaped amplification changes the neural representation of speech with noise-induced hearing loss," Hearing Research, **117**, 57–70.

Schuknecht, H. F. (**1993**). *Pathology of the Ear*, Lea & Febiger, Philadelphia, 2nd ed.

Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Laurnagaray, D., Beaulac, S., et al. (**2005**). "The Desired Sensation Level multistage input/output algorithm," Trends in Amplification, **9**, 159–197.

Sergeyenko, Y., Lall, K., Liberman, M. C., and Kujawa, S. G. (**2013**). "Age-Related Cochlear Synaptopathy: An Early-Onset Contributor to Auditory Functional Decline," Journal of Neuroscience, **33**, 13686–13694.

Shamma, S. A. (**1985**). "Speech processing in the auditory system I: The representation of speech sounds in the responses of the auditory nerve," Journal of the Acoustical Society of America, **78**, 1612–1621.

Shamma, S. A. (**1985**). "Speech processing in the auditory system II: Lateral inhibition and the central processing of speech evoked activity in the auditory nerve," Journal of the Acoustical Society of America, **78**, 1622–1632.

Shamma, S., and Klein, D. (**2000**). "The case of the missing pitch templates: how harmonic templates emerge in the early auditory system," Journal of the Acoustical Society of America, **107**, 2631–2644.

Shera, C. A., Guinan, J. J., and Oxenham, A. J. (**2002**). "Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements," Proceedings of the National Academy of Sciences, **99**, 3318–3323.

Shera, C. A., Guinan, J. J., and Oxenham, A. J. (**2007**). "Otoacoustic estimates of cochlear tuning: Validation in the chinchilla," Journal of the Association for Research in Otolaryngology, **11**, 343–365.

Shera, C. A., Tubis, A., and Talmadge, C. L. (**2008**). "Testing coherent reflection in chinchilla: Auditory-nerve responses predict stimulus-frequency emissions," Journal of the Acoustical Society of America, **123**, 3851.

Shi, L.-F., Carney, L. H., and Doherty, K. A. (**2006**). "Correction of the peripheral spatiotemporal response pattern: a potential new signal-processing strategy," Journal of Speech, Language, and Hearing Research, **49**, 848–855.

Shinn-Cunningham, B. G., and Best, V. (**2008**). "Selective attention in normal and impaired hearing," Trends in Amplification, **12**, 283–99.

Shofner, W. (**1999**). "Responses of cochlear nucleus units in the chinchilla to iterated rippled noises: analysis of neural autocorrelograms," Journal of Neurophysiology, **81**, 2662–2674.

Silkes, S. M., and Geisler, C. D. (**1991**). "Responses of 'lower-spontaneous-rate' auditory-nerve fibers to speech syllables presented in noise I: General characteristics," Journal of the Acoustical Society of America, **90**, 3122–3139.

Simpson, A. M., Moore, B. C., and Glasberg, B. R. (**1990**). "Spectral enhancement to improve the intelligibility of speech in noise for hearing-impaired listeners," Acta Oto-Laryngologica, **469**, 101–107.

Smith, J. (**2007**). *Introduction to digital filters: with audio applications*, W3K Publishing.

Souza, P. E. (**2002**). "Effects of Compression on Speech Acoustics, Intelligibility, and Sound Quality," Trends in Amplification, **6**, 131–165.

Souza, P. E., and Tremblay, K. L. (**2006**). "New perspectives on assessing amplification effects," Trends in Amplification, **10**, 119–143.

Stevens, G., and Flaxman, S. (**2013**). "Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries," The European Journal of Public Health, **23**, 146–152.

Summers, V., and Leek, M. R. (**1998**). "FO processing and the separation of competing speech signals by listeners with normal hearing and with hearing loss," Journal of Speech, Language, and Hearing Research, **41**, 1294–1306.

Wang, G. I., and Delgutte, B. (**2012**). “Sensitivity of Cochlear Nucleus Neurons to Spatio-Temporal Changes in Auditory Nerve Activity,” Journal of Neurophysiology, **108**, 3172–3195.

Wang, J., Powers, N. L., Hofstetter, P., Trautwein, P., Ding, D., and Salvi, R. (**1997**). “Effects of selective inner hair cell loss on auditory nerve fiber threshold, tuning and spontaneous and driven discharge rate,” Hearing Research, **107**, 67–82.

Weiss, T. F., and Rose, C. (**1988**). “A comparison of synchronization filters in different auditory receptor organs,” Hearing Research, **33**, 175–179.

Wong, J. C., Miller, R. L., Calhoun, B. M., Sachs, M. B., and Young, E. D. (**1998**). “Effects of high sound levels on responses to the vowel /ɛ/ in cat auditory nerve,” Hearing Research, **123**, 61–77.

Woolf, N. K., Ryan, A. F., and Bone, R. C. (**1981**). “Neural phase-locking properties in the absence of cochlear outer hair cells,” Hearing Research, **4**, 335–346.

Yang, J., Luo, F.-L., and Nehorai, A. (**2003**). “Spectral contrast enhancement: Algorithms and comparisons,” Speech Communication, **39**, 33–46.

Young, E. D. (**2008**). “Neural representation of spectral and temporal information in speech,” Philosophical Transactions of the Royal Society B: Biological Sciences, **363**, 923–945.

