UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

VIABILITY OF COCHLEAR TRAVELLING WAVE SIGNAL PROCESSING FOR COCHLEAR IMPLANTS

by

Larry Schmidt

Submitted in partial fulfilment of the requirements for the degree
Master of Engineering (Bio-engineering)

in the

Department of Electrical, Electronic and Computer Engineering
Faculty of Engineering, Built Environment and Information Technology

UNIVERSITY OF PRETORIA

February 2016

SUMMARY

VIABILITY OF COCHLEAR TRAVELLING WAVE SIGNAL PROCESSING FOR COCHLEAR IMPLANTS

by

Larry Schmidt

Supervisor: Prof Johan Hanekom

Department: Electrical, Electronic and Computer Engineering

University: University of Pretoria

Degree: Master of Engineering (Bio-Engineering)

Keywords: Cochlear implant, auditory modelling, travelling wave, cochlear implant processing, pitch encoding

The travelling wave encodes acoustic information by stimulating the auditory nerve fibres. Understanding the travelling wave and its process is important for the development of cochlear implants speech processors. The development of a normal hearing auditory model, using a hydrodynamic model of the travelling wave to predict the nerve fibre spiking diagrams, marked the first stage of this study. This study then proceeded to look at the development of a travelling wave speech processing algorithm and model the electrical response due to the stimulation from the vocoder speech processor, and the travelling wave speech processor. The final stage was to predict whether temporal encoding occurred during cochlear implant stimulation for the vocoder speech processor and the travelling wave speech processor.

The results showed that the travelling wave normal hearing model was able to predict the nerve fibre characteristics seen in measurements from literature. This showed that the mechanical encoding performed by the travelling wave is vital to the encoding of information in auditory nerve fibres. The travelling wave speech processor was able to encode temporal cues for pitch up to 1060 Hz, where the results for the vocoder speech processor showed the 300 Hz limit seen in other literature of phase locking. Mimicking the travelling wave in cochlear implant speech processors may potentially benefit the delivery of information to the auditory cortex for cochlear implant users. However, these results must be legitimised using animal models and psychoacoustic experiments.

OPSOMMING

LEWENSVATBAARHEID VAN KOGLEÊRE-LOOPGOLF SEINVERWERKING VIR KOGLEÊRE INPLANTINGS

deur

Larry Schmidt

Studieleier: Prof J.J. Hanekom

Departement: Elektriese, Elektroniese en Rekenaaringenieurswese

Universiteit: Universiteit van Pretoria

Graad: Magister in Ingenieurswese (Bio-Ingenieurswese)

Sleutelwoorde: Kogleêre inplanting, ouditiewe modellering, loopgolf, kogleêre inplanting verwerking , toonhoogte enkodering

Die loopgolf enkodeer akoestiese inligting deur stimulasie van die gehoorsenuwees. 'n Begrip van die loopgolf is belangrik vir die ontwikkeling van kogleêre inplanting spraakverwerkers. Die eerste fase van hierdie studie was die ontwikkeling van 'n ouditiewe model van normale gehoor. Hierdie model maak gebruik van 'n hidrodinamiese model van die loopgolf om senuweevuurpatrone te voorspel. Verder het die studie die ontwikkeling van 'n loopgolf spraakverwerkingsalgoritme ondersoek, en het die elektriese respons in reaksie op stimulasie van vokoder- en loopgolfspraakverwerkers gemodelleer. Die finale fase was om te voorspel of temporale enkodering ontlok is tydens kogleêre inplanting stimulasie met die vokoder- en loopgolfspraakverwerkers.

Die resultate toon dat die loopgolfmodel vir normale gehoor in staat is om die senuwee-eienskappe van gepubliseerde metings te voorspel. Dit wys dat meganiese enkodering deur die loopgolf van kardinale belang is in die enkodering van inligting in die gehoorsenuwee. Die loopgolfspraakverwerker is in staat om temporale leidrade vir toonhoogte te enkodeer tot en met 1060 Hz, terwyl die vokoderspraakverwerker die 300 Hz limiet gedemonstreer het wat ook in ander literatuur oor fasesluiting gevind word. Om die loopgolf na te boots in kogleêre inplanting spraakverwerkers mag potensieel voordele inhou vir die aanbied van inligting aan die ouditiewe korteks van kogleêre inplantinggebruikers. Hierdie resultate moet wel bevestig word deur gebruik van diermodelle en psigo-akoestiese eksperimente.

LIST OF ABBREVIATIONS

AN	Auditory nerve
CF	Characteristic frequency
CI	Cochlear Implant
IHC	Inner hair cell
IHC-AN	Inner hair cell-auditory nerve
ISI	Inter-spike interval
LIN	Lateral inhibitory networks
ODE	Ordinary differential equation
PDE	Partial differential equation
PPS	Pulses per a second
SPL	Sound pressure level
TW	Travelling wave

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

This chapter examines the motivation for the research, the research questions, the approach to the study, and the contribution of the study to research in cochlear implants.

1.1 MOTIVATION FOR RESEARCH

The complex auditory system uses the mechanical and electrical system of the human body to convert sound pressure waves into nerve fibre spiking patterns which can be understood by the human brain to interpret sound. The study of these systems will allow for a greater understanding of how spatial-temporal information is transferred from the mechanical pre-processing to the electrical signals of the auditory nerve fibres. Present cochlear implant users struggle with (among others issues): perception of pitch, speech in noisy environments and high variability of speech recognition scores amongst cochlear implant users. Understanding the transfer of information to the auditory nerve fibres enable the development of more robust and superior electrical stimulation patterns to be used in cochlear implants.

1.2 CONTEXT OF THE PROBLEM

During normal hearing the malleus displacement, causes the basilar membrane of the cochlea to move. The displacement of the basilar membrane represents a travelling wave moving from the base of the cochlea to the apex. The movement of the basilar membrane is vital, as it is the mechanical processor to the auditory nerve fibres that encodes information to be interpreted by the auditory cortex. The encoding captured in the travelling wave is presumably richer than a simple place code, a notion that will be explored in this dissertation. This dissertation has a specific focus on travelling waves and how to encode them into electrical stimuli in cochlear implants. This section provides an overview of the relevant literature to illustrate the scientific context of the study on travelling waves and their role in the normal auditory system, how they are modelled in normal hearing auditory models, and the potential application in cochlear implant speech processors.

Auditory models can provide insight into the delivery of information from basilar membrane displacement to the auditory nerve fibres by analysing the auditory encoding in the mechanics of the ear (Carney, 1993). Popular auditory models use filter banks to

describe the mechanical filtering and the characteristics of the basilar membrane when stimulated by a sound pressure wave (Carney, 1993, Zilany and Bruce, 2006, Zilany and Carney, 2010). There are very few mechanical models able to describe the mechanical process using a biophysical response of the travelling wave (Shamma, Chadwick, Wilbur, Morrish and Rinzel, 1986, Eguia, Garcia and Romano, 2010).

Filter bank speech processing models typically describe the different processes observed in the travelling wave separately, including: the travelling wave delay (movement of the travelling wave from base to apex and the delay of the peak at a particular place relative to the start of the travelling wave at the base of the cochlear), basilar membrane displacement, the active process (gain of the basilar membrane displacement wave due to outer hair cell electro-motility (Kemp, 1979)), non-linearity, and the broadening of the travelling wave peak with increasing tone intensity. A model that describes travelling wave mechanics, often implemented as a hydrodynamic model of the fluid mechanics of the cochlea, usually includes all of these properties in a single unified model. This work uses a single unified model to describe the basilar membrane displacement in order to study if a single unified model can gain the output of the nerve fibres' spiking patterns measured in the literature.

This work then moves its focus to cochlear implants. The implementation of cochlear implants in profoundly deaf users has been relatively successful in restoring partial hearing but with great variability (Loizou, 1999). This can be seen in noisy environments where speech recognition deteriorates greatly. Speech recognition scores of 70% to 80% for sentences in quiet environments have been recorded (Stickney, Zeng, Litovsky and Assmann, 2004), but these scores deteriorate rapidly in noisy environments (Zeng and Galvin III, 1999) to an average of approximately 50% for a signal to noise ratio of 10 dB (Battmer, Reid and Lenarz, 1997). This work focuses on using the hydrodynamic model in a cochlear implant speech processing strategy and validates the possible improvement the strategy can have on speech recognition scores through the delivery of speech cues.

In using a hydrodynamic travelling wave model to implement a CI speech processor that transforms an input acoustic waveform into space-time firing patterns, the way in which the properties of a travelling wave affects the delivery of the encoded information to the auditory nerve fibres become evident. The use of decoding models to extract information contained in the auditory nerve fibres' spiking patterns facilitates a close investigation of the information contained in the resulting spike train patterns (Cariani and Delgutte, 1996a, Larsen, Cedolin and Delgutte, 2008).

The stimulation characteristics of the cochlear implants are vastly different to those of a normal hearing stimulation. The stimulation of the nerve fibres in normal hearing occurs through a travelling wave that moves along the basilar membrane. The deflection of the basilar membrane stimulates the hair cells, which in turn causes action potentials to occur in the auditory nerve fibres. The cochlear implant, however, most often uses a vocoder-type speech processor to electrically stimulate the surviving nerve fibres in a tonotopical arrangement along the length of the cochlea (Meddis and Hewitt, 1991a, Loizou, 2006). The cochlear implant relies heavily on place encoding for the delivery of information to the auditory nerve fibres. Place encoding is, however, not the only process that delivers auditory information to the nerve fibres in normal hearing. Processes such as temporal encoding, rate encoding, and phase encoding all carry information about the acoustic input signal. However, place encoding is used in most CI processors as the principle is simple, the signal processing is easily developed and it complements the hardware of the present generation of cochlear implants (i.e. more processing intensive processing would need better processors that use more power).

There are vast differences between the normal hearing stimulation of the auditory nerve and cochlear implant nerve fibre stimulation: the firing rates of electrical stimulation are higher, there is less deviation in the firing times, and the dynamic range is much smaller (Rattay, 1990).

The failure of the nerve fibre firing pattern to mimic the response in normal hearing nerve firing patterns while using cochlear implant speech processors, is the main cause of the deterioration of speech recognition in noise. The lack of speech recognition performance, and the vast differences in neural excitation patterns produced, indicates that an understanding of how normal hearing speech cues are delivered by the travelling wave (which is thought to contain rich frequency information) is required.

The delivery of speech information in vocoder speech processors occurs by means of place encoding, but these processes are unable to deliver temporal encoding due to limitations such as stimulation rate and stimulation patterns which do not include temporal encoding (Varsavsky, 2010b). Cochlear implants processing strategies modulate the pitch onto the rate of the stimulation pulse. However, the transfer of pitch is limited to 300 to 400 Hz (McDermott and McKay, 1997, Shepherd and Javel, 1997). The tonotopical placement of electrodes used for encoding of the acoustic signal, which is used in cochlear implant processors, does not account for the acoustic sources in a wide range of environmental

conditions (speech in noisy environments and music perception). The lack of acoustic sources include temporal encoding used in multi-tone pitch delivery (Greenberg, 1997). Therefore, the lack of acoustic sources is the possible limit for speech perception in noise, indicating the need for a more complete method of delivery of acoustic cues (including temporal encoding and place encoding, by the cochlear implants). A more thorough and holistic method involves making the cochlear implant robust to noise and increasing speech perception in non-ideal environments by including better temporal encoding in the cochlear implant speech processors. The travelling wave in normal hear, that may be implemented in cochlear implant speech processors, which captures both place and temporal information simultaneously is a more complete method of delivery of speech cues.

The delivery of spatiotemporal information by existing cochlear implants lacks the stimulation characteristics used by the travelling wave to deliver speech cues. This includes, but is not limited to, the travelling wave delay (which desynchronises the spectral components along the length of the cochlea) (Wojtczak, Beim, Micheyl and Oxenham, 2012), the latency between the different spectral components of the stimulation signal and the phase shift along the length of the cochlea (Greenberg, 1997). A vocoder distribution speech processor implements only one of these characteristics: the spatial spreading of spectral components of the acoustic signal along the length of the basilar membrane, known as place encoding (Loizou, 1999). It is hypothesised that an increase in the number of travelling wave characteristics used in cochlear implant speech processors will increase the robustness of cochlear implants in noise by observing nerve fibre firing diagrams which closely mimic those in cochlear implant speech processors (Varsavsky, 2010b).

In order to gain insight into the response of the auditory nerve fibres to different forms of stimulation and the information presented to the brain, this study calls for the development of auditory models. An accurate model is expected to have the ability to predict physical parameters (e.g. spike train patterns or BM displacement) that cannot be measured easily, and in this sense allows for an understanding of the processing of the system (i.e. in a sense, a high quality model provides a way of measuring "inside" a system in which physical measurements may be impractical or impossible). This present study set out to develop accurate models of the normal hearing auditory system and the electrically stimulated auditory system. While inputs to the former were multi-toned, missing fundamental signals or simple pure tones, inputs to the latter electrically stimulated model consisted of typical vocoder-type speech processor stimulation patterns, and a proposed

travelling wave speech processor electrical stimulation patterns. The model is used to study nerve fibre spike train patterns for these different types of inputs.

1.3 RESEARCH GAP

An improved understanding of the delivery of spatial-temporal information by cochlear implants is crucial, as it could potentially lead to robust cochlear implant processing strategies for CI users in noise. Existing cochlear implant processing strategies focus on the vocoder processing strategies that result in neural activation, which varies a great deal from normal hearing activation.

Normal hearing auditory models use filters to model the displacement of the basilar membrane (Carney, 1993, Zhang, Heinz, Bruce and Carney, 2001, Zilany and Bruce, 2006). Very little research has gone into implementing a nerve fibre auditory model using the travelling wave as the processing model for the basilar membrane displacement. Eguia et al. (2010) has contributed significant work to this field of using travelling waves to describe cochlear filtering in auditory model.

This research looks at plausibility of drawing processing techniques in cochlear implants closer to that of normal hearing and specifically, the plausibility of using a travelling wave processor (rather than a vocoder) as an encoder to arrive at electrical stimulation patterns. In the present work, this was considered for pure tones and missing fundamental tones. Pure tones were used as evidence that pitch can be encoded using the travelling wave while retaining place encoding in the processing by the use of peak displacement of the basilar membrane caused by the travelling wave. However, temporal processing was used for decoding of the pitch. The missing fundamental was then used to see if the pitch of multi-tones could be encoded without the requirement of place encoding. Very little research has been done in this area of implementation of the travelling characteristics in cochlear implants, and the possible benefits of the travelling wave characteristics to cochlear implant speech processing (Taft, Grayden and Burkitt, 2009, Du Preez, 2011, Harczos, Chilian and Husar, 2013).

There are many limitations in the cochlear implant (Loizou, Dorman and Tu, 1999, Loizou, 2006, Du Preez, 2011) that restrict the delivery of temporal and spatial information to the auditory nerve fibres. These limitations include (but are not limited to): the electrode resolution, pulse rate, nerve fibre survival, and electrodes array length. In order to counter these limitations, many different processing techniques have been developed: current

steering to increase the spatial resolution (limited by current spread), the addition of travelling wave delay to place encoding in order to introduce add phase encoding to the cochlear implant (Taft et al., 2009), and using rate encoding to encode pitch in the stimulus signal (limited by phase locking of nerve fibres to stimulation electrical pulse). Better stimulation spike patterns are required to gain the nerve fibre response which mimics normal hearing spike patterns.

1.4 HYPOTHESIS AND RESEARCH QUESTIONS

The main hypothesis of this study is that it should be possible to develop a travelling wave speech processor that will closely resemble the nerve fibre response seen in normal hearing. It is further hypothesized that such a travelling wave processor could potentially be a robust speech processor for the delivery of speech cues in noise. To investigate these hypotheses, the following specific research questions were considered.

- **Basilar membrane modelling for use in auditory nerve fibre firing prediction uses measured basilar membrane displacement to predict the basilar membrane filters. These filters are then used as the input for nerve fibre models. Does a more realistic model of the cochlear filtering, by using the travelling wave as a signal processing technique, allow for the accurate prediction of nerve fibre spiking patterns and cochlear filters?**
- **What stimulation pulse rate and number of electrodes (Du Preez, 2011) are required to find an optimal solution for implementing the cochlear travelling wave model in a cochlear implant?**
- **Can the implementation of a travelling wave processor for cochlear implant users potentially be used to improve pitch perception?**

1.5 APPROACH

The practical implementation of the study follows a modelling process in order to elevate the understanding of how spatial-temporal information is delivered to the human ear. The process became the basis for the design of a spatial-temporal processor which could be potentially used in cochlear implants to deliver temporal information to the auditory nerve fibres. During cochlear implant stimulation, nerve fibres lock onto the electrical stimulation pulses of the cochlear implant. In the case of the vocoder speech processor, the phase locking of the nerve fibres is related to the stimulus pulse rate of the cochlear implant. However, this does not deliver the pitch of the acoustic signal as it is not the

temporal nerve fibre firing rate relating the pitch of acoustic signal, but the temporal pitch (nerve fibre firing rate) relating to the pulse rate of the electrical stimulus. This study furthers the exploration of this concept.

The development of a human auditory model was done using a realistic hydrodynamic model of the travelling wave to fully understand the mechanical processing seen in the normal hearing ear. This could potentially allow for an understanding of the encoding of acoustic information in the travelling wave and therefore the delivery of acoustic cues to the auditory nerve fibres. Nerve fibre activation patterns were predicted using the travelling wave model as the signal processing input to the nerve fibre model. The results of nerve fibre data was then used as an input into decoding prediction models to predict the encoding of pitch information available in the nerve fibre spike train patterns. The pitch decoding model results was used to understand the mechanical pre-processing done by the normal cochlea for encoding of temporal information in the input acoustic signal.

Mechanical pre-processing, seen in normal hearing ears, was then replaced by pre-processing used in cochlear implant speech processors. A modelling approach facilitated this replacement, implementing models of the speech processor and electrical stimulation of nerve fibres. Different speech processors were tested: a typical vocoder-type processor, a processor that attempts to encode the entire travelling wave, and a processor that encodes only some aspects of the travelling wave. The results of the electrical nerve fibres model were then introduced to a pitch prediction decoding model to predict the temporal pitch that a cochlear implant user may perceive. The absence of temporal pitch in cochlear implant processing techniques, which focus on the delivery of place pitch, motivated this approach.

1.6 RESEARCH CONTRIBUTION

The main contributions of this work are:

- The development of a hydrodynamic travelling wave model of the cochlea basilar membrane displacement that captures travelling wave dynamics. This model was adapted from (Duifhuis, 2012), Duifhuis, Hoogstraten, Van Netten, Diependaal and Bialek (1987), and Duke and Jülicher (2003);
- The development of neural models that can predict neural activation from BM displacement. This model was adapted from models of: the inner hair cell model

from Mountain and Hubbard (1995), the AN-IHC model from Zilany, Bruce, Nelson and Carney (2009) and the auditory nerve model from Carney (1993);

- The implementation of neural models that can take electrical stimulation patterns as input. This model was adapted from Bruce, White, Irlicht, O'Leary, Dynes, Javel and Clark (1999b),Bruce, White, O'Leary, Dynes, Javel and Clark (1999a), and Smith (2011);
- A central auditory processing model – a phenomenological model that performs the single function of extracting temporal pitch from neural spike trains, adapted from the models of Cedolin and Delgutte (2005) and Cariani and Delgutte (1996a);
- Different variations of electrical stimulation input (vocoder processor, travelling wave processor, simplified travelling wave processor);
- Predictions of which processing has the most potential for conveying accurate temporal pitch information to CI listeners.

1.7 OVERVIEW OF STUDY

The next chapter motivates the proposed research by looking at the present literature in the field of travelling wave dynamics, auditory models, and cochlear implants. Chapter 3 describes the development of the auditory model, using the travelling wave model to describe the basilar membrane displacement. A prediction of the nerve fibres' spiking patterns occur and its results are compared to those measured in literature. Chapter 4 discusses the development of a travelling wave cochlear implant processor model used in conjunction with an electrical nerve fibre model to predict the nerve fibre spiking patterns for electrical stimulation hearing. The results from the travelling wave processor are then compared to the normal hearing auditory nerve fibre spiking patterns and the vocoder processor predictions.

The prediction by inspection used in Chapter 4 is then confirmed in Chapter 5 using a temporal pitch prediction algorithm to determine if pitch can be decoded from nerve fibre spiking patterns. Chapter 6 examines the findings from Chapters 3 to 5 and highlights important results. Chapter 7 reconsiders the research questions, takes a look into the contribution of the work in the research field, and reflects on how the work can be used in future research.

CHAPTER 2 LITERATURE STUDY

2.1 CHAPTER OBJECTIVES

This chapter is a review of the relevant literature for this study. The chapter introduces the normal hearing dynamics of the ear, in order to give the readers a brief overview of how the auditory systems encode sound information for interpretation.

Subsequently, the chapter provides an overview of the cochlear implant to discuss the functionality of the cochlear implant processor. There is a large difference between the processing of the normal hearing ear and the cochlear implant. The chapter illustrates a modelling approach for both normal hearing and cochlear implant nerve fibre predictions in order to gain an understanding of how pitch can be delivered to cochlear implant users.

2.2 NORMAL HEARING MECHANICS

2.2.1 Ear mechanics

The human ear consists of the outer ear, the middle ear, and the inner ear (cochlea). The outer ear consists of the pinna and the ear canal. This boosts the resonance to a maximum level of sounds in the mid to high frequency ranges in accordance with the angle between the acoustic source and the head. As a result, the origin source of the sound is identifiable (Shaw, 1974).

The middle ear comprises of the ear malleus, incus, and stapes which operate as a system of levers to overcome the impedance mismatch between the air and the fluids of the cochlea, as fluids in the cochlea would not create enough movement to produce an audible stimulation if the air stimulation from the eardrum was applied directly to the cochlea.

The inner ear (which is significant to cochlear implant research) consists of the vestibular apparatus and the cochlea. The cochlea consists of three sections, the scala vestibule, the scala media, and the scala tympani. These sections are important to the understanding of the normal stimulation for hearing as pressure from the round window, produced by the stapes, onto the fluid in the cochlea causes movement of the basilar membrane in the form of a travelling wave (Von Békésy, 1960). This causes the depolarisation of the hair cells which are connected to the basilar membrane, thus stimulating the auditory nerve fibres along the length of the basilar membrane (Von Békésy, 1960, Rattay, 1990).

The stimulation of the auditory nerve fibres by the hair cells causes these nerve fibres to fire, producing action potentials which are interpreted by the auditory system as sound. When a periodic stimulation signal is applied under normal hearing stimulation conditions, the entire basilar membrane will start to move from the base towards the apex with each section moving periodically according to the stimulus (Von Békésy, 1960), which produces a travelling wave with a phase lag relative to that of the stapes along the basilar membrane and a delay from the base to the apex which varies along the length of the cochlea.

2.2.2 Travelling wave

This section further explains the mechanics of the travelling wave as it is the mechanical pre-processing that causes the nerve fibres to be stimulated. The mechanics of the travelling wave is important to understand as it gives insight into the stimulus required in order to encode acoustic cues in nerve fibre spike train patterns.

The basilar membrane conducts a travelling wave, which starts at the base of the cochlea, finds its maximum at the point where the basilar membrane is tuned to a specific frequency and then decreases in amplitude towards the apex (Von Békésy, 1960, Rhode, 1971, Greenwood, 1990). The point at which the maximum deviation of the basilar membrane will occur is dependent on the acoustic stimulus frequency and the intensity of the pure tone, as the basilar membrane is tonotopically organised (in such a way that high frequencies are extracted at the base of the basilar membrane and low frequencies at the apex). This systematic relationship between peak displacement and cochlear location serves as a dynamic representation of spectral information of acoustic stimulation (Greenberg, 1997).

The travelling wave is rich in temporal information about the acoustic stimulation signal (Von Békésy, 1960), as it contains information about the space and temporal information of the stimulation signal. Neither the classical place nor temporal models alone, can account for many of the important properties of acoustic perception in the cochlea, such as the stability (the cochlea under normal hearing conditions is highly robust to noise) and apparent continuity of sound sources under a wide range of environmental conditions (Varsavsky, 2010a).

Spatiotemporal information is combined in various mechanisms within the travelling wave, including the travelling wave delay (which varies with frequency), tonotopical arrangement for the critical frequency with basilar membrane displacement, and the non-linear relationship of the basilar membrane displacement to the displacement at the malleus.

The tonotopic arrangement of the basilar membrane indicates that at a specific observation point, the deflection of the basilar membrane will be observed. The maximum deflection is achieved when the stimulation signal frequency correlates to the tuned frequency of the observation point. This shows that the basilar membrane acts as a filter along the frequency spectrum, by containing tuned bands that correlate to points along the basilar membrane. It was originally discovered that the travelling wave deflection along the basilar membrane had a linear relationship to the deflection of the malleus when measured *in vitro* (Von Békésy, 1960). However further measurements taken *in vivo* showed that the travelling wave deflection was non-linear in displacement as well as when measured *in vitro,* the non-linear distortion appeared to be absent (Rhode, 1971).

The travelling wave delay allows an estimation of the latency difference between the spectral components for the acoustic stimulus to be obtained (Greenberg, 1997), as the spectral components are tonotopically organised. This allows a frequency cue which can be decoded to be delivered to the auditory cortex.

The delay properties of the cochlear travelling wave also provide a potential means of organising spectral information across a dynamic range of sound pressure levels and acoustic environmental conditions. The differential delays associated with specific periodicities appear to be magnified in the primary auditory cortex. This is useful in the encoding of different sound pressure levels across the length of the cochlea and encoding spectral information using the delay across the cochlea (Greenberg, 1997).

The velocity of this travelling wave is rapid by the base of the cochlea, being nearly instantaneous (the latency from the stapes at the base is less than 5 ms with a velocity of 7 m/s to 24 m/s (Donaldson and Ruth, 1993)) for frequencies above 4kHz, but slowing dramatically for peak displacements in the apex. The transmission time for the total length of the basilar membrane requires approximately 10 ms to 14 ms (Greenberg, 1997),(Donaldson and Ruth, 1993).

Phase difference, caused by the travelling wave delay between the deflection of the basilar membrane and the motion of malleus, changes as the frequency at the observation point is increased towards the tuned stimulation frequency. The phase difference remains constant at approximately 1.6 radians for low frequencies and increases to approximately 28 radians where the observation point is stimulated with the tuned frequency (Von Békésy, 1960). The phase difference then decreases rapidly (similarly to the basilar membrane deflection) once the stimulation frequency increases beyond the optimal tuned frequency of the observation point.

The nerve fibres are sensitive to the phase of the stimulus signal which transfers pitch information to the auditory nerve fibres over a wide band along the length of the cochlea moving away from place pitch theories (Meddis and Hewitt, 1991a, Meddis and Hewitt, 1991b). The auditory nerve fibres are sensitive to the phase difference of the auditory stimulation over a wide tonotopical region (McGinley, Charles Liberman, Bal and Oertel, 2012), which will influence the firing pattern of these nerve fibres. The travelling wave is a processing mechanism which affects the nerve fibres. The reader will now gain insight into how the travelling wave acts as a pre-processor which influences the nerve fibre spiking patterns.

2.3 TRAVELLING WAVE AS A PRE-PROCESSOR

The travelling wave along the basilar membrane results from acoustic stimulation and provides spatiotemporal information to the auditory cortex. The spatiotemporal information is rich in acoustic cues able to encode pitch information.

As mentioned before, the basilar membrane acts as a filter that spatially spreads the acoustic stimuli into spectral components along the length of the tonotopically organised basilar membrane (Wojtczak et al., 2012). The frequency components are delayed along the length of the cochlea - the high frequency components have the smallest delay at the base of the cochlea and the low frequency components have the largest delay at the apex of the cochlea (Von Békésy, 1960). The delay is used to create a spatial spread of peak excitation sites relating to the turned frequency of the basilar membrane.

The increase in basilar membrane deflection toward the tuned frequency of the observation point illustrates how the basilar membrane acts as a filter to tonotopically arrange the frequency along the length of the cochlea. The basilar membrane deflection when the stimulation frequency is not at the tuned frequency of the observation point. These non

tonotopical deflections hints at the possibility that information about the acoustic signal is not only present at the peak deflection sites, but also at sites leading up to the observation point where temporal information can be presented to the nerve fibres by synchronous activity of nerve fibre firings

The spectral representation of a sound may represent a combination of highly synchronous activity distributed over the basilar membrane (Greenberg, 1997). Although the encoding of the travelling wave is not yet fully understood, it may be a pre-processing strategy for the stimulation of the auditory nerve fibres so that features pertaining to the pitch, intensity, and spectral information can be accurately and efficiently extracted, in order to stimulate the auditory nerve fibres for decoding by the auditory cortex.

Pitch perception of normal hearing lies in the coding by the auditory nerve fibres (Rattay, 1990). The change in neural firing rates, and the number of fibres active during stimulation represents an increase in auditory sound pressure stimuli (Berenstein, Vanpoucke, Mulder and Mens, 2010). The critical frequency bands heard by normal hearing are represented tonotopically, in a very fine densely packed spectral resolution across the length of the cochlea in normal hearing listeners. It may, therefore, be that the physiological basis of the critical bands may rest on the synchronicity of neural activity across a contiguous extent of a tonotopically organised map.

Using the travelling wave characteristics to mimic the travelling wave in a cochlear implant processor may have potential benefits, including higher speech recognition in noise due to the redundant delivery of the cues about the acoustic input contained in the travelling wave. Acoustic information is represented by the basilar membrane as a stable spectral representation (Taft, Grayden and Burkitt, 2010) that can be extracted from multiple points along the length of the cochlea. It cannot be extracted through place code, but through the various extraction methods presented by the travelling wave such as the temporal and phase coding present in the mechanics of the travelling wave.

The encoding of pitch will be discussed later in the chapter to illustrate the possible pitch encoding captured in the travelling wave and encoded for the auditory nerve fibres.

2.4 NERVE FIBRES

Nerve fibres, stimulated by the travelling wave, contain information of the acoustic signal encoded in neural spike trains by the travelling wave mechanics. The spike train patterns are then decoded later in the auditory cortex for interpretation. The nerve fibres under

normal hearing circumstances are encoded with acoustic information from the travelling wave mechanics which causes the displacement of the basilar membrane. The encoded information in the nerve fibres is explored below.

During pure tone acoustic stimulation, in which the travelling wave spatially spreads the acoustic signal into spectral components (Wojtczak et al., 2012), the nerve fibres in a normal hearing ear will fire synchronously with the stimulation frequency to which the nerve fibre population is tuned, making the synchronously firing nerve fibres a potentially reliable frequency cue as it can be decoded by the auditory cortex (Greenberg, 1997, Cedolin and Delgutte, 2010).

Synchronous firing of the nerve fibres is not affected by changes in sound pressure level. Only an increase in area of nerve fibre excitation along the length of the basilar membrane results from sound level changes. The width of the area of nerve fibre excitation will increase due to changes in sound pressure level. The limited band of synchronised activity, without change due to sound pressure level, defines the frequency information and the limits of the resolution range which contain the frequency cue for the tune stimulus (Greenberg, 1997).

Synchronised activity can be observed by inspection of an interspike interval (ISI) histogram. The interspike interval illustrates phase locking of individual nerve fibres to the stimulus period and multiple periods of the stimulus signal (Cariani and Delgutte, 1996a, Cedolin and Delgutte, 2005).

Information about the acoustic stimulus is delivered in a combination of synchronised (temporal encoding) and the non-synchronised activity (place encoding) of the nerve fibre firing patterns as a result of the stimulus. The synchronised activity delivers detailed information about the acoustic stimulus (by causing nerve fibres to phase lock to the auditory signal). This can be seen during pure tone stimulation (Dynes and Delgutte, 1992)) where the overall synchronised and non-synchronised activity delivers the energy level of the signal for interpretation (Greenberg, 1997, Duifhuis, 2012). However, this does not disregard the possibility for place cues also being present in the encoding of detailed acoustic information.

Auditory nerve fibres are tuned to drive biophysical characteristics (McGinley et al., 2012) to respond to stimuli with high temporal precision (Smith, Massie and Joris, 2005). These tuned nerve fibres are distributed along the length of the cochlea in a tonotopical

arrangement, with each frequency having an optimal position for stimulation. The cochlear nucleus (CN), however, has cells which are not tuned solely to a specific place of stimulation (Smith et al., 2005). The CN contains different cells, of which some are the octopus cells – named as such due to the striking morphology of the cells (Smith et al., 2005). The travelling wave of the basilar membrane is distributed along the length of the octopus cells, where the stimulus with the shortest travelling wave delay (the high frequency components) stimulates the distal tips and the longest travelling wave delay (the low frequency components) stimulates the proximal tips. These octopus cells each span different tonotopic lengths of the cochlea. The octopus cells will experience only a portion (approximately one fifth to two thirds) of the travelling wave stimulus. Moreover, the octopus cells found at the lower frequencies of the tonotopical region span the smallest ranges across the length of the cochlea (McGinley et al., 2012).

These cells are activated to fire most reliably during a stimulus sweep across the length of the octopus cell (McGinley et al., 2012), as can be seen in the excitatory postsynaptic potentials (EPSP) measurements. The EPSP have rapidly rising voltage times with the travelling wave delay as the input stimuli across the length of the octopus cell (McGinley et al., 2012). As sweep times vary from the natural travelling wave delay, the EPSP amplitudes decrease and simultaneously decrease the temporal precision of the neural fibres. Frequency components of the acoustic signal will be desynchronised in time along the length of the basilar membrane during travelling wave stimulation, and could potentially be resynchronised by the octopus cells (McGinley et al., 2012).

The octopus cell is one example of how the travelling wave presentation to the auditory nerve fibres may potentially enable the delivery of reliable cues to the CN. An exploration into the delivery of auditory cues using cochlear implants follows.

2.5 VOCODER-TYPE SPEECH PROCESSORS

This section looks at the present vocoder speech processors and processing strategies to encode speech information in order for the reader to understand how present cochlear implants convey pitch to the auditory nerve fibres and how it differs from the proposed strategy used in this work. Subsequently, the proposed strategy will be compared to that of the vocoder speech processor. Later chapters of this work will shed light on the proposed processing strategy.

2.5.1 Multichannel speech processors

This section discusses multichannel vocoder speech processors used in cochlear implants. This section highlights. It highlights the process of information delivery from cochlear implants to the nerve fibres of the cochlear implant user and the attempted improvements of speech perception performance of implants for implant users. Many advancements in cochlear implant speech processors and hardware have improved speech performance only slightly but were limited in the benefits as the improvements never gain performance of speech recognition seen in normal hearing listeners. This section ultimately suggests that cochlear implants need to include better processing techniques that rely on more than place cues for speech processing.

Multichannel cochlear implant strategies consist of multiple stimulation electrodes for the stimulation of the basilar membrane at multiple sites. The multichannel speech electrode array allows the exploitation of the frequency place mechanics of the basilar membrane (Loizou, 1999), through the use of the tonotopical arrangement of the cochlea. The electrode variation in the number of current cochlear implant electrode arrays ranges from 12 electrodes to 24 electrodes between different companies' brands. The use of multiple electrodes for stimulation potentially increases the number of information channels for an electrode array implant, which can range from 12 to 31 channels on a cochlear implant (Shapiro and Bradham, 2012). Speech recognition is shown to increase as the number of information channels increase (Loizou et al., 1999). It has been shown that sentence recognition requires at least 5 to 8 channels to reach asymptotic performance (Loizou et al., 1999). This was found in a quiet environment. However, Friesen, Shannon, Baskent and Wang (2001) showed that cochlear implant users don't use the number of channels available to them and speech recognition performance did not increase above 7 channels. This could be due to electrode interaction, stimulation rate, or spectral mapping. A closer look suggests a deterioration in performance by an average of around 50% for a signal-to-noise ratio of 10 dB (Battmer et al., 1997). Improvement in this performance has not been observed with newer cochlear implant devices (Spahr and Dorman, 2004). There are many reasons for the lack of improvement in speech perception with the improvement of electrode hardware and cochlear implant technology. These include, but are not limited to, electrode implant depth, frequency place mismatch, and channel interaction. The information that follows elaborates on each particular reason.

Electrode implant depth does not extend the entire length of the cochlea. Instead only short insertions of 7.2 mm to deep insertions of 21.6 mm (Baskent and Shannon, 2005), that do not reach the upper turn of the cochlea. The length of the electrode does not guarantee that the electrode has been placed at its maximum depth in the cochlea (Stöver, Issing, Graurock, Erfurt, ElBeltagy, Paasche and Lenarz, 2005). It is not possible to ensure that the electrode array has not caused trauma during implantation (such as perforating the basilar membrane) (Stöver et al., 2005) causing damage to the organ of corti affecting the tonotopical layout of the frequency map along the length of the cochlea.

Channel interaction distorts the acoustic information by spatially smearing the temporal information across the tonotopically arranged cochlea (Shannon, Cruz and Galvin, 2011). A continuous interleaved sampling approach (where electrodes are stimulated in a pulsatile manner) is used to reduce the distortion from channel interaction (Loizou, 1999). Electrode arrays (such as bipolar electrode array (Hartmann and Klinke, 1990) and tri-polar electrode array (Berenstein et al., 2010)) have been developed to minimise the current spread due to stimulation and to gain higher stimulation focusing accuracy (therefore, decreasing channel interaction and showing improvement in rate discrimination (Goldsworthy and Shannon, 2014)). Bipolar and tri-polar electrode arrays need to be stimulated with higher currents to gain the same perceived loudness as that of the bipolar electrode array when compared to the tri-polar electrode arrays, as the loudness perceived is a function of the number of activated nerve fibres. The higher stimulation currents required limits the benefits of these electrode array designs (Bipolar and tri-polar electrode arrays). Therefore, an elevated loudness level causes an increase in current spread and minimises the benefits of multiple electrode stimulation. However, using current steering has showed an improvement in speech recognition (Srinivasan, Padilla, Shannon and Landsberger, 2013) and current steering has shown an improvement in rate discrimination (Goldsworthy and Shannon, 2014).

High stimulation rates have been used to achieve improved temporal sampling of the acoustic signal and to increase the dynamic range (Shannon et al., 2011). The increase in stimulation rates did not, however, show significant improvement in speech recognition in quiet or noisy environmental conditions; an improvement that did occur varied across subjects (Shannon et al., 2011).

The increase in stimulation rate increased the dynamic range of electrical stimulation due to a lowering of the minimum threshold (Shannon et al., 2011). The trade-off enables the

higher stimulation rate in the increase in stimulation pulses, however, does not allow the neuron to fully recover as the stimulation rate is faster than the neuron refractory period (Plonsey and Barr, 2007). Increasing the stimulation rates, therefore, has slight benefit to the increasing in dynamic range of a cochlear implant user.

Many different approaches have been undertaken for delivering spectral and temporal information in the acoustic signal to the cochlea in speech processing strategies. The two methods proposed were a compressed-analogue approach and a continuous interleaved sampling approach (Loizou, 1999). The method for stimulating the cochlea using multiple electrodes is typically a continuous interleaved sampling approach, where electrodes are stimulated in a pulsatile approach to reduce channel interaction (Loizou, 1999).

These methods attempt to convey important information required for speech recognition (Blamey, Dowell, Clark and Seligman, 1987). The first speech processors conveyed the fundamental voice frequency (used to find the voice pitch of the acoustic signal), the second formant frequency, and the first formant amplitude (Loizou, 1999). Speech processors later expanded to include the first formant frequency and the second formant amplitude (Blamey et al., 1987).

The stimulation pulse period was, in these specific vocoder-type speech processors controlled by the fundamental voice frequency, used to deliver temporal information about the voice signal. The continuous interleaved sampling approach could be varied in numerous parameters. These include fundamental frequency detection, stimulation pulse rate, shape compression function, and filter spacing (Loizou, 2006).

These parameters are most often varied to be subject specific. Pulse rate, for example, shows a large variance in its benefits for different users. An increase in pulse rate would assist with the delivery of temporal information, while the filters applied, which relate to the number of information channels available, would influence the spectral resolution.

The stimulus pulse rate presented at an electrode results in higher and lower pitches perceived as pulse rate is increased and decreased for the cochlear implant user. Rate pitch encoding appears to work only for low stimulation rates of below 300 Hz (Loizou, 2006), although Venter and Hanekom (2014) achieved rate pitch perception that was very similar to normal hearing rate pitch perception when applying stimuli at the same rate on multiple electrodes. Their work indicates the possibility of improving the perception using rate pitch, which is explored later in this work.

The lynchpin of cochlear implant processors is the tonotopical arrangement of the cochlea (Greenberg, 1997). The effects which limit the present implant processors on the tonotopical areas of the cochlea include a limiting of the frequency spectrum (Shapiro and Bradham, 2012), trauma to the basilar membrane from the implant of the electrode array (Stöver et al., 2005), small dynamic range due to electrical stimulation (Shannon et al., 2011), and a lack of spatial resolution due to the number of electrodes and channel interaction (Bingabr, Espinoza-Varas and Loizou, 2008). The present cochlear implant processors have no means of overcoming this lack of encoding mechanisms, which are observed in normal hearing listeners, to deliver speech cues to the auditory nerve fibres. Presented and explored in this dissertation is a novel method of delivery of pitch and its potential use as a cochlear implant speech processor.

This section summarised the literature behind the multi-channel vocoder speech processor. Pointing out the successes and the downfalls of the processing strategy. Using the vocoder speech processor in the development of better processing strategies, which have a better delivery of speech cues, may require mimicking the biological system of the inner ear. There is current literature in which processing strategies have mimicked the travelling wave in order to develop processing strategies.

2.5.2 Processing strategies mimicking the travelling wave processing

The current vocoder-type processors do not use the full travelling wave characteristics to deliver fine structural information, such as travelling wave delays. However, some have tried to use the travelling wave delay as a processing technique (Taft et al., 2009). Taft et al. (2009) used the travelling wave delay as a period between sound arriving at the speech processor microphone to the time of the electrical stimulation pulse at the place specific location. This usage showed a minor improvement in speech perception. However, Taft et al. (2010) found that implementation of the travelling wave delay was just an improved processing technique and not the exploitation of the travelling wave mechanism.

Wang, Koickal, Hamilton, Cheung and Smith (2015) developed a CMOS analog cochlea filter which could be used with cochlear implants and had comparable responses to measured results in terms of basilar membrane gain and delay. Yang, Lyon and Drakakis (2015) developed an analog biomimetic cochlear implant processor filter bank architecture with across channels AGC (automatic gain control). The processor was able to predict the basilar membrane displacement in terms of gain, but characteristics such as delay and

phase are not compared. However, the results of the analog cochlea filter have not been tested using psychoacoustic tests or correlation of nerve fibre spiking patterns.