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,” Journal of the Acoustical Society of America, **66**, 1381–1403.

Zeyl, T., and Bruce, I. (**In Press**). “Analysis of Spatiotemporal Pattern Correction using a Computational Model of the Auditory Periphery,” Ear and Hearing.

Zhang, X., Heinz, M. G., Bruce, I. C., and Carney, L. H. (**2001**). “A phenomenological model for the responses of auditory-nerve fibers: I Nonlinear tuning with compression and suppression,” Journal of the Acoustical Society of America, **109**, 648–670.

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,” Journal of the Acoustical Society of America, **120**, 1446–1466.

Zilany, M. S. A., and Bruce, I. C. (**2007a**). “Predictions of speech intelligibility with a model of the normal and impaired auditory-periphery,” Proceedings of the 3rd International IEEE EMBS Conference on Neural Engineering, Kohala Coast, Hawaii, USA.

Zilany, M. S. A., and Bruce, I. C. (**2007b**). “Representation of the vowel /ɛ/ in normal and impaired auditory nerve fibers: model predictions of responses in cats,” Journal of the Acoustical Society of America, **122**, 402–417.

Zilany, M. S. A., Bruce, I. C., Nelson, P. C., and Carney, L. H. (**2009**). “A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics,” Journal of the Acoustical Society of America, **126**, 2390–412.

Zwicker, E. (**1974**). “On a Psychoacoustical Equivalent of Tuning Curves,” In E. Zwicker and E. Terhardt (Eds.), Facts and Models of Hearing, Springer-Verlag, Berlin, pp. 132–141.

VITA

VITA

Jonathan Boley
Graduate School, Purdue University

Research Interests

Auditory perception, psychoacoustics, auditory neurophysiology, computational modeling, digital signal processing

Education

Ph.D. in Biomedical Engineering, December 2013
Purdue University, West Lafayette, IN, USA.
Master of Science in Music Engineering Technology, May 2005
University of Miami, Coral Gables, FL, USA.
Bachelor of Science in Electrical Engineering, May 2003
University of Illinois at Urbana-Champaign, Champaign, IL, USA.

Professional Experience

2012 - present	Research Scientist	GN ReSound, Glenview, IL
2003 - 2007	Senior DSP Engineer	Shure Inc., Niles, IL
2000 - 2002	Engineering Intern	Motorola Inc., Schaumburg, IL

Sponsored Research Experience

Effects of hearing aid amplification on robust speech coding, NIH (NIDCD, F31-DC010966), 1/1/10-12/31/11. $83,180 Principal Investigator.

Book Chapters & Conference Proceedings

Boley, J., C. Danner, and M. Lester, "*Measuring Dynamics: Comparing and Contrasting Algorithms for the Computation of Dynamic Range*," in Proceedings of the 129th Convention of the Audio Engineering Society, November 2010.

Gaston, L., **J. Boley**, S. Selter, and J. Ratterman, "*The Influence of Individual Audio Impairments on Perceived Video Quality*," in Proceedings of the 128th Convention of the Audio Engineering Society, May 2010.

Heinz, M., J. Swaminathan, **J. Boley**, and S. Kale, "*Across-Fiber Coding of Temporal Fine-Structure: Effects of Noise-Induced Hearing Loss on Auditory Nerve Responses,"* in The Neurophysiological Bases of Auditory Perception. Springer (New York), March 2010.

Boley, J. and M. Lester, "*Statistical Analysis of ABX Results Using Signal Detection Theory*," in Proceedings of the 127th Convention of the Audio Engineering Society, October 2009.

Lester, M. and **J. Boley**, "*The Effects of Latency on Live Sound Monitoring*," in Proceedings of the 123rd Convention of the Audio Engineering Society, October 2007.

Boley, J. "*Auditory Component Analysis*," in Proceedings of the 121st Convention of the Audio Engineering Society, October 2006.

Conference Presentations

Boley, J. and M. Heinz, "*Impaired Spatiotemporal Coding of Vowels in Noise*," International Hearing Aid Research Conference, August 2012.

Boley, J. and M. Heinz, "*Predicted Effects of Amplification on Spatiotemporal Coding of Vowels in Noise*," International Hearing Aid Research Conference, August 2010.

Kale, S., **J. Boley,** J. Swaminathan, M. Heinz, *"Within and across fiber temporal fine structure coding following noise induced hearing loss,"* 33rd Midwinter Meeting of the Association for Research in Otolaryngology, February 2010.

Boley, J. and M. Heinz, *"Quantifying the Effects of Hearing Aid Dynamics on Temporal Coding in the Auditory Nerve,"* First International Symposium on Audible Acoustics in Medicine and Physiology, September 2008.

Patent Applications

Rits, S. K., **J. Boley**, O Masciarotte, R Bichariju. "Mobile device application." U.S. Patent Application 12/932,620, filed February 28, 2011.