This work looks at fully integrating the travelling wave as a speech processor for cochlear implants using a software approach in the model of the electrically stimulated ear. This allows for insight into the potential benefits of a travelling wave speech processor.

2.6 PITCH

Pitch, in the auditory system, can be delivered by place and timing codes. The dominance of each encoding mechanism is still largely debated when placed in the context of other criteria, such as the importance and dominance of each mechanism in the auditory system (Oxenham, 2013) and the range of frequency to which encoding mechanism is used.

2.6.1 Place pitch

Place pitch relies on use of the tonotopical arrangement of the cochlea. The travelling wave uses mechanical filtering along the length of the basilar membrane during stimulation. The basilar membrane is tuned from the base to the apex of the cochlea (Von Békésy, 1960, Rhode, 1971, Greenwood, 1990). Stimulus of a pure tone will cause the vibration of the basilar membrane at a specific location, especially during low and medium sound levels. The position that vibrates due to a pure tone stimulus is called the characteristic frequency. This process is place frequency mapping and transfers pitch to the auditory system.

2.6.2 Temporal pitch

Temporal coding is a reliance on nerve fibres firing at a certain phase during a period of stimulation. The nerve fibres will phase lock with the period of the stimulus. The brain can then interpret the pitch of the stimulus by extracting the period of the signal through analysis of the time between the two successive spikes in the spike train patterns.

2.6.3 Trends of place and temporal pitch

The argument for the dominance with which mechanism the brain uses to extract pitch is widely contested in literature (Cariani and Delgutte, 1996a, McKay, McDermott and Carlyon, 2000, Carlyon, Deeks and McKay, 2010). In order to analyse the possible contribution of each mechanism in the encoding of pitch in the auditory system we will analyse the trends and the results found in literature.

Phase locking of nerve fibres occurs in frequencies up to 2 to 4 kHz in other mammals, strongly relating to the upper limit of phase locking of nerve fibres onto a stimuli (Rose, Brugge, Anderson and Hind, 1967). Frequency discrimination tests show the trend of frequency discrimination thresholds decreasing to the frequency of 8 kHz and subsequently staying consistent as the frequency increased. The trend shows a strong correlation to the temporal phase locking of nerve fibres to stimulus, which suggests that temporal pitch is dominate up to 8 kHz. Analysing the nerve fibre results from Cariani and Delgutte (1996a), which utilize the pooled interspike interval histograms to extract the temporal pitch, showed that temporal pitch can be easily extracted from frequencies up to approximately 1300 Hz. However, at a frequency beyond 1300 Hz, temporal pitch is harder to extract, implying that place pitch plays a role at a frequency much lower than 8000 Hz.

Transposed signals of stimulation that cause nerve fibre activation without relating to the place pitch of the stimulus, lead to worse pitch discrimination among listeners. These results suggest that place performs a role in the pitch delivery (Dreyer and Delgutte, 2006). The salience of a pure tone is almost always ranked higher on a pitch salience ranking (Fastl and Stoll, 1979). However, Cariani and Delgutte (1996a) found that when using temporal decoding to rank pitch salience, they found that the pitch salience was lower than that of complex stimuli containing the same pitch. Their findings again point to the idea that place code plays a role in delivery of pitch information.

Spatiotemporal encoding of pitch, which encodes pitch in both time and place along the length of the cochlea, could be an answer to the mechanism of encoding pitch in the auditory system (Cedolin and Delgutte, 2010). There is strong evidence in trends for the support of both place and temporal encoding in the auditory system. Improvement of the cochlear implant requires study of the use of both encoding mechanisms in the auditory system, leading to an understanding of the encoding mechanisms in the cochlear implant.

2.7 TRAVELLING WAVES IN COCHLEAR IMPLANTS

This section looks at the plausibility of mimicking the travelling wave in cochlear implants as well as questions why cochlear implants have poor speech performance and variability in different environmental conditions.

2.7.1 The importance of the delivery of acoustic cues to cochlear implant users

A limitation of the present place model, used by vocoder-type speech processors, is the lack of delivery of multiple speech cues in the representation of speech (Greenberg, 1997).

The auditory cortex possibly takes a Bayesian approach regarding the identification of the acoustic signal (Knill and Pouget, 2004). The tonotopical vocoder-type processor lacks this ability to deliver multiple speech cues to the auditory cortex and enable the Bayesian mechanism in the auditory cortex. In cochlear implants with vocoder-type speech processors, only place and rate encoding is presented to the auditory nerve fibres. However, this presentation differs vastly from the normal hearing presentation to the auditory nerve fibres.

The firing properties of the neurons in electrical cochlear implant vocoder-type speech processors have several differences compared to those of a normal hearing listener. These differing neuron firing properties can result in poor speech recognition in noise due to speech cues not being delivered by the tonotopical filter arrangement used in vocoder-type speech processors to the higher auditory structures. For example, speech recognition and the perception of music by cochlear implant users in a noisy environment are poor compared to normal hearing listeners (Battmer et al., 1997, Zeng and Galvin III, 1999, Stickney et al., 2004). The cochlear implant user requires a higher signal-to-noise ratio and more cues for speech in noise to experience the same performance as a normal hearing listener (Zeng, Nie, Stickney, Kong, Vongphoe, Bhargave, Wei and Cao, 2005, Loizou, 2006).

Cochlear implant users require more than simple pure tone stimuli and speech, such as music perception. Music perception occurs through the delivery of fine structural information of the acoustic signal (Eskridge, Galvin Iii, Aronoff, Li and Fu, 2012). Cochlear implant processors deliver only the envelope of the acoustic signal which does not contain its fine structural information (Loizou, 1999). The travelling wave could potentially delivery fine structural information to the auditory nerve fibres by making use of both place and time cues simultaneously.

The travelling wave probably contains speech cues which are not implemented in vocoder-type CI speech processors. Characteristics of the travelling wave include the travelling wave delay, the measurement of latency between the different spectral components of the stimulation signal, tonotopical desynchronisation (spread of frequency components not only in place, but also with delays between spectral components), and the phase shift along the length of the cochlea. These characteristics could potentially deliver the fine structural information required for music perception.

Present cochlear implants vocoder-type speech processors illustrate a loss of precision using place cues due to high current spread (therefore, limiting the accuracy of the place frequency encoding). The travelling wave sweep could lower the current required to gain an action potential across the area of the cell. The use of a travelling wave-based sweep in the electrical stimulation pattern to cause nerve fibre spiking patterns that activate the mechanisms that cause the octopus cell to fire with precision may decrease the high current used in cochlear implants to elicit an action potential. The travelling wave sweep would also increase the precision of place stimulation of the cochlear implant by limiting the current spread, requiring a lower stimulation current but achieving the same loudness level. The nerve fibre would fire due to the correct timing caused by the travelling wave stimulation. Using the travelling wave in cochlear implants offers to improve the delivery of pitch information to the auditory nerve fibre for cochlear implant users.

A premise of the present work is that the travelling wave could act as a pre-processor for a cochlear implant to optimally deliver speech cues, such as pitch, intensity and spectral information, to the neural fibres. The implementation of a travelling wave and an accurate neural model will add to the understanding of the mechanisms which underlie the delivery of speech cues to the auditory cortex, such as those represented by the octopus cells in the CN (assumedly involving the extracting of temporal pitch) (McGinley et al., 2012) and the delivery of information from which temporal pitch can be extracted by these cells.

A computer model of a present cochlear implant processing offers to help optimise the delivery of the speech cues in cochlear implants and aid the study of the limitations of present vocoder-based cochlear implant speech processors.

2.8 MODELLING

2.8.1 Modelling of normal hearing

The modelling of normal hearing engages the prediction of the nerve fibre spiking patterns from acoustic stimulus. The modelling approach predicts the basilar membrane mechanical processing, (basilar membrane displacement) and uses the processing to predict the IHC voltage used as an input to predict nerve fibre spiking patterns.

Modelling a normal hearing listener is most often done, using filters and/or signal processing, to describe the displacement of the basilar membrane (Moore, Peters and Glasberg, 1990, Carney, 1993, Zhang and Carney, 2005). The basilar membrane filters are designed using the measurements of basilar membrane displacement (Von Békésy, 1960,

Rhode, 1971). These filters that match the displacement of the basilar membrane are not realistic representations of the basilar membrane. The filter bank can accurately model a linear travelling wave. The signal processing filters are used to describe the basilar membrane displacement and further developed to describe the other travelling wave characteristics (Shamma et al., 1986, Carney, 1993, Zhang et al., 2001). However, Duifhuis (2012) shows that the cochlea cannot be modelled by second order or even forth order filters. This is due to the complex coupling of the fluid in the cochlea. The complex description of the travelling wave requires the understanding that the solution to the fluid movement in the cochlea duct is both a long wavelength (representing the interaction between the basilar membrane partition and the fluid), and a short wavelength (representing the compressibility of the fluid) solution.

Once the basilar membrane displacement has been modelled, the inner hair cell is modelled to relate the basilar membrane displacement to inner hair cell voltage (Hartmann, 2005). A synapse model is used as the link between the inner hair cell and the auditory nerve fibre.

The IHC-AN synapse adds adaptation to the nerve fibres firing seen in normal hearing nerve fibre measurements. An exponential model is used in most IHC-AN synapse models (Westerman and Smith, 1988, Zhang and Carney, 2005). The exponential model of the IHC-AN synapse uses two exponential time constants to describe the adaption response seen in auditory nerve synapses. The problems with the exponential model are that firing rate drops below the spontaneous rate of the nerve fibre at offset, the recovery of the auditory nerve synapse is not variable, and the auditory nerve synapse fails when stimulated with a long duration stimulus (Zilany et al., 2009).

To solve the problems of the exponential IHC-AN synapse model, a power-law adaptation model was developed by (Zilany et al., 2009, Zilany and Carney, 2010). Power-law adaptation parallels an adaptation of discharge rate that follows a fractional power of time or frequency response rather than an exponential decay (Chapman and Smith, 1963). The power-law function solves the problem of stimulus additive response where a second stimulus affects the nerve fibres is dependent on the first stimulus. It solves the problem at stimulus offset where the firing rate drops below the spontaneous rate of the nerve fibres' specified spontaneous rate. The auditory nerve synapse is scale-invariant which means that the recovery of the nerve fibres from stimulation is varied according to the stimulus duration.

The auditory nerve fibre can be modelled in various ways, such as the leaky integrate (Bialek, Rieke, De Ruyter Van Steveninck and Warland, 1998) and fire model or the entire complex complete compartmental model (Hodgkin and Huxley, 1952). These models are perhaps chosen based on their representation of the response of the nerve fibre and their computation time.

2.8.2 Modelling of cochlear implant users

In order to gain insight into the dynamics response of the nerve fibres due to electrical stimulation, modelling of the cochlear implant electrical auditory system occurs. This allows for insight into the workings of the cochlear implant and the nerve fibre response due to electrical stimulation.

Modelling the cochlear implant bypasses the normal hearing mechanical encoding and models the electrode stimulus current interface between the nerve fibres and the electrode current. Cochlear implant models assume the electrode position placement constant in the cochlea. However, placement varies in cochlear implant users and is subject specific.

The mapped stimulation pulses are mapped to the threshold-comfort current level (CL) units with a loudness growth function (LGF) (Nogueira, Büchner, Lenarz and Edler, 2005), compensating for the logarithmic increase in perceived loudness as a result of the increased stimulation current. The stimulation pulses' (mapped to threshold-comfort current level units) current spread is then predicted from the stimulation current output from the LGF. The spread of current is predicted, in order to see the flow of current to predict the stimulation of the individual nerve fibres along the length of the electrode array inside of the cochlea (Strydom and Hanekom, 2011).

Nerve fibre response was predicted by deterministic descriptions. However, stochastic activity is present in nerve fibre measurements and these models cannot predict the variance measured in nerve fibres which are electrical stimulation (Bruce et al., 1999b). Bruce et al. (1999b) illustrated that a stochastic single-pulse model better predicts a range of psychophysical measures of loudness, such as discharge probabilities, than a deterministic model (a poor model to describe the response of a nerve fibre). Subsequently a prediction of the nerve fibres' spiking patterns arises through the use of a stochastic model.

2.9 GAPS IN PRESENT LITERATURE

Modelling of an auditory system requires the prediction of the mechanical and nerve fibre filtering. Present models predict the mechanical response of the basilar membrane displacement using signal processing techniques, such as signal processing filters. This technique forsakes the biophysical characteristics of the mechanical response of the basilar membrane displacement. Using filtering characteristics does not fully allow for study into the effects of basilar membrane displacement on encoding of signal processing as filter banks connot accurately represent the complex fluid coupling in the cochlea, as will be discussed below, or accurately represent the basilar membrane displacement, which is important for understanding the role of the travelling wave in the auditory system.

Cochlear implants use the vocoder-type speech processing strategies to deliver information to nerve fibres in the auditory system that rely on place encoding. This leaves opportunity for investigation into the use of temporal encoding in speech processing strategies in cochlear implants, as the cochlear implant speech processor enables the control over timing of pulses which could mimic the temporal encoding of the travelling wave observed in normal hearing listeners.

Cochlear implants can only deliver low frequency pitch information to cochlear implant users. Studying the possibility of delivery of temporal information using realistic normal hearing processing in cochlear implants users may allow the delivery of a greater range of pitch information in cochlear implants by encoding both spatial and temporal information.

CHAPTER 3 IMPLEMENTATION OF A TRAVELLING WAVE MODEL FOR NORMAL HEARING

3.1 CHAPTER OBJECTIVES

The previous chapter laid the foundation for an understanding of how the human auditory system works. This chapter considers the development of a computational model for normal hearing listeners, which includes outer ear, middle ear, and inner ear models. The results from the normal hearing auditory model will be discussed by evaluating basilar membrane displacement and nerve fibres spiking patterns resulting from the stimulation by pure tones at the input of the model.

Gaining insight into auditory nerve fibre spiking patterns, without the need for invasive measurement methods such as animal models, requires a normal auditory model. The normal hearing model developed in this chapter will facilitate study of the role of the travelling wave in auditory perception. The model-predicted nerve fibre patterns will be used in later chapters as input to pitch perception models, and to evaluate proposed CI processing implementations.

Modelling of normal hearing listener is often done by representing cochlear signal processing by a bank of filters. From a cochlea mechanics viewpoint, these represent the displacement of the basilar membrane (Moore et al., 1990, Carney, 1993, Zhang and Carney, 2005), specifically the BM movement envelope, which has a peak excitation place that corresponds to the input acoustic frequency described by a tonotopical map. Basilar membrane filters (Giguere and Woodland, 1993) are typically designed using measurements of basilar membrane displacement (Von Békésy, 1960, Rhode, 1971). The filters that describe the displacement of the basilar membrane are not realistic representations of the basilar membrane motion. The BM conducts a travelling wave, and while a filter bank can accurately model some characteristics of a linear travelling wave, Duifhuis (2012) showed that cochlea signal processing cannot be modelled by second order or even forth order filters. This is due to the complex fluid mechanics of the cochlea. The complexity of the hydrodynamics of the travelling wave requires an understanding that the solution to the differential equations that describe the fluid movement in the cochlea

duct is both a long wavelength solution (that represents the interaction between the Organ of Corti partition and the cochlea fluid), and a short wavelength solution (representing the compressibility of the fluid).

The model presented in this dissertation differs from many existing models of normal hearing that are used to understand cochlea signal processing, primarily in the use of a travelling wave model to describe the basilar membrane displacement instead of using signal filters. A travelling wave model provides a more realistic representation of the signal processing in normal hearing model. It will be shown that this is a valid way to describe cochlea filtering and to gain insight into nerve fibre spiking patterns. It is necessary to show that a travelling wave model of cochlea processing in normal hearing provides cochlea filtering because this model will be applied in following chapters to consider the processing in cochlear implants. The specific focus of this dissertation is to explore travelling wave-based CI speech processing.

The travelling wave model in this work is a model from Duifhuis et al. (1987) and later updated in Duifhuis (2012), authenticated through correspondence to human Otoacoustic data and not basilar membrane displacement measurements. This travelling wave model is not updated in this study, but is compared to literature in which basilar membrane displacement was measured. However, the basilar membrane displacement data is not taken from human measurements. Therefore, no model parameters were adjusted to match this basilar membrane displacement data, as the model parameters of the Duifhuis et al. (1987) model, derived from available human otoacoustic data, is expected to be more accurate than parameters that match the model to cat or chinchilla data. Later on, this study will illustrate that an adjustment of the model parameters to match the measured basilar membrane displacement from a chinchilla (Ruggero, Rich, Recio, Narayan and Robles, 1997).

3.2 METHODS

This section describes the normal hearing model used in this dissertation to predict nerve fibre spike patterns on the auditory nerve. The various parts or processing steps of the model are derived from existing models, shown as different sections or processing blocks in Figure 3.1. An active, non-linear travelling wave model (third block) is used to obtain the nerve fibre spiking patterns, although travelling wave models are not commonly used in these types of models as explained in the chapter objectives section above.

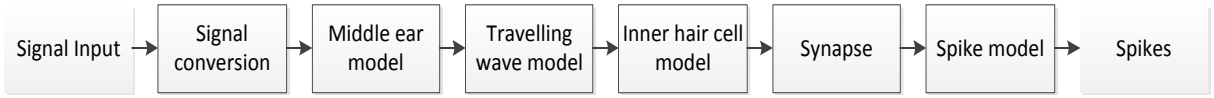

Figure 3.1. Normal hearing model used to generate the nerve fibre spike trains.

The normal hearing model consists of: a signal conversion stage that relates the sound pressure level (in dB SPL) input to sound pressure (in Pascal units) at the tympanic membrane, a middle ear model as the interface that matches the air impedance with the impedance of the fluid in the cochlea (O'Connor and Puria, 2008, Duifhuis, 2012), a travelling wave model which predicts the basilar membrane displacement (Duke and Jülicher, 2003, Duifhuis, 2012), an inner hair cell model which relates basilar membrane displacement to a membrane potential (Mountain and Hubbard, 1995), a synapse model which provides neural adaptation characteristics (nerve fibre threshold changes due to a stimulus of a period of time) to the model (Zilany et al., 2009), and a neural spiking model which generates nerve fibre spike train patterns (Carney, 1993). These model sections are explained in detail below.

The middle ear model is required to gain the correct basilar membrane response from the travelling wave model used. Duifhuis et al. (1987) modelled the travelling wave without correct middle ear coupling, resulting in the travelling wave failing to reflect the harmonic distortion products of the stimulus (Diependaal and Viergever, 1983).

The travelling wave model simulates the natural basilar membrane response of the transmission of the acoustic signal from the middle ear to the hair cells (Von Békésy, 1960), allowing for the evaluation of travelling wave properties effects on information transfer to the auditory nerve fibres.

3.2.1 Signal conversion

The input signal was converted from SPL (in dB) to linear pressure units (in Pascal) at the tympanic membrane using the following equation (3.1) (Duifhuis, 2012):

$$Pressure\ Level = p_0 10^{\frac{level}{20}}, \tag{3.1}$$

where the P*ressure level* is the pressure in Pascal at the tympanic membrane, p_0 is the reference pressure of 0.2 μN/m^2, and *level* the sound level at the tympanic membrane in dB SPL. The pressure level was then used as the input into the middle ear model.

3.2.2 Middle ear model

The middle ear of the auditory system is the interface between the acoustic environment and the cochlea (Duifhuis, 2012). This section describes the middle ear model used. This particular middle ear model was chosen for simplicity of implementation (as it has been previously used with the travelling wave model implemented below), correct middle ear coupling between the acoustic environment and the cochlea, and because this particular middle ear model was matched to measure otoacoustic data in Duifhuis (2012).

The middle ear was modified by Duifhuis (2012) to match human data. The middle ear matches the air pressure level with the impedance of the fluid in the cochlea. Figure 3.2 shows the compartmental representation of the middle ear model used (Duifhuis, 2012).

The middle ear stiffness (S_{me}), mass (m_{me}), and acoustic impedance of air (Z_a) are all described by a transmission line network. The physical representation parameters for the middle ear model in Figure 3.2, are given in Table 5.2 below, taken from Duifhuis (2012). Figure 3.2 describes the transformation of the pressure level from equation 3.1 to the pressure at the stapes of the cochlea using a crude bandpass filter to represent the middle ear. The middle ear can represent the bandpass filter as the hearing range is from 30 Hz to around 20 kHz. However, this crude bandpass filter does not accurately represent the filter cut-off characteristics seen in normal middle ear operation. It does, however, accurately represent the frequency spectrum which we are concerned: from 0.7 kHz to 5.7 kHz.

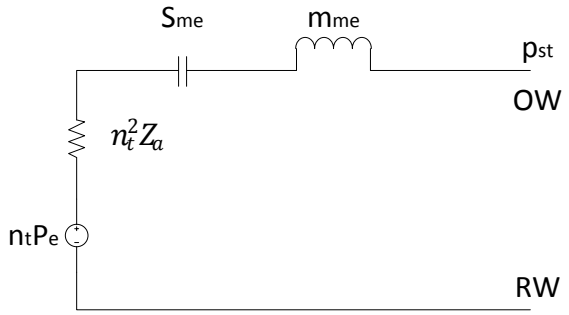

Figure 3.2. Simple middle ear model. The middle ear stiffness (S_{me}), mass (m_{me}), acoustic impedance of air (Z_a), pressure level (P_e), Pressure at the stapes (P_{st}), and the cochlear coupling constant (n_t) are shown in the figure. Parameter values are provided in table 5.2. This model provides a bandpass filter, centered around 2 kHz, with cut-offs at 0.7 and 5.7 kHz, and relates the acoustic properties of air at the ear canal entrance to those at the entrance of the cochlea (round window (RW) and oval window (OW)).