Review of Manuscripts

Audio Engineering Society
Ear and Hearing
Institute of Electrical and Electronics Engineers

Professional Affiliations

Audio Engineering Society, Member, 2001-Present

Technical Committees: Perception and Subjective Evaluation, Signal Processing, Audio Coding

Leadership Roles: Founder & past chair of the University of Illinois student section, past Treasurer of the Chicago section

IEEE Signal Processing Society, Member, 1998-Present
Motion Picture Experts Group (MPEG), Member, 2003-2005
Acoustical Society of America, Member, 2001-Present

PUBLICATION

Chapter 56
Across-Fiber Coding of Temporal Fine-Structure: Effects of Noise-Induced Hearing Loss on Auditory-Nerve Responses

Michael G. Heinz, Jayaganesh Swaminathan, Jonathan D. Boley, and Sushrut Kale

Abstract Recent psychophysical evidence suggests that listeners with sensorineural hearing loss (SNHL) have a reduced ability to use temporal fine-structure cues. These results have renewed an interest in the effects of SNHL on the neural coding of fine structure. The lack of convincing evidence that SNHL affects within-fiber phase locking has led to the hypothesis that degraded across-fiber temporal coding may underlie this perceptual effect. Spike trains were recorded from auditory-nerve (AN) fibers in chinchillas with normal hearing and with noise-induced hearing loss. A spectro-temporal manipulation procedure was used to predict spatiotemporal patterns for characteristic frequencies (CFs) spanning up to an octave range from the responses of individual AN fibers to a stimulus presented with sampling rates spanning an octave range. Shuffled cross-correlogram analyses were used to quantify across-CF fine-structure coding in terms of both a neural cross-correlation coefficient and a characteristic delay. Neural cross-correlation for fine-structure decreased and the estimated traveling-wave delay increased with increases in CF separation for both normal and impaired fibers. However, the range of CF separations over which significant correlated activity existed was wider, and the estimated traveling-wave delay was less for impaired AN fibers. Both of these effects of SNHL on across-CF coding have important implications for spatiotemporal theories of speech coding.

Keywords Auditory nerve • Sensorineural hearing loss • Across-fiber coding • Temporal fine structure • Traveling wave delay

M.G. Heinz (✉)
Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN 47907, USA
e-mail: mheinz@purdue.edu

E.A. Lopez-Poveda et al. (eds.), *The Neurophysiological Bases of Auditory Perception*,
DOI 10.1007/978-1-4419-5686-6_56,

56.1 Introduction

Recent psychophysical studies suggest that listeners with sensorineural hearing loss (SNHL) have a reduced ability to use temporal fine-structure cues, which is correlated with their reduced understanding of speech in complex backgrounds (Lorenzi et al. 2006; Hopkins and Moore 2007). These perceptual results have renewed an interest in the effects of SNHL on neural coding of temporal fine structure, both within single auditory-nerve (AN) fibers and across fibers with different characteristic frequencies (CFs). There is conflicting evidence as to whether within-fiber encoding of fine-structure (i.e., phase locking) is degraded following SNHL (Harrison and Evans 1979; Woolf et al. 1981; Miller et al. 1997). Thus, it has been hypothesized that degraded across-fiber temporal coding due to broader tuning and associated shallower phase responses could underlie these perceptual deficits, e.g., as implicated in spatiotemporal theories of speech coding (e.g., Shamma 1985). However, effects of SNHL on across-CF coding have been difficult to examine because of experimental limitations associated with sparse CF sampling in AN population studies and variability in CF estimates (Chintanpalli and Heinz 2007).

The present study compared the effects of noise-induced hearing loss on within- and across-fiber coding of temporal fine structure. Across-fiber variability was minimized by using responses of individual AN fibers to frequency-shifted stimuli to predict responses of a population of AN fibers with differing CFs to a single stimulus. Shuffled auto- and cross-correlograms were used to quantify across-CF temporal coding in terms of both a neural cross-correlation coefficient and a characteristic delay (CD) that estimates the traveling-wave delay between two CFs.

56.2 Methods

56.2.1 Experimental Procedures

All procedures were approved by the Purdue Animal Care and Use Committee. Neural recordings were made from AN fibers in two anesthetized chinchillas using standard procedures (e.g., Heinz and Young 2004; Chintanpalli and Heinz 2007). Spike times were measured with 10-μs resolution. Isolated fibers were characterized by an automated tuning-curve algorithm to determine fiber CF, threshold, and Q_{10}. Impaired-fiber CFs were chosen by hand near the steep high-frequency slope, which better estimates the original CF prior to SNHL (Liberman 1984). Spontaneous rate was determined over 20 s and PST histograms were measured to verify AN responses based on latency. Noise-induced hearing loss was produced in one chinchilla by presenting a 50-Hz-wide noise band centered at 2 kHz continuously for 4 h at 115 dB SPL, after which the animal recovered for 6 weeks. Consistent with previous studies in which noise exposure produced mixed outer- and inner-hair cell damage (Liberman 1984; Heinz and Young 2004), a moderate hearing loss was produced with thresholds elevated by ~30–50 dB and broadened tuning in all fibers (Q_{10}s below the normal range for chinchillas).

Neural responses were recorded to both a broadband noise and a speech sentence. Both stimuli were 1.7-s in duration at the baseline sampling frequency of 33 kHz. Positive and negative polarity versions of both stimuli were presented at 7 or 9 different sampling frequencies spanning a range of up to 1 octave. All stimuli were presented in an interleaved manner, with a new stimulus presented every 2.9 s. Both stimuli were presented to each AN fiber at 10 or 20 dB above stimulus threshold for that fiber, as determined by measured rate-level functions. Stimuli were repeated until ~2,000 spikes were recorded for all stimuli, or until the fiber was lost. Data are presented from 17 normal-hearing fibers and 19 hearing-impaired fibers.

56.2.2 Predicting Spatiotemporal Patterns from Individual AN Fibers

The ability to quantify across-CF temporal coding is significantly limited by sparse sampling and across-fiber variability inherent in AN population studies, as well as by variability in CF estimates from automated tuning-curve algorithms (Chintanpalli and Heinz 2007). These limitations are particularly true with SNHL. To overcome these limitations, a spectro-temporal manipulation procedure (STMP) was used to predict the spatiotemporal response of a population of AN fibers with a range of CFs responding to a single stimulus from responses of a single AN fiber to frequency-shifted stimuli (Heinz 2007). The STMP relies on scaling invariance in cochlear mechanics and is similar to procedures that have been used to study spatiotemporal coding of pitch (Larsen et al. 2008). Although some cochlear properties are not scaling invariant (e.g., roll-off in phase locking, refractoriness, adaptation), these effects are likely to be negligible in comparisons between normal and impaired responses over ±0.5 octaves (Larsen et al. 2008).

56.2.3 Within-CF and Across-CF Temporal Analyses

Shuffled correlogram analyses (Louage et al. 2004; Joris et al. 2006; Heinz and Swaminathan 2009) were used to quantify within- and across-CF fine-structure coding from single AN-fiber responses to broadband noise and speech. Within-fiber temporal coding was evaluated based on normalized shuffled auto correlograms (SACs, thick lines, Fig. 56.1a, b), which were computed by comparing spike times between all possible pairs of stimulus presentations for a given effective CF from the STMP. For each pair, intervals between every spike in the first spike train and every spike in the second spike train were tallied with a 50-μs binwidth to create a shuffled all-order interval histogram. For each AN fiber, SACs were computed for each effective CF from the STMP. Figure 56.1 shows correlogram analyses for two effective CFs separated by 0.5 octaves based on spike trains recorded in response to broadband noise. Responses to positive and negative polarity versions of each stimulus were recorded because polarity inversion inverts stimulus fine-structure while not affecting stimulus envelope. Cross-polarity auto correlograms (XpACs,

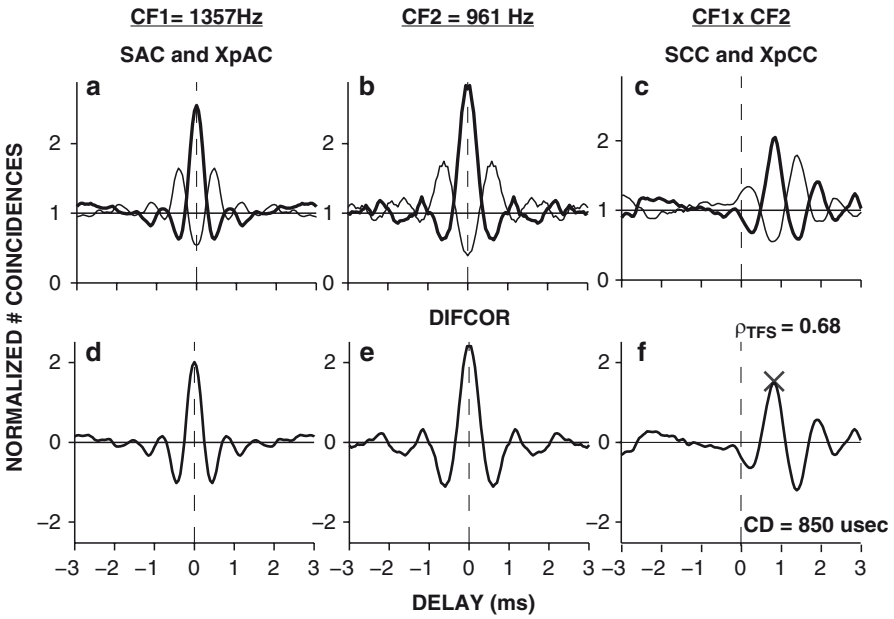

Fig. 56.1 Within- (cols. 1–2) and across-CF (col. 3) temporal coding based on shuffled correlograms. (**a**, **b**) Auto correlograms: SACs (*thick line*), XpACs (*thin line*). (**c**) Cross correlograms: SCC (*thick line*), XpCC (*thin line*). (**d–f**) DIFCORs emphasize fine structure by subtracting XpAC from SAC (or XpCC from SCC). Auto-correlogram DIFCOR peak heights quantify within-fiber fine structure. Across-CF coding was quantified with neural cross-correlation coefficients (ρ_{TFS}), computed as the ratio of cross-correlogram DIFCOR peak height (**f**) to the geometric mean of auto-correlogram DIFCOR peak heights (**d**, **e**). Characteristic delay (CD) (× in panel **f**) estimates traveling-wave delay between the two effective CFs 0.5 octaves apart. Spike trains recorded from one impaired AN fiber responding to a broadband noise with two sampling rates that differed by 0.5 octaves. STMP was used to predict responses of two effective CFs 0.5-octaves apart. CF = 1.36 kHz, thresh. = 49 dB SPL, $Q_{10} = 0.9$, spont. rate = 64 spikes/s

thin lines, Fig. 56.1a, b) were computed by tallying intervals between all spikes in response to positive and negative polarity versions of the stimulus. DIFCORs computed by subtracting the XpAC from the SAC thus emphasize fine structure coding, which was significant for both effective CFs in Fig. 56.1d, e.