The parameters used to obtain the values substituted into equation 3.20 and 3.22 (the equations used to solve the travelling wave model which are described later in the chapter under the travelling wave model) are solved using the stapes stiffness, mass, and damping from the middle ear model (Duifhuis, 2012). Damping (d_{st}) is described by (3.2):

$$d_{st} = n_t^2 Z_a A_{fp}, \tag{3.2}$$

where n_t is the cochlea coupling constant(which accounts for the coupling between the external ear air pressure and the cochlea fluid pressure), A_{fp} is the stapes footplate area, Z_a is the acoustic impedance of air. The stapes mass is described by (3.3):

$$m_{st} = \frac{Q_{3dB,me} d_{st}}{2\pi f_{c,me}}, \tag{3.3}$$

where $f_{c,me}$ is the centre frequency, and $Q_{3dB,me}$ is the quality factor of the middle ear (indicated with subscript *me*). The stiffness of the stapes is described by (3.4):

$$s_{st} = m_{st} \times \left(2\pi f_{c,me}\right)^2. \tag{3.4}$$

The variable (A) used in the matrix of equation 3.22 is given by equation 3.7, which is solved using equations 3.5 and 3.6 using the holding variable γ_0 and A_{sq}:

$$\gamma_0 = \frac{m\left(\left(\frac{m_{sT}}{A_{fp}}\right)+n_t^2 m_{me}\right)}{BM_{width}\times dx}, \tag{3.5}$$

$$A_{sq} = dx\frac{2\rho BM_{width}}{mA_{fp}}, \tag{3.6}$$

where ρ is the fluid density, *m* is the basilar membrane mass, *dx* is change in distance of the x axis (which has a resolution of 0.043 mm steps) and BM_{width} the basilar membrane width. The variable A, used in the matrix of equation 3.22 to describe the middle ear response for the basilar membrane displacement in the travelling wave model, is:

$$A = 1 + \gamma_0 A_{sq} \tag{3.7}$$

where γ_0 is given from equation 3.5 and A_{sq} is from equation 3.6.

3.2.3 Travelling wave model

This section describes the travelling wave model implemented in this study. This travelling wave model describes the basilar membrane deflection using a one-dimensional, non-linear partial differential equation. The model also describes both long and short wavelengths of the cochlea fluid motion, representing the basilar membrane displacement (Duifhuis, 2012). The characteristics that this particular travelling wave model displays are the travelling wave delay from the base of the cochlea to the apex, frequency decomposition along the length of the cochlea where high frequencies are extracted at the base and low frequencies are extracted at the apex, active gain of the basilar membrane with low intensity acoustic stimulation, and broadening of the basilar membrane displacement with an increase in the acoustic stimulation intensity.

The motivation for using the Duifhuis et al. (1987) model, the same one dimensional, active nonlinear model of the basilar membrane model from Duifhuis (2012), is to gain understanding of the effects of the active and nonlinear parameters of cochlea mechanics when one presents the basilar membrane deflection to the nerve fibres of a plausible nerve fibre model. The travelling wave model implemented in this work (Duifhuis, 2012) represents the travelling wave characteristics, such as cochlea nonlinearities, the travelling wave delay, basilar membrane displacement, and active gain of the basilar membrane (Von Békésy, 1960, Rhode, 1971, Rhode, Roth and Recio-Spinoso, 2010). The model's validation requires a comparison of the modelled characteristics obtained from the travelling wave model to measured results from literature later in the chapter.

The partial differential equation (PDE), describing the displacement of the basilar membrane (eq 3.8), is solved by converting the PDE into two coupled ordinary differential equations, (ODEs) which can be solved using numerical methods, as discussed below.

$$\frac{d^2\left(m\frac{d^2w(x,t)}{dt^2}+d_{damping}(x)\frac{\mathrm{dw(x,t)}}{dt}+s(x)w(x,t)\right)}{dx^2}-\frac{2\rho}{h}\frac{d^2w(t)}{dt^2}=0 \qquad (3.8)$$

where m is the mass of the basilar membrane (kg/m^2), $d_{damping}$ is the damping of the basilar membrane, s the stiffness of the basilar membrane, x is the distance along the length of the basilar membrane, w is the basilar membrane displacement at x, ρ is the fluid density for all the canals, and h is the height of the height of the scala vestibule duct. The height of the cochlea does not change in a one-dimensional model.

Figure 3.3 shows a portion of the travelling wave model that was discretised into n compartments from the base of the cochlea to the apex. The value of n is dependent on computational power available; each compartment is described by the mass (m), damping ($d_{damping}$), and stiffness (s) of that individual compartment. The present implementation of the travelling wave model used 800 compartments along the length of the cochlea. This was sufficient to ensure stability of the model during simulation and to minimise simulation computation time. U_i represents basilar membrane displacement, P_i represent trans-partition pressure at i, OW is the oval window and RW is the round window.

The travelling wave model (eq 3.8) was solved using numerical methods as shown below. Figure 3.3 is a representation of how each discretised step is described by a unique parameter set.

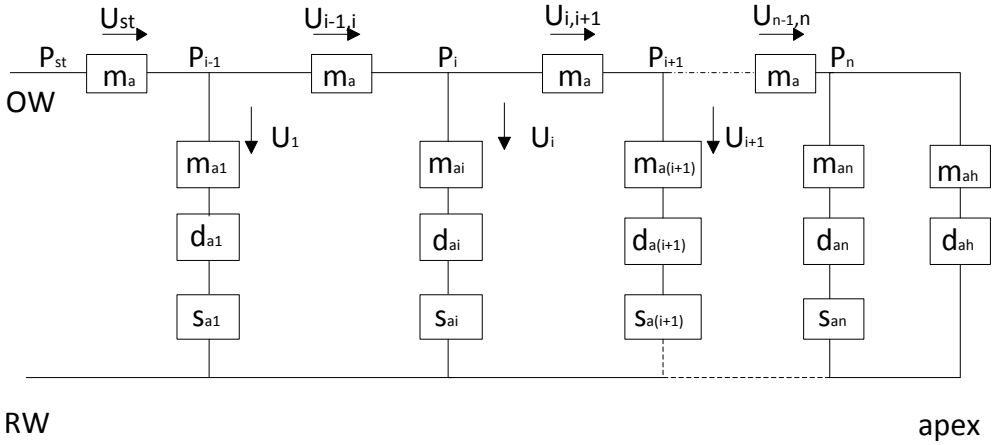

Figure 3.3. Compartmental representation of the travelling wave model, reconstructed from Duifhuis (2012). The compartment is a visual discretisation of the sections along the length of the basilar membrane, where m_a is the fluid mass coupling as each section is equally spaced, U_i represents trans-partition volume velocity of the basilar membrane, P_i is the pressure of each section, m_{ai} is the basilar membrane mass, d_{ai} is the basilar membrane damping, and s_{ai} is the basilar membrane stiffness of each section. h represents the helicotrema.

The chosen parameters for the mass, damping, and stiffness will affect the linearity, and active or passive mechanics of the travelling wave model. Mass, damping, and stiffness are altered using mathematical descriptions of each compartment. Duifhuis (2012) described the damping of each compartment d_{ai} mathematically by a van der Pol oscillator. The van der Pol oscillator was then critically damped; by changing the model to a Hopf-profile travelling model (Duke and Jülicher, 2003). A critically damped model (Hopf-bifurcation model) will not oscillate randomly when no stimulation is applied to a section of the basilar membrane. The van der Pol oscillator would cause random oscillations when no stimulus is applied, causing nerve fibres to fire at higher spontaneous rates than the nerve fibres found in the human auditory system. These higher spontaneous rates could be incorrectly perceived by the auditory system as an applied stimulus.

The mass of each compartment for the Duifhuis (2012) model is described by a constant 0.5 kg/m^2. The damping factor is described by a van der Pol oscillator, equation 3.9. The van der Pol equation is a function of velocity of the particular section of the basilar membrane and is used as feedback loop which controls the gain observed in measurements for different intensities of stimulation. This Hopf-bifurcation model accurately describes

basilar membrane displacement and the travelling wave characteristics seen during stimulation (Duifhuis et al., 1987, Duifhuis, 2012), and therefore this study utilizes a Hopf-bifurcation model.

The Hopf-bifurcation model, which has the same equation as the van der Pol oscillator, describes the damping (Pa.s/m) factor for equation 3.8, $d_{damping}$, in the form:

$$d_{damping}(x) = -d_1(x) + d_2 \frac{\mathrm{dw(x,t)}^2}{dt} \tag{3.9}$$

where $\frac{\mathrm{dw(x,t)}}{dt}$ represents the average basilar membrane partition velocity, and x the distance along the basilar membrane measuring from the base of the cochlea. d_1 is given by:

$$d_1(x) = \varepsilon\sqrt{ms(x)}, \tag{3.10}$$

where ε=0.05 is a crucial van der Pol parameter which determines the active gain of the basilar membrane response to input stimuli, m is the mass of 0.5 kg/m^2 and s(x) is the membrane stiffness. If ε is sufficiently small then the limit cycle is very close to a sinusoid, with a frequency close to that of the undamped oscillator. If ε > 1, then the limit cycles change shape dramatically and the oscillators become relaxation oscillators, and if Ɛ is 0 then the model is critically damped. The stiffness (Pa/m) of the basilar membrane is described by:

$$s(x) = 2\pi(s_0 10^{-(\gamma x)} - apex\ correction\ constant), \tag{3.11}$$

where s_0 is the Greenwood (1990) upper limit frequency for human frequency-position predictions at the base of the cochlea (20682 Hz), γ the stiffness decay constant (60 m^{-1}), and x is the distance along the length of the basilar membrane. The s_0 constant is multiplied by 1/Hz to keep the unit of the stiffness to Pa/m. d_2 is given by:

$$d_2 = (\alpha)^2, \tag{3.12}$$

where α is a dimensionality constant with the value of 5×10^{12}.

Implementation of the van der Pol oscillator on each section will see each discretised section along the length of the basilar membrane oscillate at its limit cycle velocity (natural resonant frequency) (equation 3.18). ε affects the limit cycle velocity of the section. A lower ε value will see a lower limit cycle velocity and an increase in gain observed when the BM is stimulated.

The model of the van der Pol oscillator is modified to use the Hopf-bifurcation profile (Duke and Jülicher, 2003), which is a van der Pol oscillator that has been critically damped. The Hopf- bifurcation profile is scaled (by change in ε) so that small deflections have a limit cycle value of exactly zero during no stimulation. The limit cycle velocity (μm/s) can be calculated by d_2/d_1. This limits the value of the random oscillations and keeps them to a minimum displacement amplitude (Duifhuis, 2012). The model version of the Hopf-bifurcation model, first implemented by Duke and Jülicher (2003), was modified by Duifhuis (2012) to obtain reasonable middle ear impedance matching (by correlating to otoacoustic emissions) between the cochlea and the stimulus, and proper apical termination at the final section of the basilar membrane model.

3.2.4 Method for solving the travelling wave model

The differential equation 3.8 of describing basilar membrane displacement was solved by a Gauss elimination method to solve the spatial differential equations and then a fourth-order Runge-Kutta method was used to solve the PDEs which represent the model in Figure 3.3 above for the displacement and velocity of the basilar membrane (Diependaal, Duifhuis, Hoogstraten and Viergever, 1987, Duifhuis, 2012). The procedure is described below. Diependaal et al. (1987) found the fourth-order Runge-Kutta method to be superior in stability and efficiency to other numerical methods, such as Heun and Sielecki, for solving the one-dimensional cochlea model in the time domain.

Equation of 3.8, which describes basilar membrane displacement, appears in a shorthand format by introducing variable ϕ and g (equations below), describing the basilar membrane damping and stiffness,

$$\varphi(x,t) = m\frac{d^2w(x,t)}{dt^2} + d_{damping}(x)\frac{dw(x,t)}{dt} + s(x)w(x,t), \tag{3.13}$$

and from equation 3.13 g is shown below,

$$g(x,t) = d_{damping}(x)\frac{dw(x,t)}{dt} + s(x)w(x,t)\,, \tag{3.14}$$

The difference between equation 3.13 and 3.14 gives

$$m\frac{d^2w(x,t)}{dt^2} = \varphi(x,t) - g(x,t), \tag{3.15}$$

this allows equation 3.8 to be rewritten as:

$$\frac{d\varphi(x,t)}{dx^2} - K\varphi(x,t) = -Kg(x,t), \tag{3.16}$$

where *K* combines the other parameters of equation 3.8 and is described by:

$$K = \frac{2\rho}{hm}. \tag{3.17}$$

Introduction of the velocity v which is equal to $\dot{w}$ allows rewriting equation 3.15 as:

$$\dot{v} = \frac{\varphi - g}{m}, \tag{3.18}$$

because

$$v = \dot{w}\ , \tag{3.19}$$

where

$$g = d(x)v + s(x)w\ , \tag{3.20}$$

equation 3.18 and 3.19 are now solvable using a forth order Runga-Kutta method and equation 3.16 is solvable using Gauss elimination procedure. Equation 3.13 is solved by transforming it into a spatial differential equation:

$$\ddot{\varphi}(i) = \frac{1}{\Delta^2}\big(\varphi(i-1) - 2\varphi(i) + \varphi(i+1)\big), \tag{3.21}$$

where Δ is the spatial discretisation step size of equal length along the length of the basilar membrane. Describing the ϕ for each section along the length of the basilar membrane in a vector and using equation 3.16. Equation 3.16 is multiplied by Δ^2, K x Δ = β, and replacing -2 – β by K gives:

$$\left(\begin{bmatrix} A & 1 & 0 & & 0 & 0 & 0 \\ 1 & K & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & K & & 0 & 0 & 0 \\ & \vdots & & \ddots & & \vdots & \\ 0 & 0 & 0 & & K & 1 & 0 \\ 0 & 0 & 0 & \cdots & 1 & K & 1 \\ 0 & 0 & 0 & & 0 & Y & Z \end{bmatrix}\right) \varphi = -\beta g + \begin{pmatrix} S \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ A_p \end{pmatrix}, \tag{3.22}$$

where Y, Z, A_p and β are described by:

$$Y = 1, \tag{3.23}$$

$$Z = -\left[1 + \frac{m_a^n}{m_{a_n}} + \frac{m_a^n}{m_{a_h}}\right], \tag{3.24}$$

$$\beta_n = \frac{m_a^n}{m_{a_n}}, \tag{3.25}$$

$$A_p = -\frac{m_a^n}{m_{a_n}} d_{a_h} V_h, \tag{3.26}$$

d_{a_h} is the damping factor at section of the apex (h) of the basilar membrane, V_h is the velocity at apex, n is the last discretised section of the basilar membrane, and m_a^i and m_{a_i} are described by:

$$m_a^i = \frac{2\rho\Delta}{bh}, \tag{3.27}$$

$$m_{a_i} = \frac{m(x)}{b\Delta}, \tag{3.28}$$

where m(x) is the mass (kg/m^2) of the basilar membrane section at the distance x along the length of the basilar membrane, h is the height of the basilar membrane, and i is reference of the discretised section.

In these equations, the deflection of the basilar membrane is dependent on the adjacent basilar membrane step. However the end sections, base and apex, are only influenced by

one section and require termination described by s and A_p. Using the Gaussian elimination rule, to solve for the solution (g) of matrix 3.22 can be found at each time step (Diependaal et al., 1987, Duifhuis, 2012). The value of g is used in equation 3.18 to solve for the basilar membrane velocity.

A 4th order Runga–Kutta method may be used to solve the PDEs of equation 3.18 and 3.19(Diependaal et al., 1987), which solves for the displacement of the basilar membrane:

$$w^{(1)} = w(t) + \frac{1}{2}\Delta t v(t)\ , \tag{3.29}$$

$$v^{(1)} = v(t) + \frac{1}{2}\Delta t \dot{v}[t, w(t), v(t)], \tag{3.30}$$

$$w^{(2)} = w(t) + \frac{1}{2}\Delta t v^{(1)}, \tag{3.31}$$

$$v^{(2)} = v(t) + \frac{1}{2}\Delta t \dot{v}\left[t + \frac{1}{2}\Delta t, w^{(1)}, v^{(1)}\right], \tag{3.32}$$

$$u^{(3)} = u(t) + \frac{1}{2}\Delta t v^{(2)}, \tag{3.33}$$

$$v^{(3)} = v(t) + \frac{1}{2}\Delta t \dot{v}\left[t + \frac{1}{2}\Delta t, w^{(2)}, v^{(2)}\right], \tag{3.34}$$

gives the final values of the displacement (u) and velocity (v) of the basilar membrane:

$$u(t + \Delta t) = u(t) + \left(\frac{\Delta t}{6}\right)\{v(t) + 2v^{(1)} + 2v^{(2)} + v^{(3)}\}, \tag{3.35}$$

$$v(t + \Delta t) = v(t) + \left(\frac{\Delta t}{6}\right)\left\{\dot{v}[t, w(t), v(t) + 2\dot{v}(t + \frac{1}{2}\Delta t, w^{(1)}, v^{(1)}) + 2\dot{v}[t + \frac{1}{2}\Delta t, w^{(2)}, v^{(2)}] + \dot{v}[t + \frac{1}{2}\Delta t, w^{(3)}, v^{(3)}]\right\}, \tag{3.36}$$

where Δt is the time step. In the results predicted, shown below (Section 3.3), the time step is 2.5 μs.

3.2.5 Inner hair cell model

The basilar membrane displacement, produced by the travelling wave hydrodynamic model of Duifhuis (2012) modified to match the Duke and Jülicher (2003) variation of the travelling wave model, was related to the inner hair cell voltage to model the information processing performed by the hair cells (Mountain and Hubbard, 1995). The ionic current flow through the hair cell apex is almost entirely caused by tension-gated channels, which open and close due to mechanical hair-bundle displacement (Mountain and Hubbard, 1995).

The inner hair cell (IHC) transducer is assumed to respond instantaneously to the mechanical input from the BM. The hair cell's potential is then solved for by using the mechanical displacement of the basilar membrane as input into the IHC model, which used 800 IHC's equally spaced of the length of the basilar membrane (35 mm). The relationship between the IHC voltage and the mechanical displacement is a non-linear function.

The model from Mountain and Hubbard (1995) uses a two stage process, relating the basilar membrane displacement to the transduction voltage of the hair cell and then processing it through a low pass filter. This process describes the nonlinearity of the IHC voltage, the output of the IHC voltage is from the low pass filter. The relationship between the IHC voltage (V_{IHC}) and basilar membrane displacement at a specific section (related through the transducer voltage (V_T)) can be described by a low pass filter with the following equation:

$$V_{IHC}[i,n] = K_A(V_T[i,n] + V_T[i,n-1]) + K_B V_{IHC}[i,n-1], \tag{3.31}$$

where V_{IHC} is the IHC voltage which describes the nonlinearity of the transduction voltage, i is the indicator of the IHC and position, n is the time sample, V_T is the inner hair cell transduction process voltage described by equation (3.33). Equation (3.33) describes the transduction process from the mechanical basilar membrane displacement to electrical signals, and K_A and K_B are filtering constants described by:

$$K_A = \frac{1}{2\tau_{IHC}F_S+1} \quad \text{, and} \quad K_B = \frac{2\tau_{IHC}F_S-1}{2\tau_{IHC}F_S+1}, \tag{3.32}$$

where Fs is the sampling rate used in the model, τ_{IHC} is the time constant of the IHC. The transducer voltage (V_T) from equation 3.31, which is the result of the inner hair cell transduction process converting the basilar membrane displacement to voltage, is described by a parabolic function where μ is the displacement of the basilar membrane and μ_0 is the lower limit of the basilar membrane displacement (19 μm) which will cause hair cell voltage to follow equation 3.33:

$$V_T[\mathrm{i,n}] = V_0 \frac{(u-u_0)^2}{(u-u_0)^2+K_S} - V_1 \quad u > u_0, \tag{3.33}$$

$$V_T[\mathrm{i,n}] = -V_1 \quad u < u_0, \tag{3.34}$$

where V_0 (14 mV), V_1 (1.8 mV), and K_S (1600 μm^2) are constants. Response of the IHC voltage to displacement of the basilar membrane is illustrated in Figure 3.4.

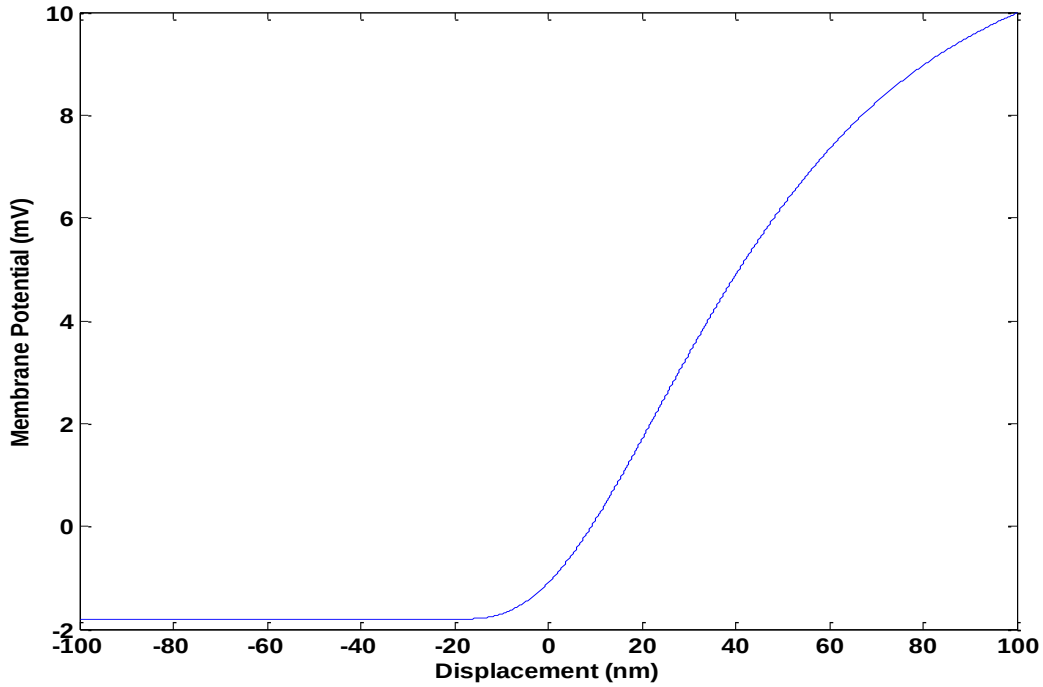

Figure 3.4. The results show the IHC membrane voltage as a function of the basilar membrane displacement as modelled. This reproduces the model predictions in Mountain and Hubbard (1995). See figure 4.7A from Mountain and Hubbard (1995) for a comparison between model predictions and IHC measurements.

3.2.6 Inner hair cell AN synapse

An exponential model is used in many inner hair cell-auditory nerve (IHC-AN) synapse models (Westerman and Smith, 1988, Zhang and Carney, 2005). The exponential model uses two exponential time constants to describe the adaptation response seen in auditory nerve synapses. The exponential model can be problematic because the firing rate drops below the spontaneous rate at offset, the recovery time of the auditory nerve synapse is not variable, and the auditory nerve synapse fails when stimulated with a long duration stimulus (Zilany et al., 2009).

The IHC AN synapse model used in the present study is that of Zilany et al. (2009), and was implemented using a power-law dependence on time followed with a short term (fast) exponential decay, to describe the adaptation components seen in measured data (Zilany et al., 2009). The model from Zilany et al. (2009) can be found at the authors website (https://www.urmc.rochester.edu/labs/Carney-Lab/publications/auditory-models.cfm)

Power-law adaptation is characterized by an adaptation of discharge rate that follows a fractional power of time or frequency response rather than an exponential decay (Chapman and Smith, 1963). The power-law function solves the problem that the exponential model has with stimulus additive response, where a second stimulus' effect on the nerve fibres is dependent on that of a first stimulus. It also solves problems at stimulus offset (where the firing rate of the nerve fibre will drop below the nerve fibre's specified spontaneous rate at stimulus offset), and the problem of the auditory nerve synapse not being scale-invariant, meaning that the recovery of the nerve fibres from stimulation varies according to the stimulus duration. The power-law function implementation of the synapses allows for predictions from modelled nerve fibres to match nerve fibre measurements.

3.2.7 AN nerve model

The auditory nerve model of Zilany et al. (2009) was used in the present study, this model can be found at the authors' website (https://www.urmc.rochester.edu/labs/Carney-Lab/publications/auditory-models.cfm). This model is based on the Carney (1993) spike generator. Carney (1993) implemented a spike generator which modelled both refractory and spontaneous rates of nerve fibres. Modelling the spontaneous rate of nerve fibres allowed for implementation of high, medium, and low spontaneous rate fibres.

The discharge generator is a Poisson discharge generator. The modelled nerve fibres have absolute refractory period of 0.75 ms, and discharge history of 40 to 50 ms which takes into account the previous response of the AN. The spontaneous rates that were used are shown in the Table 1. Random firing is influenced by the spontaneous rate. Discharge probability increases as the spontaneous rate increases, causing more random nerve fibre firings. The spontaneous rate is controlled by the synapse voltage, which increases the discharge probability.

Table 1. The spontaneous rates of nerve fibres used in the AN nerve model, used in the present study.

Spontaneous rate	Spikes/second
High	100
Medium	5
Low	0.1

The ability to model different nerve fibres of different spontaneous rates allows the modelling of different combinations of nerve fibres along the length of the cochlea, which could result in a greater understanding of how different nerve fibres encode information contained in the basilar membrane displacement. Nerve fibres are equally distributed along the length of the cochlea in the model. A total of 800 nerve fibres were chosen for the simulations shown in this chapter. The cochlea length used was 35 mm, which translates to a nerve fibre spacing of 43.75 µm.

3.2.8 Model summary

In summary, the normal hearing model consists of the following set of models from literature: the middle ear model (O'Connor and Puria, 2008) (which was modified to fit human data by Duifhuis (2012)), the travelling wave model by Duifhuis et al. (1987) (which was modified to be critically damped (Duke and Jülicher, 2003)), the inner hair cell model of Mountain and Hubbard (1995), the synapse model from Zilany et al. (2009) (which was modified to accept inputs from the inner hair voltage), and the nerve fibre model from Carney (1993).

The input to the model used in this study are sinusoidal waveforms (frequencies from 50 Hz to 10 kHz was used) and multi-tone missing fundamental signals with a sampling frequency of 400 kHz. The output is a space-time pattern of spike trains predicted for a

given input. The sampling frequency of 400 kHz is used for the travelling wave model, as lower sampling frequency causes the model to become unstable due to the non-linear nature of the travelling wave model, which requires smaller step sizes to be solved. Pure tones are used as input to compare the model to relevant literature and compare the response of the model to frequency inputs. However, the model is valid for complex signals as well. Figure 3.1 and again in Figure 3.5 shows a block diagram of the complete model. The sections that follow consider model predictions and compare those with available data.

Table 2 below lists the parameters used in the various sections of the model.

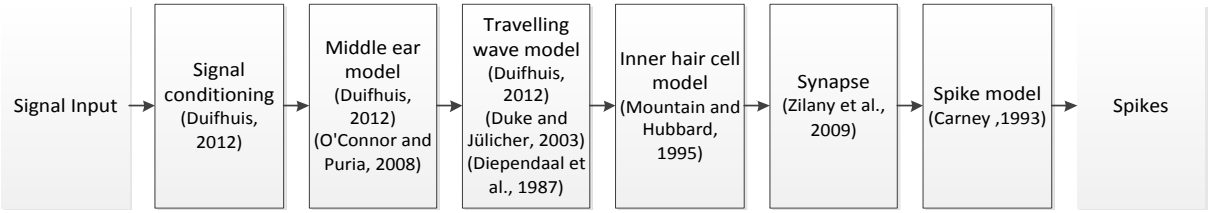

Figure 3.5. Figure 3.1 repeated with references relating to the models used to create the normal hearing model.

Table 2. List of parameters for normal hearing auditory model and the sources of the parameters. These parameters are chosen to best match human data in the respective articles.

Parameter name	Value	Origin
Signal Conversion		
Reference Pressure(ρ_0)	0.2 μN/m^2	(Duifhuis, 2012)
Middle ear Parameters		
Acoustic impedance of air (Z_a)	6.92 x 10^6 Pa s/m^3	(O'Connor and Puria, 2008)
Cochlea coupling (n_t)	30	(Duifhuis, 2012)
Middle ear mass (m_{me})	401 x 10^3 kg/m^4	(Duifhuis, 2012)
Middle ear stiffness (S_{me})	3.13 x 10^{13} N/m^5	(Duifhuis, 2012)
Quality factor of the middle ear ($Q_{3db,me}$)	0.4	(Duifhuis, 2012)
Stapes footplate area (A_{fp})	3.16 x 10^{-6} m^2	(O'Connor and Puria, 2008, Duifhuis, 2012)

Basilar membrane Parameters		
Mass (m)	0.5 kg/m^2	Duifhuis (2012)
upper limit frequency (S_0)	20682 Hz	Duifhuis (2012) which is related from the Greenwood (1990) map and modified to fit to human data.
Stiffness decay constant (γ)	60 m^{-1}	(Duifhuis, 2012)
Damping (d2)	5×10^{12}	
Ɛ	0.05	(Duifhuis et al., 1987)
Time Step (Δt)	2.5 µs	
Height (h)	1 mm	(Duifhuis, 2012)
Cochlea Length(x)	35 mm	
Apex correction constant	140.59	(Duifhuis, 2012)
IHC Parameters		
Time constant (τ_{IHC})	1 ms	(Mountain and Hubbard, 1995)
Scale factor constant (V_o)	14 mV	(Mountain and Hubbard, 1995)
Offset constant (V_1)	1.8 mV	(Mountain and Hubbard, 1995)
Displacement-offset constant (x_0)	-19 µm	(Mountain and Hubbard, 1995)
Saturation constant (K_S)	1600 µm^2	(Mountain and Hubbard, 1995)
Nerve fibres		
Number of nerve fibres	800	
Nerve fibre spacing	0.04375 mm	

3.3 RESULTS: MODEL PREDICTIONS

This section will highlight predictions from the normal hearing model and will compare these to literature. Predictions of the nonlinear active travelling wave model are compared to displacement measurements results from literature (Rhode, 1971). To evaluate whether nerve fibres spike train predictions are reasonable, the following nerve fibre characteristics

will be compared to data: spike train patterns, inter-spike interval (ISI) histograms produced for pure tone stimulation, nerve fibre synchronisation to pure tone stimulation, and nerve fibre tuning curves. These are displayed in the sections that follow.

3.3.1 Temporal response properties

Figure 3.6 below shows the predicted responses from the different stages of the model from the normal hearing model. The stages are: the stimulus input, the basilar membrane displacement from the travelling wave model, inner hair cell voltage, synapse output, and the PST histogram from the spike generator. A pure tone stimulus signal of 500 Hz with a 50 dB SPL intensity level was used as simulation input. The basilar membrane displacement is measured at position 20.96 mm, which correlates to a characteristic frequency (CF) of 1000 Hz, was used in Figure 3.6 to obtain these predictions. Figure 3.6 is used to illustrate the stages of output of the model of a single nerve fibre. The nerve fibres are spaced equally along the length of the cochlea (35 mm), and 800 nerve fibres are used for the simulations in the results section.

Figure 3.6 below shows the output of the model, Figure 3.1, at the output of different model stages. These results are compared below to measured data from literature and confirms that the model reacts correctly when compared to data which is done in the results section below.

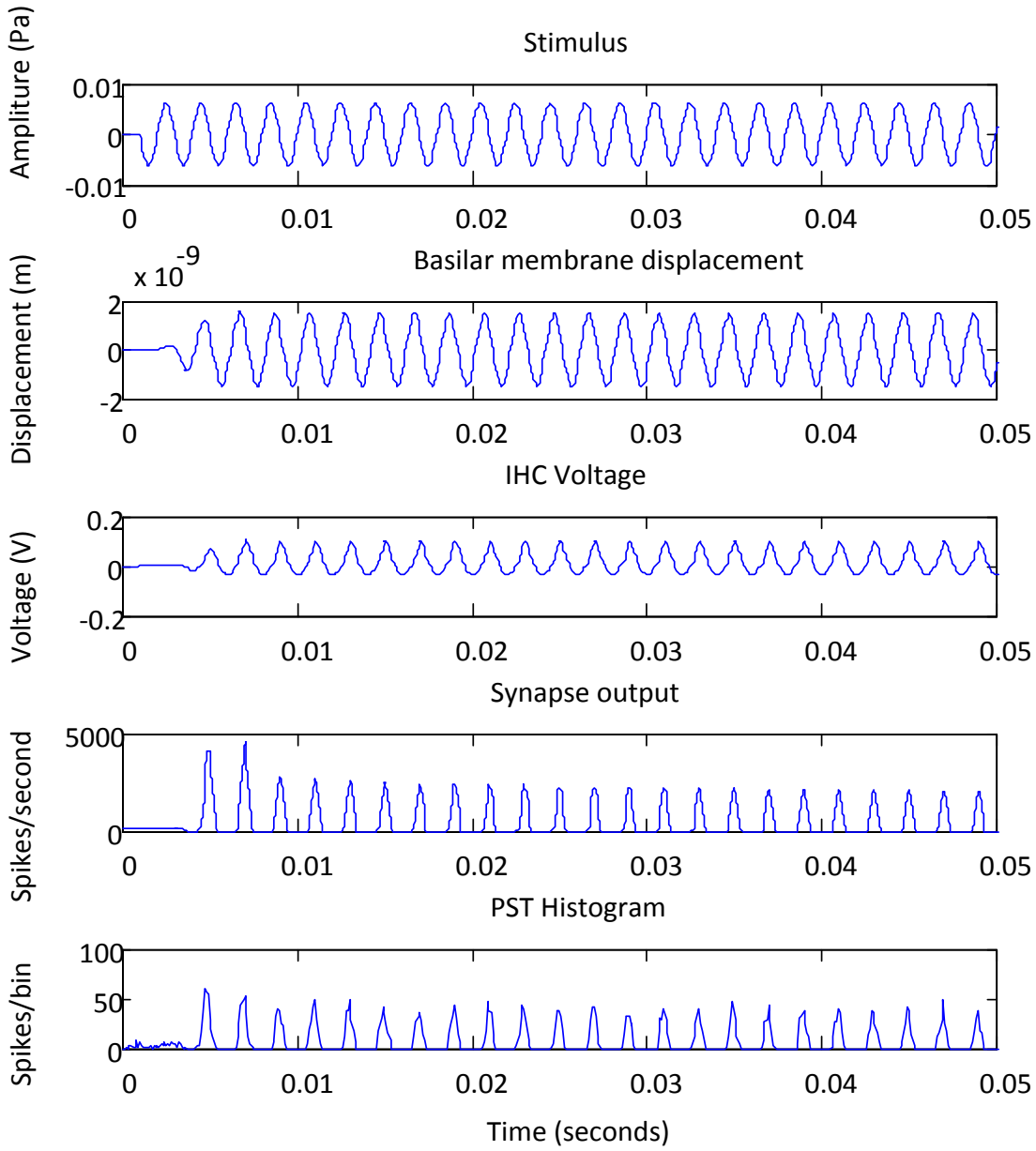

Figure 3.6. Response stages of the model for a pure tone input of 500 Hz with a stimulation loudness level of 50 dB SPL. The first stage is the stimulus signal input into the middle ear model at the tympanic membrane. The second panel is the basilar membrane displacement of a single point on the basilar membrane. The single nerve fibre used for the spike generation predictions had a CF of 1000 Hz, estimated from the human Greenwood map. The PST histogram contains the response from 400 simulation iterations for this particular nerve fibre, that were added together to form the PST histogram seen here. The figure resolution above is down sampled to be used here, as the lower resolution figure is easier to handle and a high resolution figure is not required.

3.3.2 Predicted spatial response properties: Cochlea travelling wave

When simulating not just a single cochlea position, but multiple positions, travelling wave behaviour is observed in the hydrodynamic travelling wave model. Figure 3.7 below shows a 3D representation of the basilar membrane and the displacement predicted by the hydrodynamic model over a 10 ms interval with an assumed BM length of 35mm. Mechanical stimulation is at the input to the middle ear model and in this figure, excitation at the basal end of the basilar membrane (0 mm) is conducted towards the apex of the cochlea (35 mm).

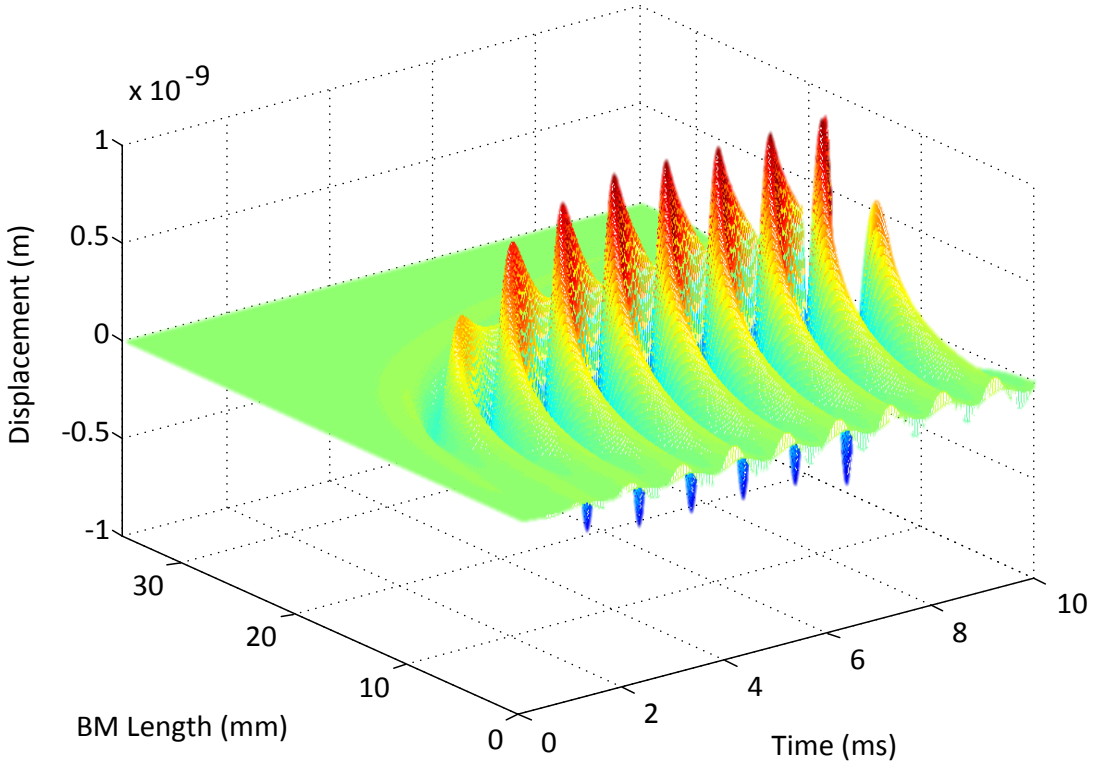

Figure 3.7. Travelling wave representation for a pure tone stimulation of 1000 Hz over 10 ms.

The point of maximum deflections, 20.96 mm, for pure tone stimulation is closely related to a tonotopical map along the length of the basilar membrane (Greenwood, 1990). The maximum deflection shifts towards the base as the intensity level increases, this is shown in Figure 3.12 and discussed below. Figure 3.8 illustrates documented frequency versus place maps for pure tone stimulation using a pulse stimulation of 0.01 ms as input to the model. This indicates where the maximum deflection of the basilar membrane may be

expected to occur. The model data is compared to the Greenwood (1990) map (which is mapped for the human species) and other travelling wave models (Zwislocki, 1950, Neely, 1981). The point of maximum deflection varies with sound pressure level when using a pure tone, and this is illustrated in Figure 3.12.

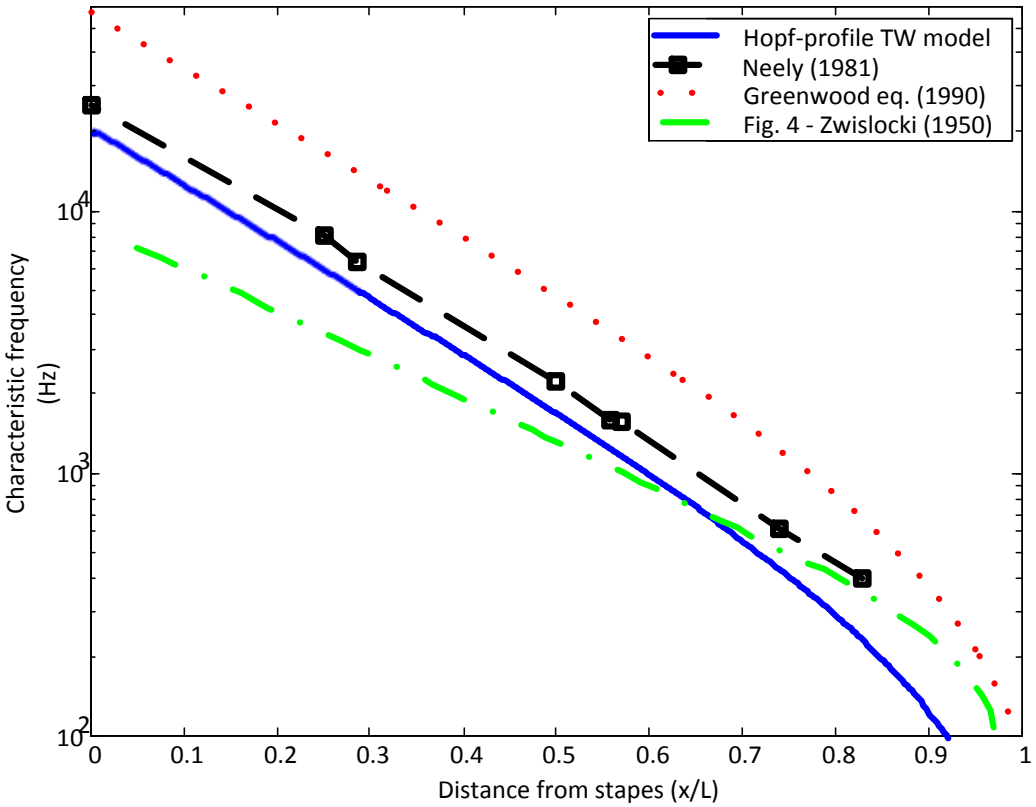

Figure 3.8. Frequency versus place map predicted by the present Hopf-profile model compared to data in literature. The figure shows the cochlea position at which the travelling wave peaks for a particular frequency pure tone input. L is the total length of the cochlea and x is the position along the cochlea.