Across-CF fine-structure coding was evaluated based on shuffled cross correlograms (SCCs, thick line, Fig. 56.1c) and cross-polarity, cross correlograms (XpCCs, thin line, Fig. 56.1c), which were computed by comparing spike trains across a pair of effective CFs from the STMP. For each effective CF pair for one AN fiber, the cross-correlogram DIFCOR was used to evaluate across-CF fine-structure coding with two metrics. A neural cross-correlation coefficient (ρ_{TFS}) was used to represent the degree of similarity between two spike-train responses (Heinz and Swaminathan 2009), and was computed as the ratio of the peak height of the cross-correlogram DIFCOR (Fig. 56.1f) to the geometric mean of the auto-correlogram DIFCOR peak heights (Fig. 56.1d, e). A significant benefit of this self-normalized similarity

metric is that the degree of cross correlation is evaluated relative to the strength of within-fiber fine-structure coding for each fiber individually, which varies with differences in CF, spontaneous rate, and stimulus level (Louage et al. 2004). The computed value of $\rho_{TFS} = 0.68$ indicates significant common temporal fine-structure in these hearing-impaired responses for effective CFs separated by 0.5 octaves. CD (× in Fig. 56.1f) of the cross-correlogram DIFCOR provides an estimate of the traveling wave delay between the two cochlear locations represented by these CFs (Joris et al. 2006). A CD of 850 μs was estimated for the two effective CFs separated by 0.5 octaves (Fig. 56.1f).

56.3 Results

Figure 56.2 illustrates the single-fiber analyses performed on each AN fiber. A normal-hearing fiber is compared to an impaired fiber with a similar CF (1.3 kHz) in terms of their tuning curves (Fig. 56.2a), the predicted effect of CF separation on cross-CF correlation (Fig. 56.2b) and CD (Fig. 56.2c). The normal-hearing tuning curve represents a low-threshold, high-spontaneous rate fiber with sharp tuning. The impaired tuning curve shows broad tuning without a defined tip. This tuning curve is representative in shape of all impaired fibers in this study, which had thresholds ranging from 40 to 60 dB SPL and CFs ranging from 0.7 to 5 kHz.

The effect of CF separation on cross-CF correlation for broadband noise is shown in Fig. 56.2b for both AN fibers. Neural cross-correlation coefficients (ρ_{TFS}) were computed for all effective CF pairs derived from the STMP for each AN fiber. Seven effective CFs predicted for the normal-hearing fibers produced 21 pairs with CF separations ranging from 0.05 to 0.55 octaves. For impaired fibers, 36 pairs with CF separations ranging from 0.05 to 1.0 octaves were obtained from nine effective CFs. The variation in ρ_{TFS} with CF separation was fit with fourth-order polynomials constrained to equal 1.0 at a CF separation of 0 octaves. Neural cross correlation decreased monotonically with increasing CF separation for all normal-hearing fibers. The example shown in Fig. 56.2b decreased to ~0.3 for the maximum CF separation of 0.55 octaves. Impaired AN fibers also showed a decrease in ρ_{TFS} as CF separation increased; however, the decrease was often less steep and sometimes did not drop below 0.6 for the largest CF separation of 1.0 octaves (especially for the speech stimulus). Based on the fitted lines, the width of the correlated region was estimated by the smallest CF separation at which ρ_{TFS} fell below 0.6. This normal-hearing fiber demonstrated correlated activity above $\rho_{TFS} = 0.6$ out to a 0.34-octave CF separation, whereas the impaired fiber demonstrated a much wider CF-separation range (0.81 octaves) of correlated activity.

The increase in CD with increased CF separation is shown in Fig. 56.2c for the same two AN fibers. CD derived from the cross-correlogram DIFCORs is represented in units of CF cycles and is plotted as a function of CF separation for all effective-CF pairs. For all normal-hearing and hearing-impaired fibers, CD increased very systematically across the entire range of CF separations and was