The travelling wave has a delay from the base of the cochlea to the peak excitation point along the length of the basilar membrane. This delay is a function of the frequency of the pure tone stimulus, as shown in Figure 3.9. The modelled data shows a larger delay at the apex (low frequencies) compared to measured data, and has a smaller delay at the base of the cochlea (high frequencies). The measured data, taken from Ruggero and Temchin (2007), was constructed from multiple measurement sources (Von Békésy, 1960, Gundersen, Skarstein and Sikkeland, 1978, Stenfelt, Puria, Hato and Goode, 2003). The travelling delay in the model is smaller than those of the measured human data at higher frequencies and follow a simple 1/f model at higher frequencies, but does not increase as

steeply as the 1/f model or the human data. The model predictions for travelling wave delay has an offset of around 1 ms compared to the human data. It may be possible to add a 1ms delay to the model to bring the model in-line with the human measured data. This technique would work for a single period stimulus but a multiple period stimulus would require knowledge of the end of a stimulus period and where the next period would start in order to apply the delay. This may be due to the middle ear, a better model of the middle ear may be required. This study does not use the 1/f model of the delay as this study is interested in more than just the travelling wave delay but also the other travelling wave characteristics. For this study, the travelling wave model is able to predict the travelling wave delay with sufficient accuracy as it mimics the trends of the travelling wave.

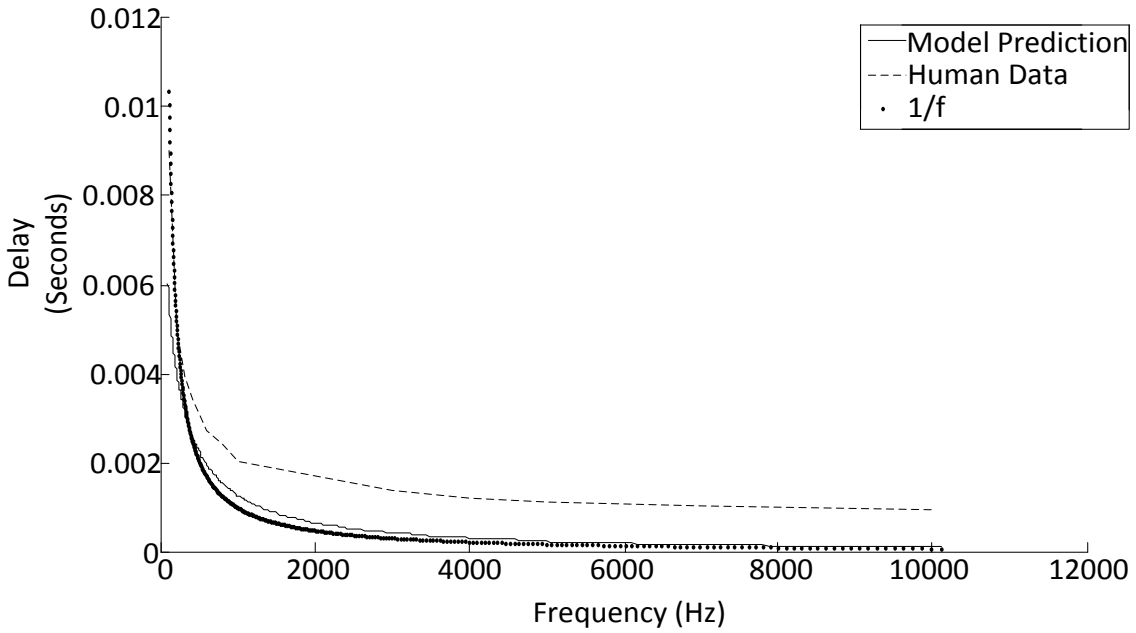

Figure 3.9. Basilar membrane peak displacement delay as a function of the frequency of the pure tone stimulus. Human measurement data produced from Ruggero and Temchin (2007) which was created from measurement results from (Von Békésy, 1960, Gundersen et al., 1978, Stenfelt et al., 2003). The maximum delay for the model is 6 ms at 100 Hz, while the measured human delay is 9 ms, and the 1/f model is 10.3 ms at 100 Hz.

Figure 3.10, Figure 3.11 and Figure 3.12 show the basilar membrane displacement relative to the malleus displacement for multiple intensity levels (5 dB to 80 dB) of a pure tone stimulation. These figures represent the basilar membrane gain (the basilar membrane displacement in relation to the malleus displacement). For a lower loudness level, the stimulation gain in Figure 3.11 is around 45 dB at the peak of excitation, resulting from the basilar membrane acting as an amplifier. This gain level decreases as the level of the pure

tone stimulation increases, as seen in Figure 3.10, Figure 3.11, and Figure 3.12. The point of maximum gain (displacement) shifts with an increase in loudness level, moving away from the predicted tonotopical place according to the Greenwood map shown in Figure 3.8. The point of maximum deflection shifts towards the base of the cochlea with an increase in stimulation intensity level for a pure tone stimulation. This characteristic is seen in measured data as well as in the present model.

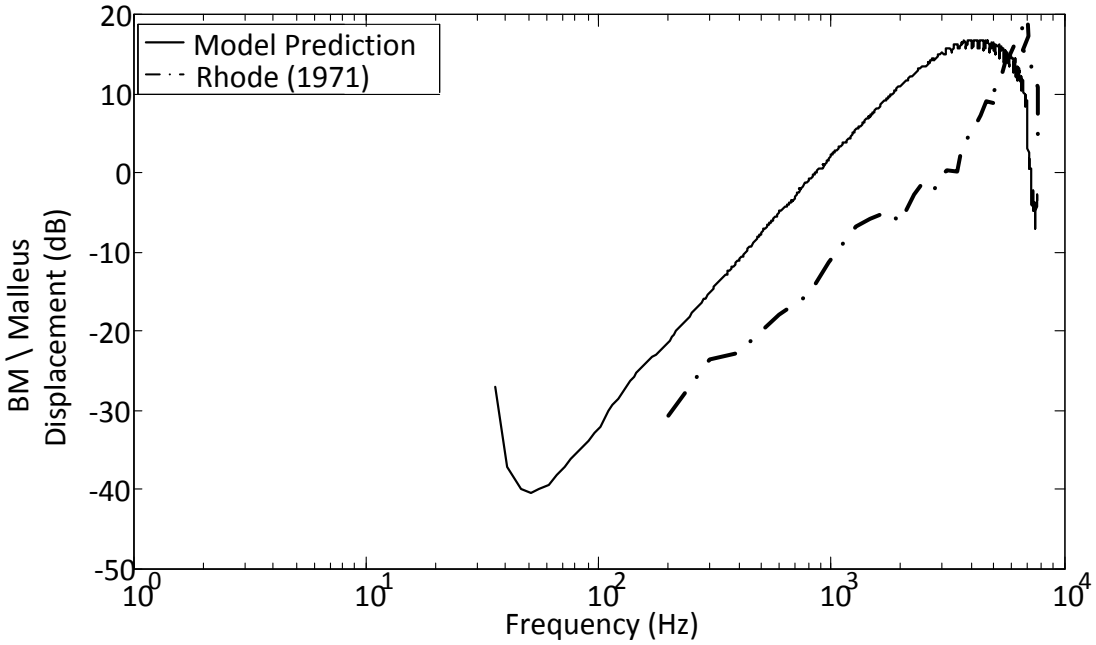

Figure 3.10. Basilar membrane displacement relative to Malleus displacement as predicted by the present model, at a point which relates to 6000 Hz on the Greenwood map, as a function of frequency for a pure tone stimulus of 60 dB SPL. Also shown are measured data from Rhode (1971), measured in squirrel monkeys.

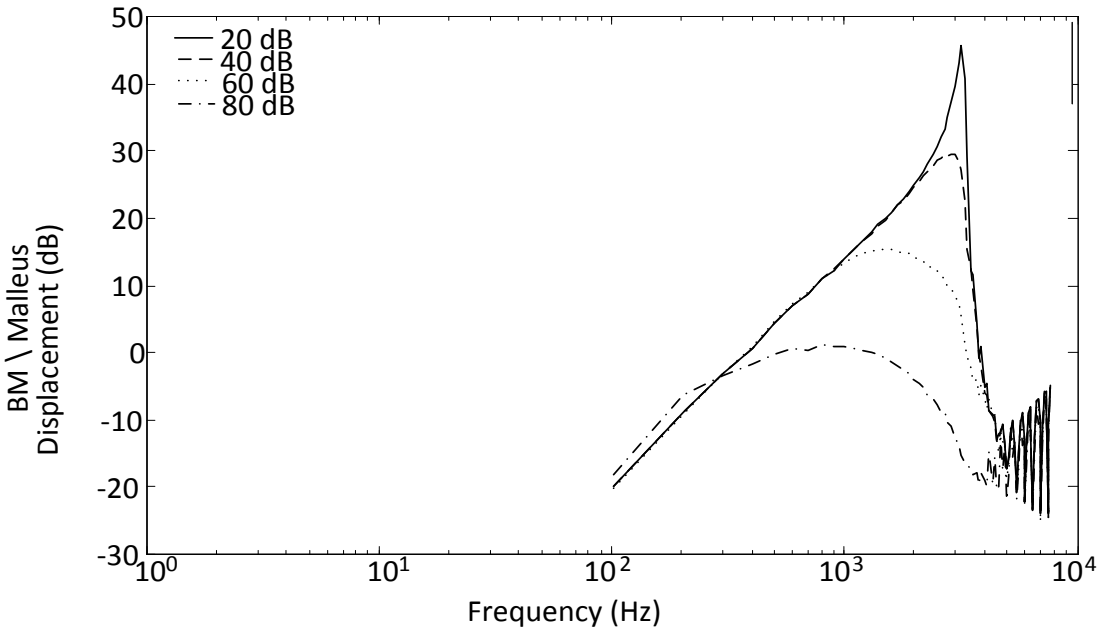

Figure 3.11. Predicted basilar membrane displacement relative to Malleus displacement, at a point which relates to 3000 Hz on the Greenwood map, as a function of frequency for a pure tone stimulations of 20, 40, 60, and 80 dB SPL.

The basilar membrane displacement versus the signal intensity is shown in Figure 3.12. Figure 3.12 shows a comparison between the modelled results and basilar membrane measured gain from a chinchilla cochlea (Ruggero et al., 1997). The chinchilla cochlea was used as it has similar size (18.31 mm) and frequency response 50 Hz to 30 kHz for a chinchilla cochlea to that of a human and no human data is available to compare to the model (Bohne and Carr, 1985). The chinchilla and the human are both primates and the chinchilla measurements be useful in the comparison of the trends in the basilar membrane displacement. Such a comparison will not be useful for accurate basilar membrane displacement of human data.

Data from literature predicts a shift in the basilar membrane position of peak activation with intensity changes, while gain is also a function of stimulus intensity. The model is able to predict the gain of measured data, where the difference in gain between peak points of excitation for the chinchilla data between 80 dB and 5 dB was 47.9 dB (Ruggero et al., 1997) and for the model prediction this difference in gain was 52.8 dB. The shift in position of the maximum amplitude is greater for the predicted basilar membrane displacement (7100 Hz) while the measured data for the chinchilla cochlea was around 2000 Hz (Ruggero et al., 1997). The difference in frequency shift could be due to morphology differences between the chinchilla cochlea (measured data) and the modelled

human cochlea. A major contributor to the steepness of the rise to the peak is the mass (m) constant of the basilar membrane. Increasing this value to 4 kg/m^2 from 0.5 kg/m^2 decreases the frequency difference to 3000 Hz. Whether this is a realistic choice for the human BM is not known. The values chosen for the mass of the basilar membrane in other studies vary across studies. Irrespective of this, however, Figure 3.12 clearly shows the expected decrease in frequency around the peak displacement of the basilar membrane with intensity increase.

In the present model from Duifhuis (2012), the mass of the basilar membrane is constant over the extent of the BM and the value is chosen to match measured data from otoacoustic emissions. Peterson and Bogert (1950) had a larger value for the mass of the basilar membrane (1.43 kg/m^2), however, the equation for the stiffness was a linear function and the width of the basilar membrane varied with distance of the base of the cochlea. Due to the linear function, a shift in the basilar membrane peak would most likely not occur. Epp, Verhey and Mauermann (2010) used the non-linear model to describe the basilar membrane damping and stiffness (as in the case with the travelling wave model used in this study) and loosely compared the results to basilar membrane displacement; Epp et al. (2010) found a mass value of 0.375kg/m^2 , which is far lower than the value used in the present study. To compare the different mass of the models would be extremely difficult at this stage as the effects of the change in mass on different models is unknown due to vastly different approaches in describing the damping and stiffness of the basilar membrane.

To compare this peak shift to psychoacoustic data would also be troublesome as the data predicting the peak shift is related to perceived pitch. Later in this document, it will be shown that pitch can be decoded from nerve fibre firing without place encoding (not relating to peak place). This could mean that the peak shift predicted by psychoacoustic data could be completely incorrect. Therefore, in this study, basilar membrane predictions will be compared to basilar membrane data.

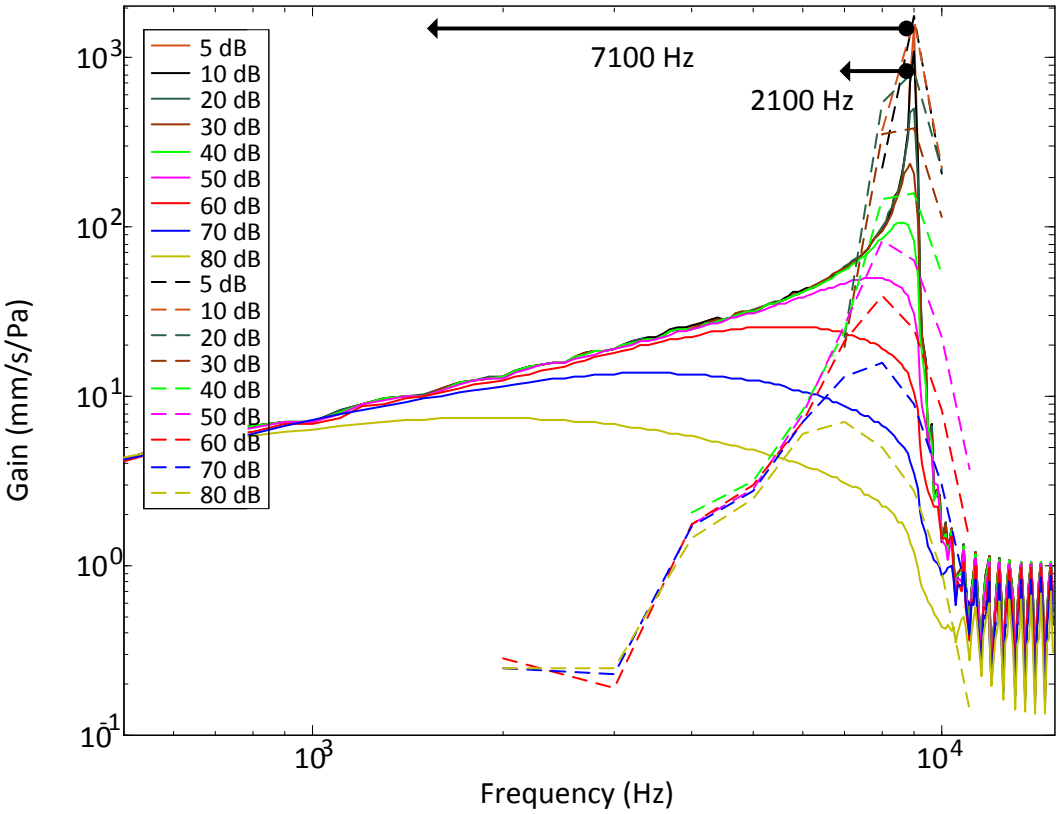

Figure 3.12. Basilar membrane displacement against signal intensity, at a point which relates to 9000 Hz on the Greenwood map. Solid lines represent model predicted data and the dashed line represents the measured data from Ruggero et al. (1997). The difference in minimum gain to maximum gain is 47.9 dB for the measured chinchilla data and 52.8 dB for the predicted modelled data, while the frequency shift of the peak is 2000 Hz and 7100 Hz respectively.

The results in Figure 3.12 suggest that the nerve fibre firing rate peak has shifted greatly towards the base of the cochlea with increased stimulus intensity. However, observing the nerve fibres spike patterns, the synchronisation diagrams, and the tuning curve in the figures that follow suggest that it may illustrate a broadening of the peak, rather than a shift. The leading edge of the travelling wave in proximity of the apex of the basilar membrane has very little shift in position with increasing intensity and this is observed in the nerve fibre reaction below. The nerve fibre spiking patterns and the correct tuning curves, which is indicated later, suggest the broadening of the basilar membrane response, which was more than what was observed for the chinchilla cochlea data in Figure 3.12. The basilar membrane response, clearly illustrates a point of peak excitation. The point of peak basilar membrane displacement is not clear in the nerve fibre tuning curves which are shown below in Figure 3.28 and Figure 3.29.

Figure 3.13 shows the predicted travelling wave along the length of the basilar membrane at different instants in time. These results compare to those of Von Békésy (1960) which modelled the travelling wave movement from the base to the peak point of excitation. In panel (a) of Figure 3.13 the peak of the travelling wave is around 21 mm, in panel (b) the peak point of excitation has moved to around 27mm. The panel (c) and (d) show the second period of the stimulus while the first stimulus peak decreases. This can be seen in Figure 3.14.

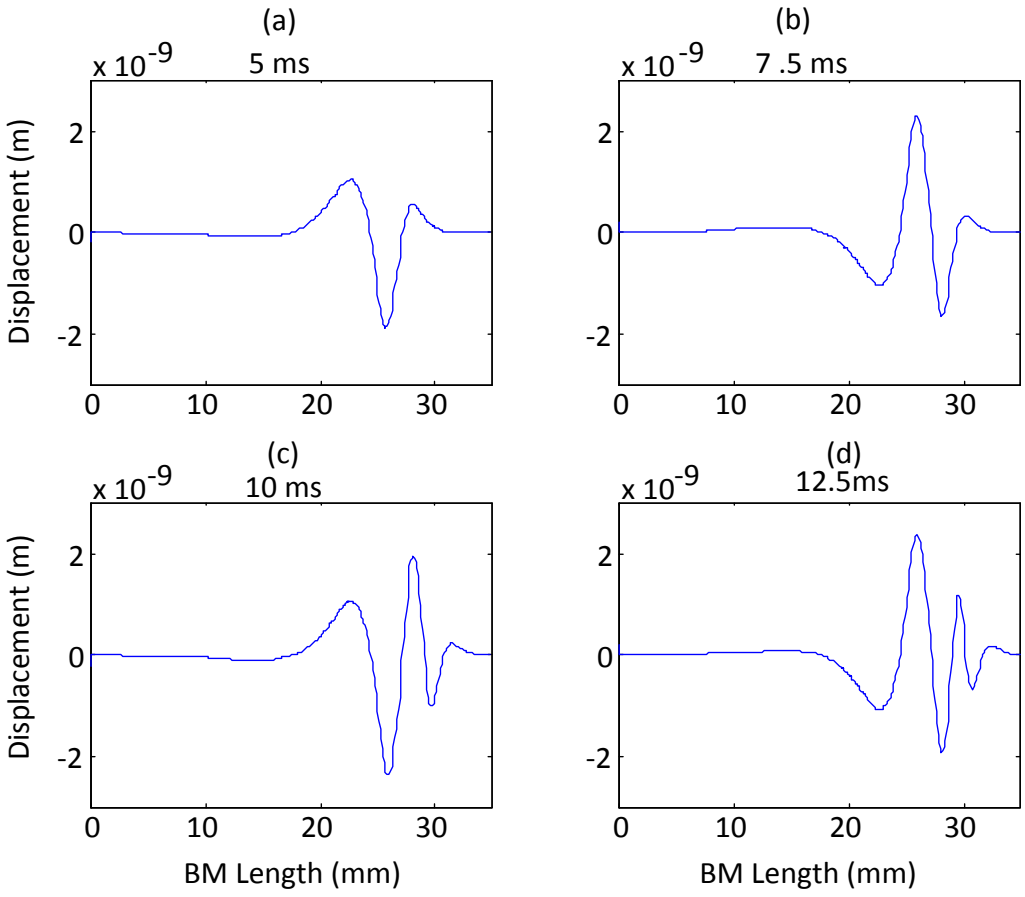

Figure 3.13. Illustration of the travelling wave with a pure tone stimulation with a frequency of 200 Hz. These results correspond to the movement of the travelling wave seen in Von Békésy (1960).

Figure 3.14 below shows the travelling wave along the length of the basilar membrane for a pure tone at three intensity levels of 20, 40, and 60 dB SPL. The Figure 3.12 and Figure 3.15 shows the shift in basilar membrane displacement towards the base as the intensity increases, the increase in basilar membrane deflection, and the movement of the travelling wave from base to apex.

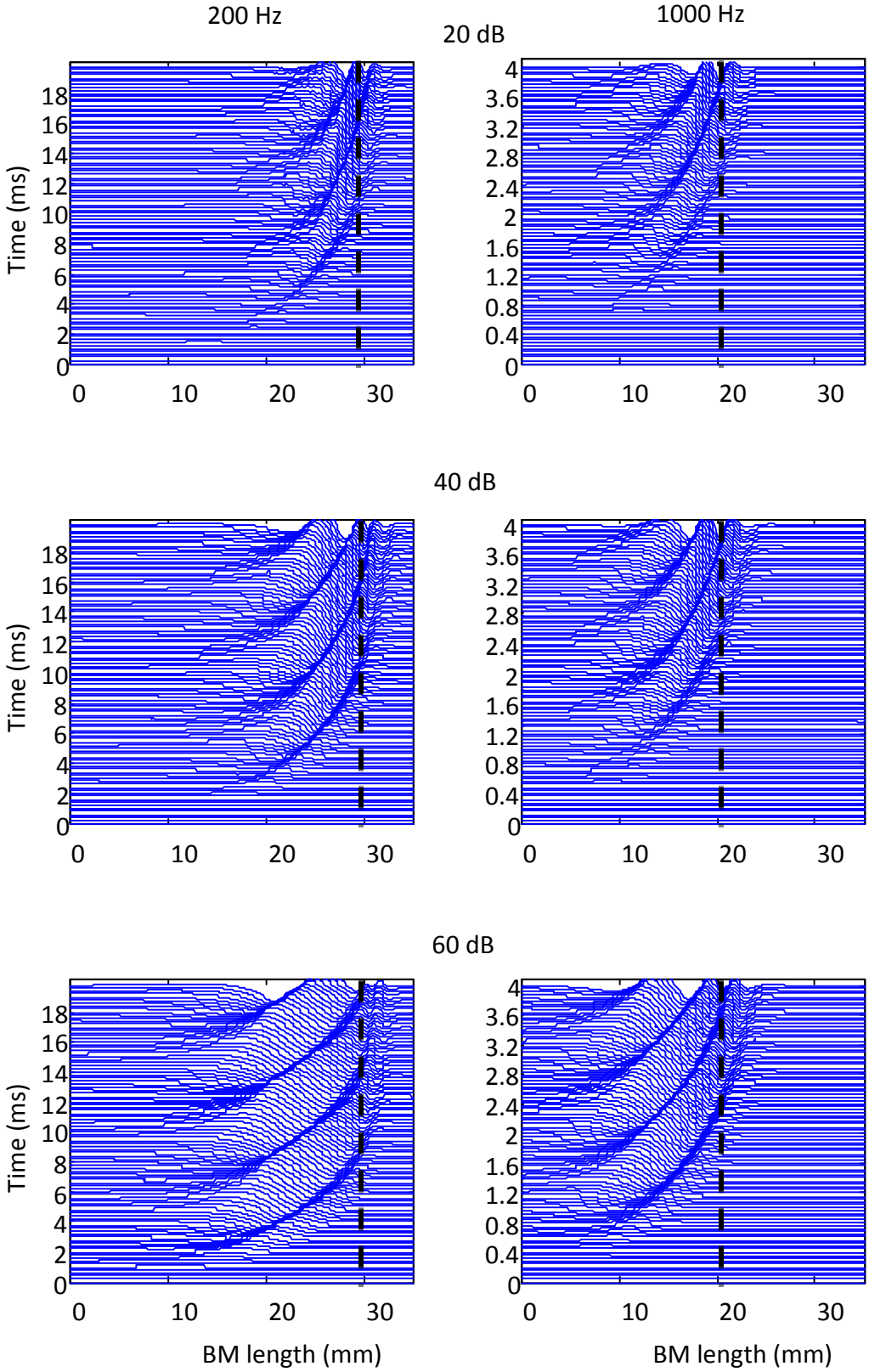

Figure 3.14. Travelling wave representation with a pure tone input of 200Hz for three different intensity levels of 20 dB, 40 dB, and 60 dB SPL respectively. The solid dashed line is the Greenwood map frequency for 200 Hz and 1000 Hz. The period can be calculated from the diagram by counting the number of travelling wave produce during the duration of the simulation, this can then be used to calculate the period of the signal and gives the frequency of stimulation in a pure tone stimulation case. In the simulation above,

4 travelling waves can be counted for all diagrams, giving a frequency of 200 Hz and 1000 Hz respectively. The non-smooth, "blockiness" of the figures is due to the decrease in resolution, over the time axis, during the plotting of the figure. Decreasing the resolution allows for less overlapping of the oscillations; this makes the figure easier to observe the travelling wave in this plot style.

The peak shift, Figure 3.12 , with changes in intensity warrants further comment. Intensity level may be encoded in the increase in the spike rate of nerve fibres locking onto the frequency of the stimulation signal. However, this mechanism does not account for the dynamic range difference between nerve fibres' firing rate and perception. The change in firing rate of nerve fibres, due to intensity increase, has a small dynamic range of 30 dB (Ruggero et al., 1997) while normal hearing has a much greater dynamic range around 80 dB. The firing rate of nerve fibres will increase with loudness level but will saturate (Rattay, 1990, Yates, Manley and Köppl, 2000, Sumner, O'Mard, Lopez-Poveda and Meddis, 2003). The loudness level could possibly be encoded in the size of the activated population of nerve fibres along the length of the basilar membrane which may represent the energy of the signal. Such a model for loudness encoding was proposed, among others, by Chatterjee and Zwislocki (1998). The model presented here suggests that loudness may be encoded by the shift in basilar membrane peak towards the base of the cochlea and the broadening of the basilar membrane displacement with loudness increase. The latter relates to a greater number of nerve fibres phase locking to the stimulus signal (Greenberg, 1997), while the shift in basilar membrane peak towards the base of the cochlea represents a potential place encoding mechanism for loudness level. However, seen from the synchronisation index graphs below, the peak of phase locking does not shift with the basilar membrane peak excitation for loudness increase. On the other hand, the broadening of the basilar membrane excitation region does increase the number of nerve fibres synchronising to the stimulus signal, which could encode loudness by the number of nerve fibres firing and phase locking to the signal.

Figure 3.15 is a 1D representation of the maximum deflection of the travelling wave for pure tone stimulation. The point of maximum deflection can be clearly seen to move towards the base (0 mm) of the basilar membrane during the increase in intensity. Ruggero et al. (1997) found that for an increase in stimulation from 10 dB to 80 dB intensity level, the shift in place frequency, from 9000 kHz to around 7000 kHz, corresponds to 0.36 octaves in the chinchilla cochlea. Chatterjee and Zwislocki (1997) found that the frequency shift in the peak can be as large as two octaves when sound pressure level is increased from 10 dB to 100 dB. The width of deflection area increases as the stimulation level

increases towards 80 dB SPL. This could indicate that the pitch information presented at lower sound pressure levels may be based on place mechanisms, but this mechanism becomes unreliable at high intensities due to broadening of the deflection area. The basilar membrane displacement broadening allows for temporal mechanisms to be used to gain information about the pitch of the tone (Greenberg, 1997, Greenberg, Poeppel and Roberts, 1998). The broader basilar membrane displacement allows for more nerve fibres to synchronise to the pure tone, as will be shown below in the synchronisation figures that follow.

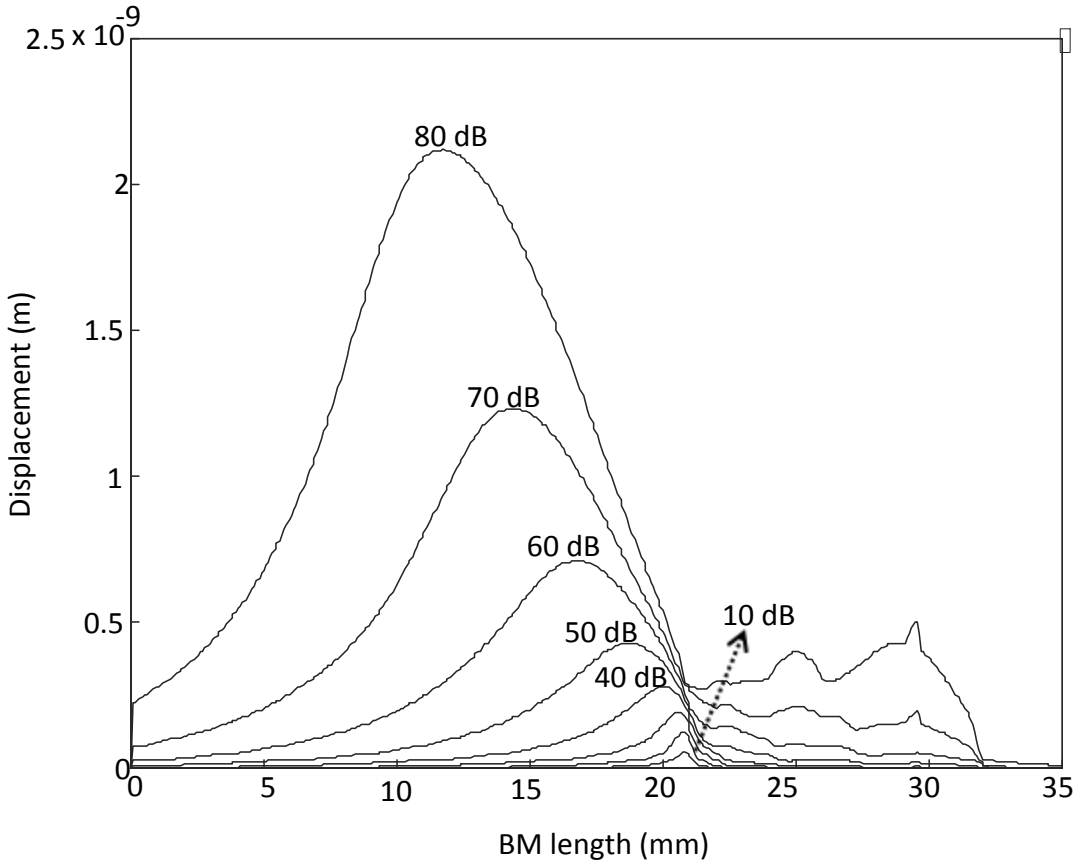

Figure 3.15. Basilar membrane maximum displacement for a pure tone input of 1000 Hz for loudness steps from 10dB to 80dB SPL.

3.3.3 Predicted neural activation patterns

A number of different views of the model-predicted neural activation patterns are considered here and compared to measured data from literature. Nerve fibres spike patterns, ISI histograms, synchronisation of nerve fibres to the stimulus, and neural tuning curves are shown below. Nerve fibre spike patterns are shown for low, medium and high spontaneous rate nerve fibres, for pure tone stimuli at different intensity levels. Figure 3.16

is the nerve fibre spiking patterns for low spontaneous rate nerve fibres along the length of the basilar membrane for a pure tone input of 1000 Hz at four different intensity levels.

In low spontaneous rate fibres very few nerve fibres fire during low intensity stimulation, as seen during 20 dB SPL stimulation. As the loudness is increased, more nerve fibres can be seen to fire around the 20.96 mm area where the CF is 1000Hz (according to the Greenwood map for human predictions (Greenwood, 1990)). Intensity level increases towards 80 dB SPL causes more nerve fibres to fire along the length of the cochlea.

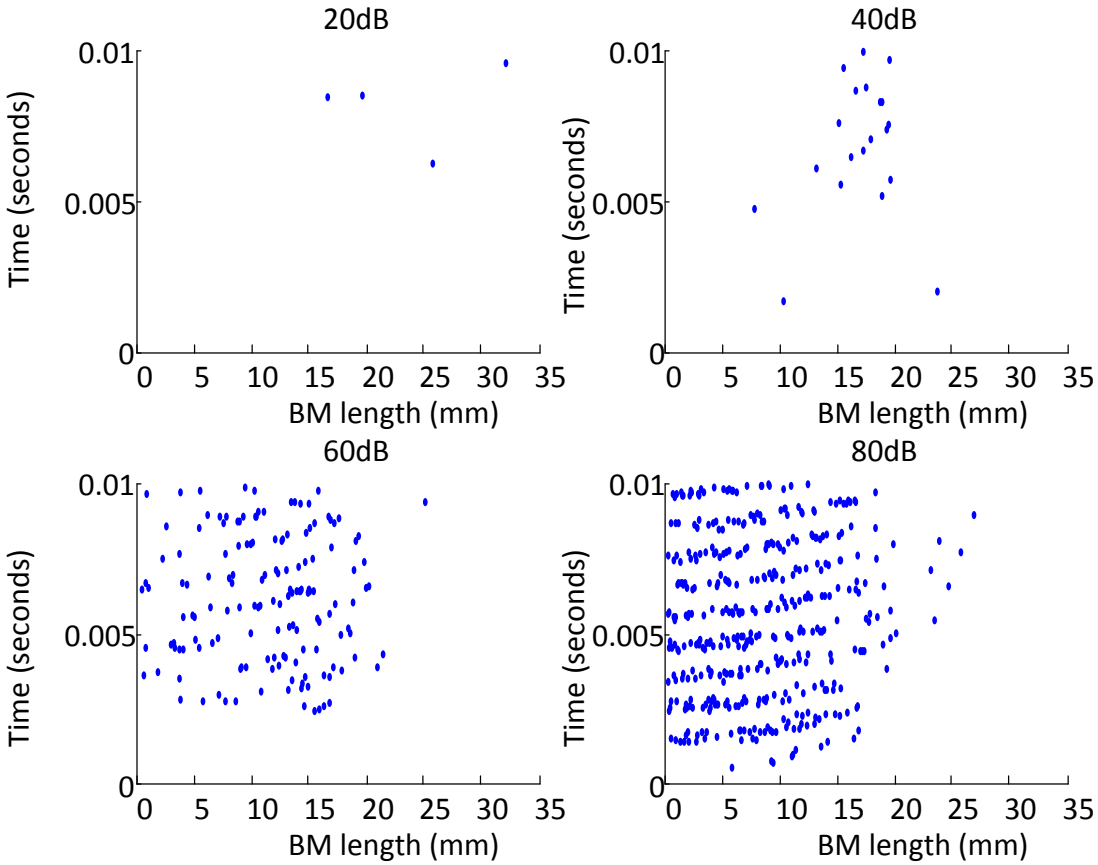

Figure 3.16. Nerve fibre firing diagram for low spontaneous rate fibres for a pure tone stimulation of 1000 Hz. The CF for the 1000 Hz pure tone is 20.96 mm.

When considering medium and high spontaneous rate nerve fibres, the same stimuli result in more nerve fibres firing along the length of the cochlea. There is a higher number of firings around the CF of the stimulation pure tone as the spontaneous rate increases. This relates to the often-observed decrease in threshold as the spontaneous rate increases (Sachs and Abbas, 1974).

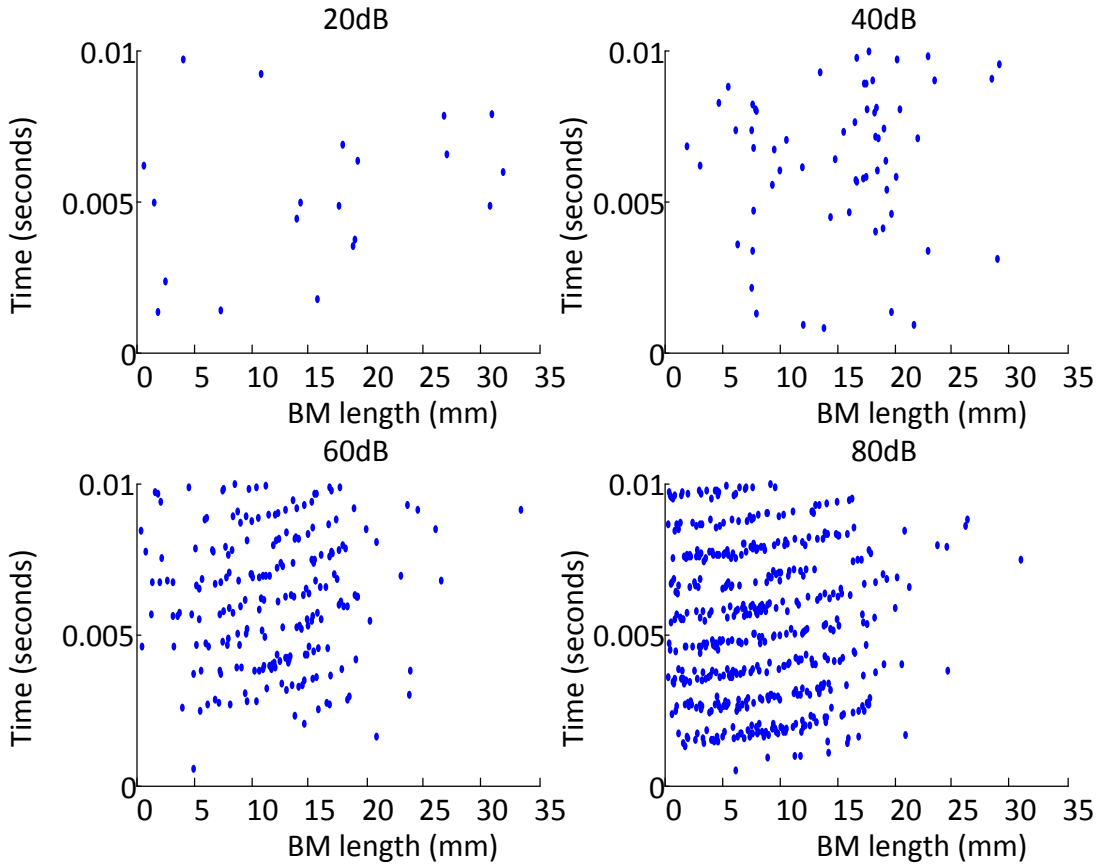

Figure 3.17. Nerve fibre firing diagram for medium spontaneous rate fibres for a pure tone stimulation of 1000 Hz.

Nerve fibre firing patterns from the low, medium, and high spontaneous rate nerve fibres in Figure 3.16, Figure 3.17, and Figure 3.18 illustrate that low and medium spontaneous rate fibres may contain place information that is available to encode pitch information, as modelled by the higher number of nerve fibres firing at the cochlea place relating to the frequency of stimulation. The nerve fibre spiking patterns for low and medium spontaneous rates show a stronger grouping of nerve fibres firing around the place relating the peak of excitation.

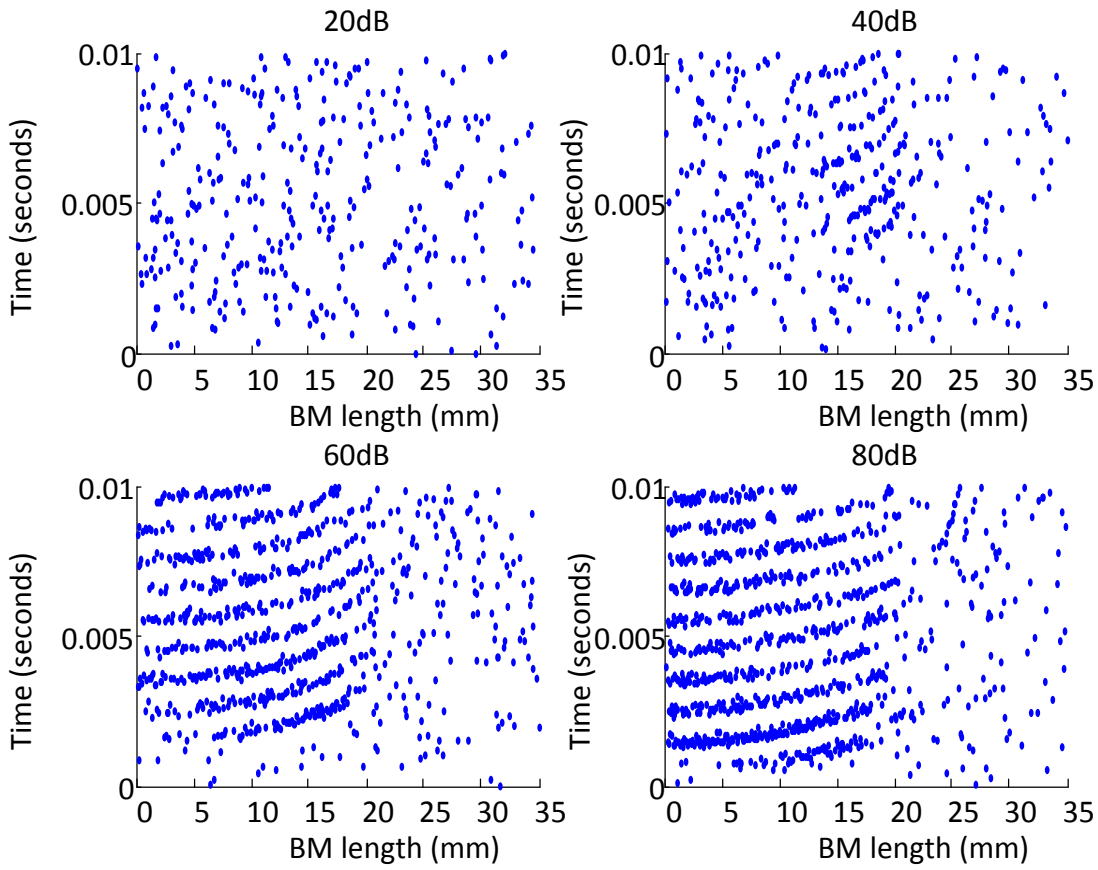

Figure 3.18. Nerve fibre firing diagram for high spontaneous rate fibres for a pure tone stimulation of 1000 Hz.

This could be due to higher thresholds of lower spontaneous rate nerve fibres (Sachs and Abbas, 1974), which then requires a higher basilar membrane amplitude for the nerve fibres to fire, causing the peak position of stimulation to dominate nerve fibre firings. For high spontaneous rate fibres, place encoding is probably lost: as seen in the predicted spike train patterns, where there is a high number of nerve fibres firing along the entire basilar membrane length and not only at the place of maximal basilar membrane displacement. The temporal characteristics are considered below.