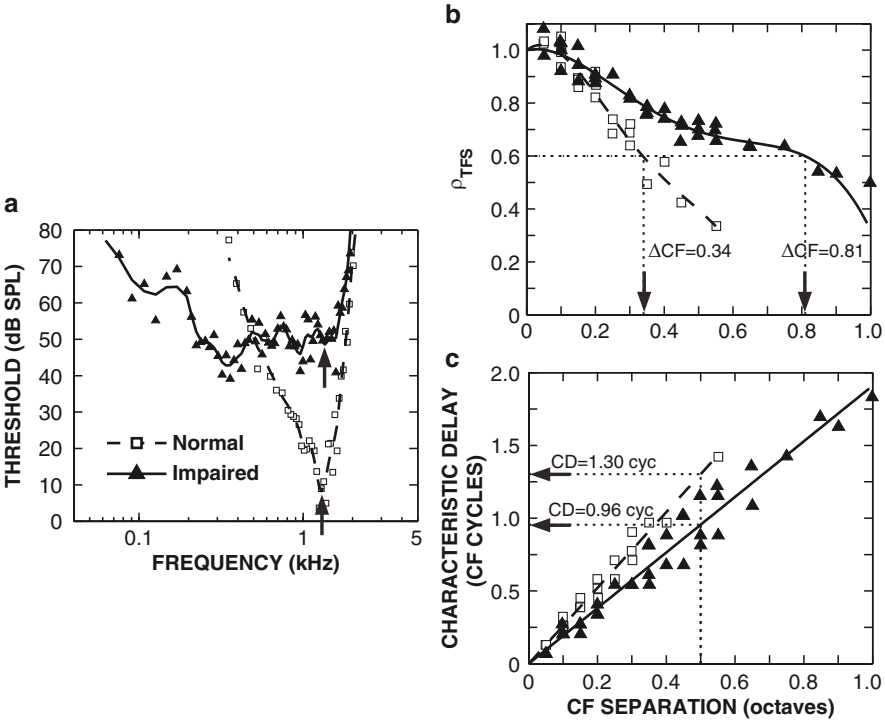

Fig. 56.2 Effect of CF separation on across-CF fine-structure coding for a normal-hearing and a hearing-impaired AN fiber with similar CFs. (**a**) Tuning curves. (**b**) Neural cross-correlation coefficients (ρ_{TFS}) as a function of CF separation. The smallest CF separation (ΔCF) at which ρ_{TFS} dropped to 0.6 was computed based on a fourth-order polynomial fit. (**c**) Characteristic delay (CD) increased linearly as a function of CF separation. CD was measured (in ms) from cross-correlogram DIFCORs and converted to CF cycles by multiplying by CF in kHz. The CD at a CF separation of 0.5 octaves was computed based on linear fits. AN fibers: normal (*open squares*, *dashed lines*): CF = 1.29 kHz, thresh. = 8 dB SPL, $Q_{10} = 4.1$, spont. rate = 91 spikes/s; impaired (*filled triangles*, *solid lines*): CF = 1.36 kHz, thresh. = 49 dB SPL, $Q_{10} = 0.9$, spont. rate = 64 spikes/s. Noise level: 10 dB above threshold for each fiber

well fit by a linear function constrained to equal 0 for no CF separation. The rate of increase in CD as CF separation increased was less for the impaired AN fiber than for the normal-hearing fiber (Fig. 56.2c). Thus, for all effective-CF separations the traveling-wave delay was predicted to be reduced following SNHL, consistent with broader tuning and the associated shallower phase transition. To quantify this effect, the CD at a 0.5-octave separation was computed for each AN fiber based on the fitted lines. For the examples shown, the CD at a 0.5-octave separation was 1.3 cycles for the normal-hearing fiber and 0.96 cycles for the impaired fiber, i.e., more than a quarter-cycle difference.

The normal-hearing and hearing-impaired populations of AN fibers are compared in Fig. 56.3 in terms of both within- and across-CF coding of temporal fine structure for broadband noise and speech responses. Auto-correlogram DIFCOR

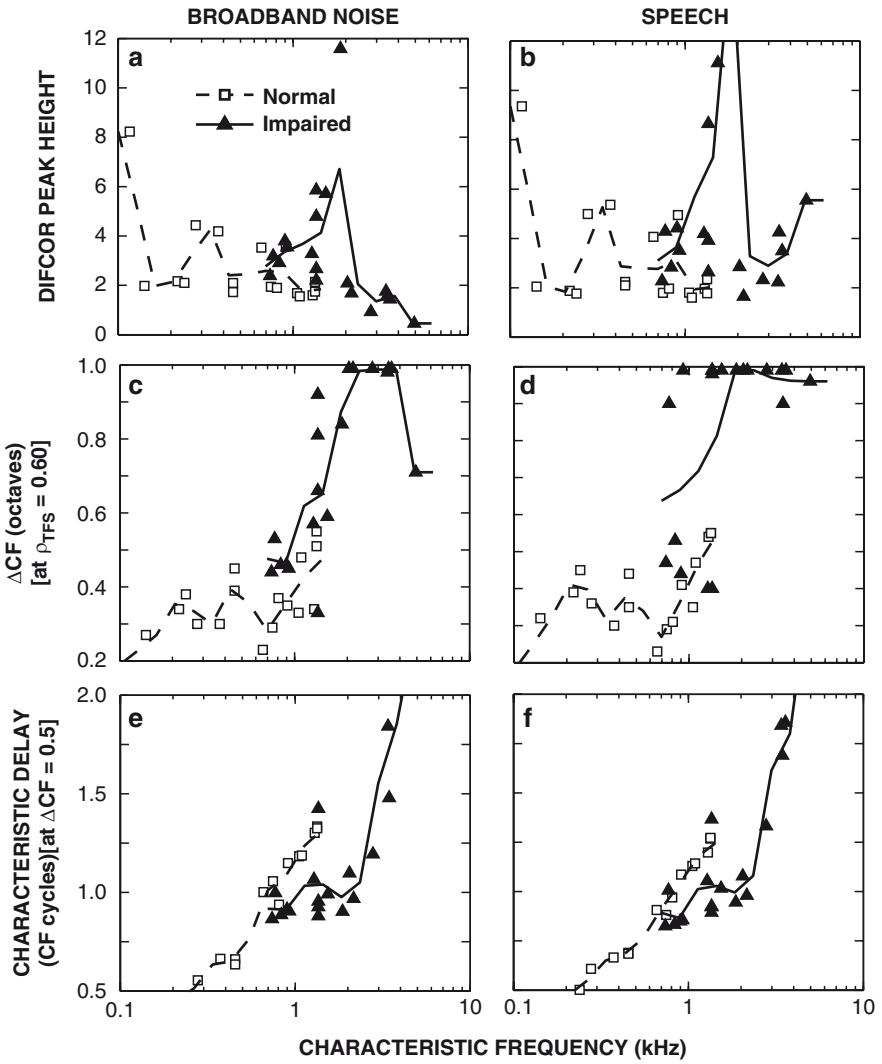

Fig. 56.3 Comparison of fine-structure coding between normal-hearing and hearing-impaired AN-fiber populations for broadband noise (*left*) and speech (*right*). (**a**, **b**) Within-fiber fine-structure coding represented by DIFCOR peak heights. (**c**, **d**) Smallest CF separation (ΔCF) at which ρ_{TFS} dropped to 0.6 represents the width of correlated activity. (**e**, **f**) Characteristic delay (CD) at a CF separation of 0.5 octaves estimates phase delay (in CF cycles) across two cochlear locations 0.5 octaves apart. Lines are weighted moving averages from a 0.7-octave-wide triangular window in steps of 0.35 octaves. All stimuli: 10 or 20 dB above stimulus threshold for each fiber

peak heights represent the strength of within-fiber fine-structure coding (Fig. 56.3a, b) and were not reduced in the hearing-impaired population for either broadband noise or speech. In fact, DIFCOR peak heights were slightly higher on average in

the impaired population within the CF region from 0.7 to 1.5 kHz, where both normal and impaired data existed in these limited populations. A few fibers in the hearing-impaired population showed much larger DIFCOR peak heights than most of the normal-hearing data. All fibers with DIFCOR peak heights above 6 were low-spontaneous rate fibers, which typically have larger DIFCOR peak heights (Louage et al. 2004) and were more prevalent in the impaired population, consistent with previous studies (e.g., Heinz and Young 2004). Thus, there was no observed degradation in the strength of within-fiber coding of temporal fine structure, consistent with several previous studies (Harrison and Evans 1979; Miller et al. 1997).

Degradations were observed in across-CF coding of temporal fine-structure. The range of effective-CF separations over which correlated activity existed above $\rho_{TFS}=0.6$ is compared between the normal-hearing and hearing-impaired populations in Fig. 56.3c, d. For most normal-hearing AN fibers, correlated activity existed over a CF separation range between 0.2 and 0.5 octaves for both broadband noise and speech. For impaired fibers with CFs between 0.7 and 1.5 kHz, the width of correlated activity for broadband noise was 0.1–0.2 octaves wider than for the normal-hearing fibers, as indicated by the trend lines. For speech responses, this degradation was more significant, with many impaired fibers showing correlated activity that did not drop to $\rho_{TFS}=0.6$ over the entire 1.0-octave range of effective-CF separations.

A decrease in CD between effective CFs was observed in the impaired population for both broadband noise and speech responses (Fig. 56.3e, f). CD at a CF separation of 0.5 octaves is plotted against fiber CF for both populations. For the normal-hearing population, the delay in CF cycles increased systematically from 0.25 cycles for fiber CFs ~ 150 Hz to more than 1.25 cycles for CFs just above 1 kHz. This trend is consistent with sharper tuning and increased cochlear delays (in cycles) with increased CF as inferred from otoacoustic emissions (Shera et al. 2002). CDs were reduced by ~ 0.25 cycles in impaired AN fibers with CFs between 0.7 and 1.5 kHz. Impaired CD was roughly constant around 1 cycle for CFs from 0.7 to 2 kHz, and increased at higher CFs. Note that unlike the cross-correlation effects (Figs. 56.3c, d), CD effects were remarkably similar between broadband noise and speech responses. The reduction in CD with SNHL was smaller for CFs below 1 kHz than for those above 1 kHz; however, this CF-dependence may simply result from the specific noise-induction procedure used (i.e., 2-kHz exposure frequency), which produces the most significant hearing loss above 1 kHz (Heinz and Young 2004). Further study is necessary to evaluate SNHL effects at low CFs, given that listeners with high-frequency hearing loss and near-normal thresholds at low CFs have been shown to have a perceptual TFS deficit for lowpass filtered speech (Lorenzi et al. 2009).

It should be noted that the effect of SNHL on cochlear phase delays was not to eliminate the traveling-wave delay, but simply to reduce the across-CF delay by roughly 0.25-cycles in the present data for a moderate hearing loss. The size of this effect is consistent with level-dependent changes in the relative phase above and below CF in AN fiber responses (Palmer and Shackleton 2009). The relative phase for a 0.5-octave frequency difference around CF can vary by a quarter to a half cycle over a 40–50 dB range of tone level, which is presumably related to normal

outer-hair-cell function associated with nonlinear cochlear tuning. Although the size of this effect is small relative to the overall phase delay of ~1 cycle for the 0.5-octave CF separation considered here, a quarter cycle phase shift (e.g., sine to cosine) represents the difference between in-phase and uncorrelated activity. Thus, characteristic-delay changes of this size would be significant in terms of any mechanism that relied on across-CF correlation at a fixed delay (e.g., cross-CF coincidence detection). Note that ρ_{TFS} represents a different effect, in that it quantifies the maximum correlation across all delays (i.e., computed at CD).

56.4 Discussion

The most significant effects of SNHL on fine-structure coding in AN fibers were in terms of across-CF coding rather than within-fiber coding, for which no degradation was observed. Across-CF coding was degraded in terms of both an increase in the cross-CF correlation and a decrease in CD between effective CFs. Broadening of the CF region over which correlated activity exists with SNHL could be perceptually significant for complex sounds because it would reduce the number of available independent neural information channels. A reduction in traveling-wave delay across CF with SNHL would result in a more coincident representation of temporal features across fibers that could degrade normal spatiotemporal response patterns. These patterns have been hypothesized to provide robust neural cues for a range of perceptual phenomena, including the coding of speech, pitch, and intensity, as well as tone detection in noise (Shamma 1985; Heinz et al. 2001; Carney et al. 2002; Heinz 2007; Larsen et al. 2008). Changes in across-CF delays would also have implications for binaural theories that rely on cochlear disparities as a source for interaural delays (Shamma et al. 1989; Joris et al. 2006).

Thus, these preliminary data suggest that the effects of SNHL on across-CF coding are significant and need to be considered when interpreting the reduced perceptual ability of listeners with SNHL to use fine-structure cues (e.g., Lorenzi et al. 2006; Hopkins and Moore 2007). If these physiological effects were perceptually relevant, they would suggest the need for new avenues into improving strategies for auditory prostheses, which currently do not attempt to restore normal spatiotemporal response patterns.

Acknowledgments This work was supported by grants from NIH (R03-DC07348) and the National Organization for Hearing Research Foundation.

References

Carney LH, Heinz MG, Evilsizer ME, Gilkey RH, Colburn HS (2002) Auditory phase opponency: a temporal model for masked detection at low frequencies. Acustica-Acta Acustica 88:334–347

Chintanpalli A, Heinz MG (2007) The effect of auditory-nerve response variability on estimates of tuning curves. J Acoust Soc Am 122:EL203–EL209

Harrison RV, Evans EF (1979) Some aspects of temporal coding by single cochlear fibres from regions of cochlear hair cell degeneration in the guinea pig. Arch Otorhinolaryngol 224:71–78

Heinz MG (2007) Spatiotemporal encoding of vowels in noise studied with the responses of individual auditory nerve fibers. In: Kollmeier B, Klump G, Hohmann V, Langemann U, Mauermann M, Uppenkamp S, Verhey J (eds) Hearing – from sensory processing to perception. Springer-Verlag, Berlin, pp 107–115

Heinz MG, Swaminathan J (2009) Quantifying envelope and fine-structure coding in auditory-nerve responses to chimaeric speech. J Assoc Res Otolaryngol 10:407–423

Heinz MG, Young ED (2004) Response growth with sound level in auditory-nerve fibers after noise-induced hearing loss. J Neurophysiol 91:784–795

Heinz MG, Colburn HS, Carney LH (2001) Rate and timing cues associated with the cochlear amplifier: level discrimination based on monaural cross-frequency coincidence detection. J Acoust Soc Am 110:2065–2084

Hopkins K, Moore BC (2007) Moderate cochlear hearing loss leads to a reduced ability to use temporal fine structure information. J Acoust Soc Am 122:1055–1068

Joris PX, Van de Sande B, Louage DH, van der Heijden M (2006) Binaural and cochlear disparities. Proc Natl Acad Sci USA 103:12917–12922

Larsen E, Cedolin L, Delgutte B (2008) Pitch representations in the auditory nerve: two concurrent complex tones. J Neurophysiol 100:1301–1319

Liberman MC (1984) Single-neuron labeling and chronic cochlear pathology. I. Threshold shift and characteristic-frequency shift. Hear Res 16:33–41

Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BC (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. Proc Natl Acad Sci USA 103:18866–18869

Lorenzi C, Debruille L, Garnier S, Fleuriot P, Moore BC (2009) Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal. J Acoust Soc Am 125:27–30

Louage DH, Van Der Heijden M, Joris PX (2004) Temporal properties of responses to broadband noise in the auditory nerve. J Neurophysiol 91:2051–2065

Miller RL, Schilling JR, Franck KR, Young ED (1997) Effects of acoustic trauma on the representation of the vowel /ε/ in cat auditory nerve fibers. J Acoust Soc Am 101:3602–3616

Palmer AR, Shackleton TM (2009) Variation in the phase of response to low-frequency pure tones in the guinea pig auditory nerve as functions of stimulus level and frequency. J Assoc Res Otolaryngol 10:233–250

Shamma SA (1985) Speech processing in the auditory system. I: The representation of speech sounds in the responses of the auditory nerve. J Acoust Soc Am 78:1612–1621

Shamma SA, Shen NM, Gopalaswamy P (1989) Stereausis: binaural processing without neural delays. J Acoust Soc Am 86:989–1006

Shera CA, Guinan JJ Jr, Oxenham AJ (2002) Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc Natl Acad Sci USA 99:3318–3323

Woolf NK, Ryan AF, Bone RC (1981) Neural phase-locking properties in the absence of cochlear outer hair cells. Hear Res 4:335–346