The increase in stimulation intensity causes the increase in nerve fibres firing towards the base of the cochlea. An increase in nerve fibres firing along the length of the cochlea may cause temporal encoding to be dominant at high intensity level, as place encoding cannot encode pitch due to the shift in peak point of excitation and broadening of the peak (Chatterjee and Zwislocki, 1998). The increase in the number of nerve fibres firing along the length of the cochlea may be a possibility of intensity encoding, due to the number of nerve fibres relating to the energy of the travelling wave, which increases with intensity

increase of stimulation causing a greater number of nerve fibres to fire (Lachs, Al-Shaikh and Bi, 1984, Duifhuis, 2012).

ISI histograms enable the consideration of phase locking of spikes onto the temporal information of the pure tone stimulus. All-order ISI histograms (Cedolin and Delgutte, 2005) are used. These histograms reflect the time difference between all nerve fibre firings and not only the difference between two consecutive nerve fibre firings. All-order ISI histograms were also used for pitch detection, described in a later chapter.

Figure 3.19 shows the first order first-order ISI histogram predictions for pure tone stimulations. The peaks cluster around the period and the period multiples of the stimulation frequency period, the peaks height decrease as the intervals increase in time. The clusters around the peaks decrease in width with increase in frequency. These characteristics compare well to those in Figure 3.20 from (Rose, Brugge, Anderson and Hind, 1968). Comparison between the predicted and measured ISI histograms shows: fewer spikes for the predicted results in Figure 3.19 compared to the measured data (this could be due to stimulus duration differences), a decrease in number of intervals for longer interval periods, and for 1000 Hz and 1150 Hz the measured results of Figure 3.20 have a higher first peak whereas the predicted results have a higher second peak (this is due to the refractory period of the auditory model, where the stimulus period approaches the refractory period of nerve fibres used in the model).

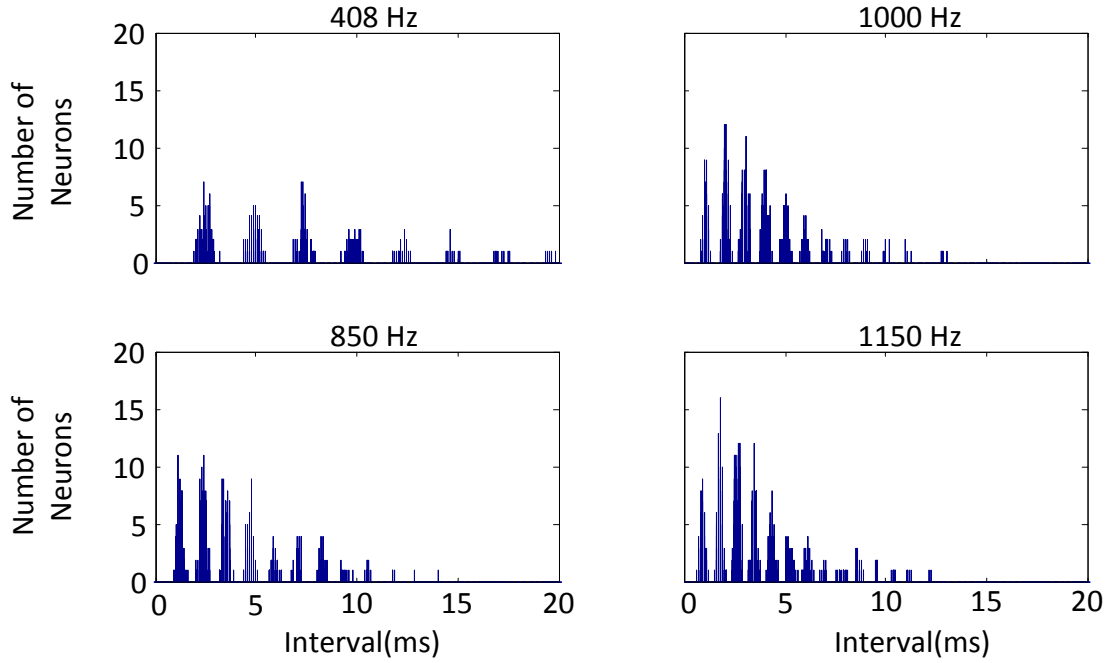

Figure 3.19. Model prediction of first-oder ISI histograms for a nerve fibre at 2.1 mm from the base of the cochlea at multiple stimulus frequencies. Stimulus duration was 500 ms.

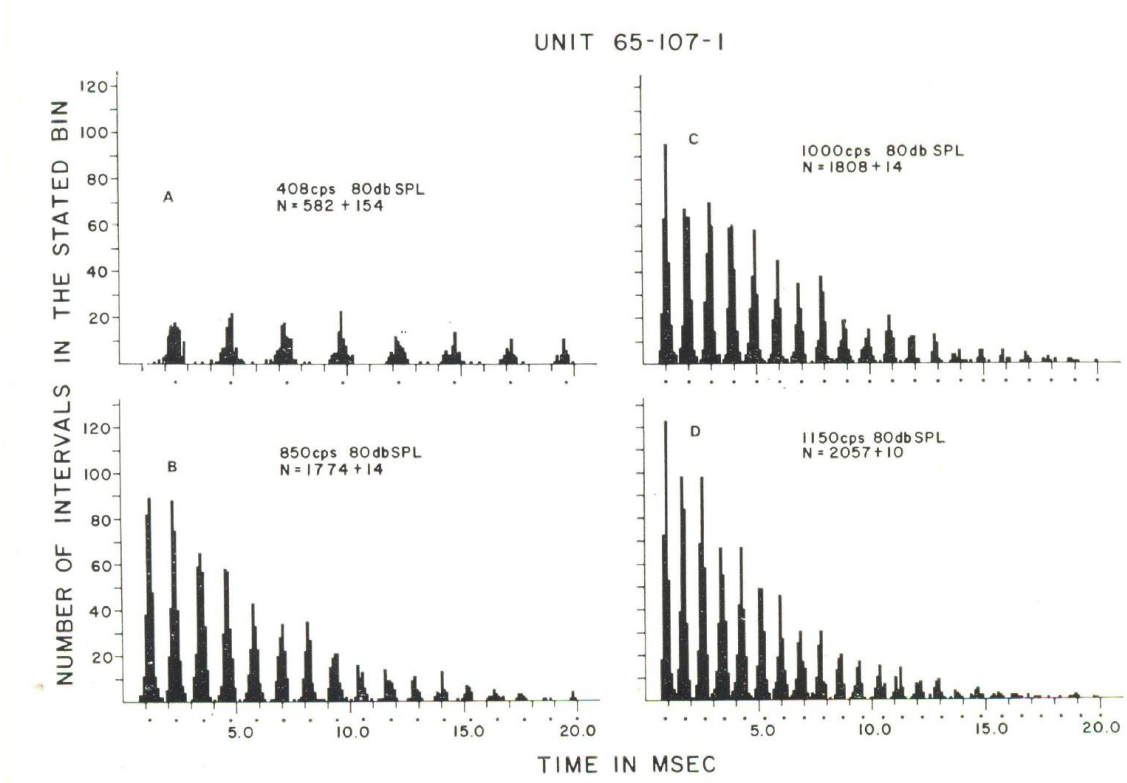

Figure 3.20. ISI histograms when pure tones of different frequencies activated the neuron measured in Rose et al. (1968). The stimulus was at 80 dB SPL with a stimulus duration of 1 second. From Rose et al. (1968) with permission.

Figure 3.21 shows predictions for the population of nerve fibres along the length of the cochlea for pure tones from 100 Hz to 4000 Hz. Interval locking of nerve fibres to the

period of the pure tone signal can be seen clearly up to 2000 Hz. The consideration of the synchronisation index later in the chapter, explores this principle further. A noise floor can be clearly seen in the prediction of the intervals.

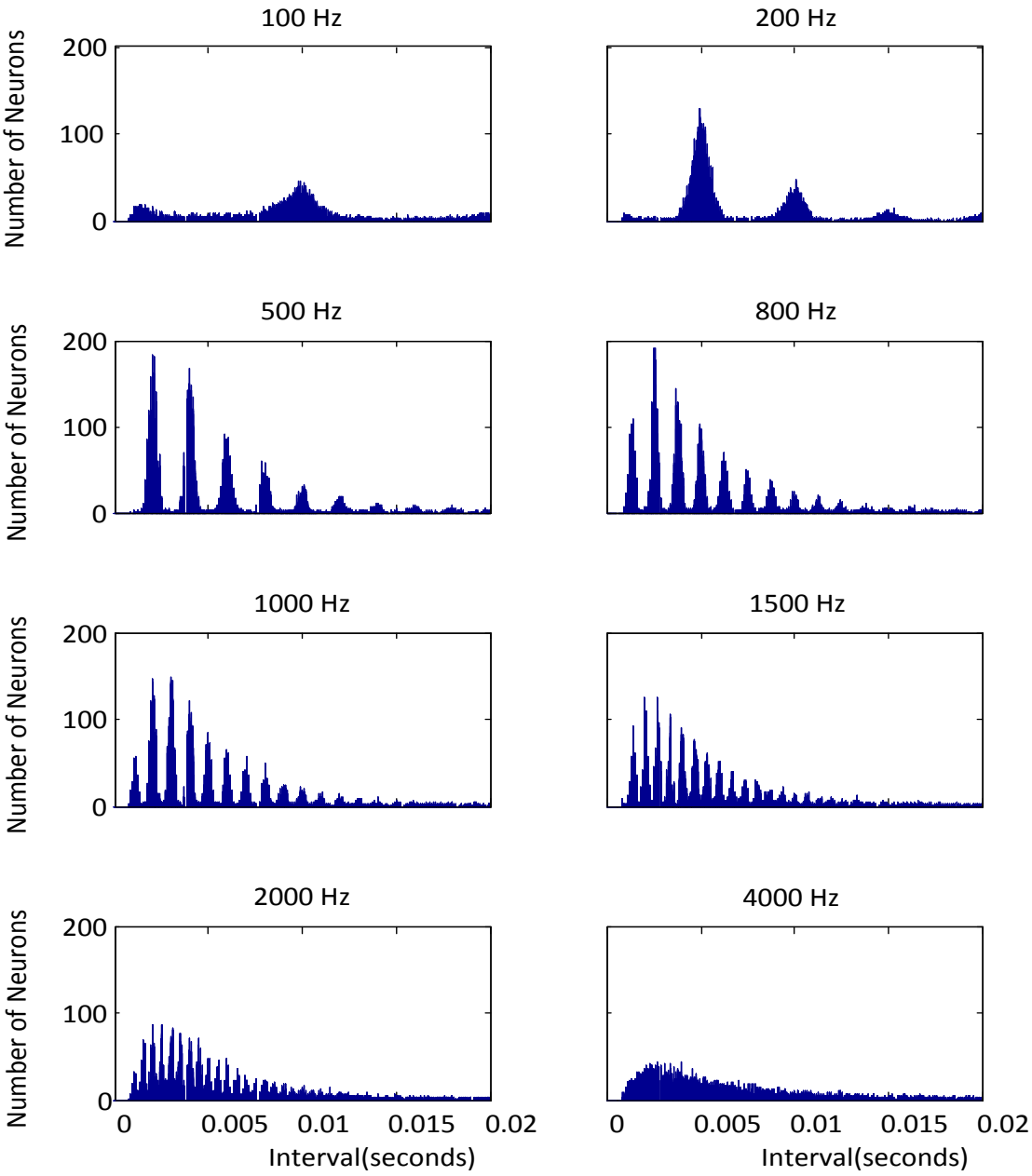

Figure 3.21. First order ISI histograms for pure tone stimulations of 100, 200, 500, 800, 1000, 1500, 2000, and 4000 Hz at 70 dB.

Figure 3.22 below show ISI histograms for pure tone stimulation by a 1000 Hz tone from 20 dB to 80 dB in 20 dB steps, for low, medium, and high spontaneous rate fibres. An increase in sound pressure level results in an increase in the number of nerve fibres locked onto 1 ms time intervals (corresponding to the 1000 Hz stimulus frequency) and its multiples. With higher spontaneous rates, it is observed that more random firings occur between the 1 ms time intervals, along with an increase in the peak number of neurons, and an increase in the width of histogram clusters.

The thresholds of the nerve fibres differ according to fibre type; high spontaneous rate fibres have a lower threshold and the threshold is higher in lower spontaneous rate fibres. The results of these threshold differences can be observed in the firing diagrams (lower number of firings on the nerve fibre), and the ISI histograms (lower peaks surrounding the interval of the pure tone).

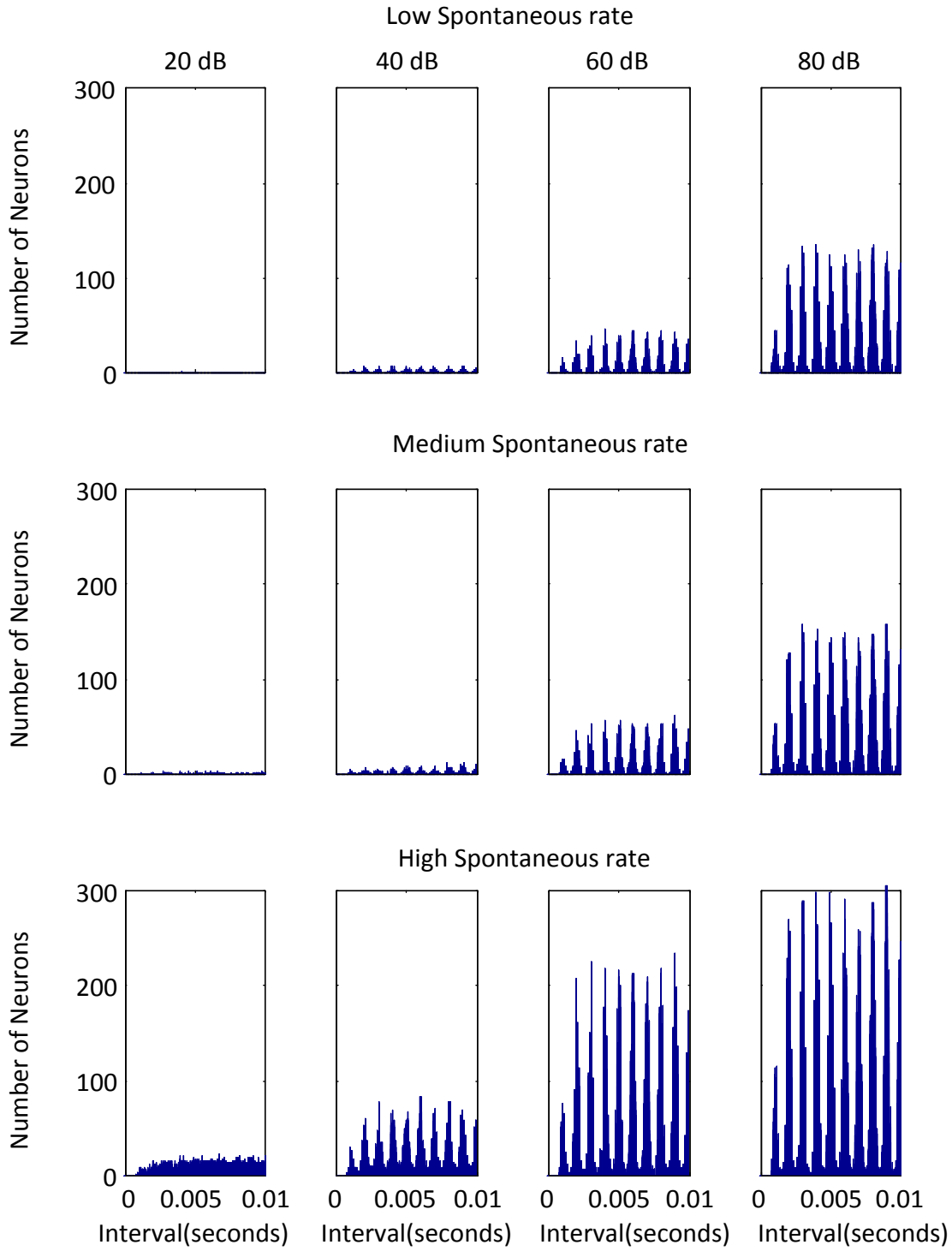

Figure 3.22. ISI histograms for nerve fibres of different spontaneous rates for 1000 Hz pure tone.

3.3.4 Predicted synchronisation index of nerve fibres

The synchronisation index is a measure of the nerve fibre firing clustering around the pure tone signal stimulation peak (seen in a period histogram), when observing one full period of the stimulation signal. Synchronisation is calculated by finding the period histogram from spike trains and then calculating the number of fibres locking onto the period of the pure tone by using the following equation:

$$S_i = \sum_{n=1}^{N} \frac{P_n \sin(2\pi f(t_n - \emptyset_i))}{N}, \tag{3.35}$$

where S_i is the synchronisation of the nerve fibre i, n is the bin number, N the total number of bins for a nerve fibre over one period, f is the frequency of the stimulation signal, P_n is the bin count value (number of spikes in the bin) in the period histogram, t_n is the time index of the period histogram, and $\emptyset_i$ is the phase shift of the travelling wave delay along the length of the BM with reference to the position of the nerve fibre (i).

The synchronisation of nerve fibres along the length of the basilar membrane is shown in Figure 3.23. A 1000 Hz pure tone stimulation was used at four different intensity levels, 20 dB to 80 dB with 20 dB steps. The place of stimulation, relating to the 1000 Hz CF is around 20.96 mm (Greenwood's map adapted to fit human predictions for place corresponding to a particular frequency (Greenwood, 1990)). During low sound pressure level, synchronisation is confined in proximity to the point of the CF, as observed in Figure 3.23, for a sound pressure level of 20 dB SPL. An increase in sound pressure level results in an increase in the number of nerve fibres along the length of the basilar membrane synchronising to the pure tone stimulation.

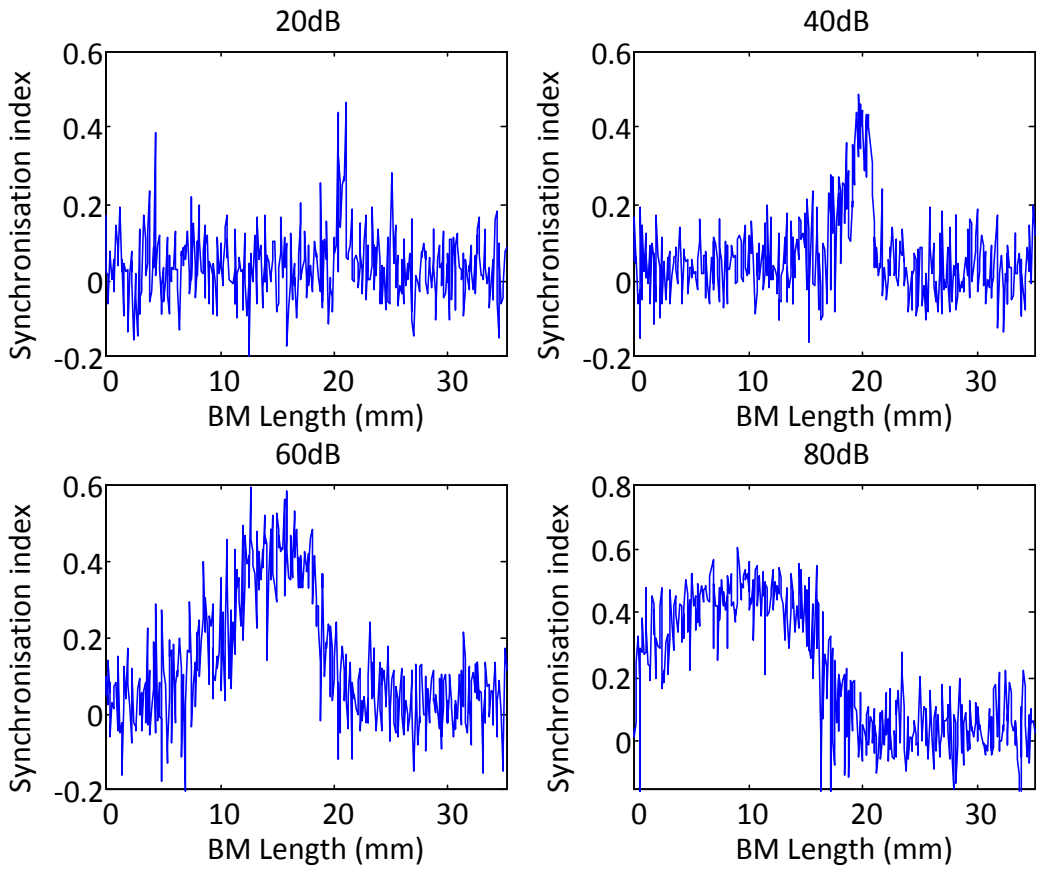

Figure 3.23. Neural synchronisation index versus basilar membrane position for 1000Hz pure tone input stimulation.

Figure 3.24 illustrates a comparison between nerve fibres of different spontaneous rates. The high spontaneous rate fibres have a spatial spread of low synchronised nerve fibres along the length of the basilar membrane, where the lower levels of spontaneous rate nerve fibres are more likely to have a synchronisation of zero and where no stimulation from the basilar membrane occurs. Higher levels of stimulation result in a spatial spread of synchronisation along the length of the basilar membrane towards the base of the cochlea, which decreases the sharpness of the place of excitation. The shift in peak synchronisation is not as clearly defined as the shift observed in basilar membrane displacement, illustrated earlier in the study. During high levels of stimulation, the peak of synchronisation is completely lost, suggesting the loss of place information and the dominance of temporal encoding of the nerve fibre.

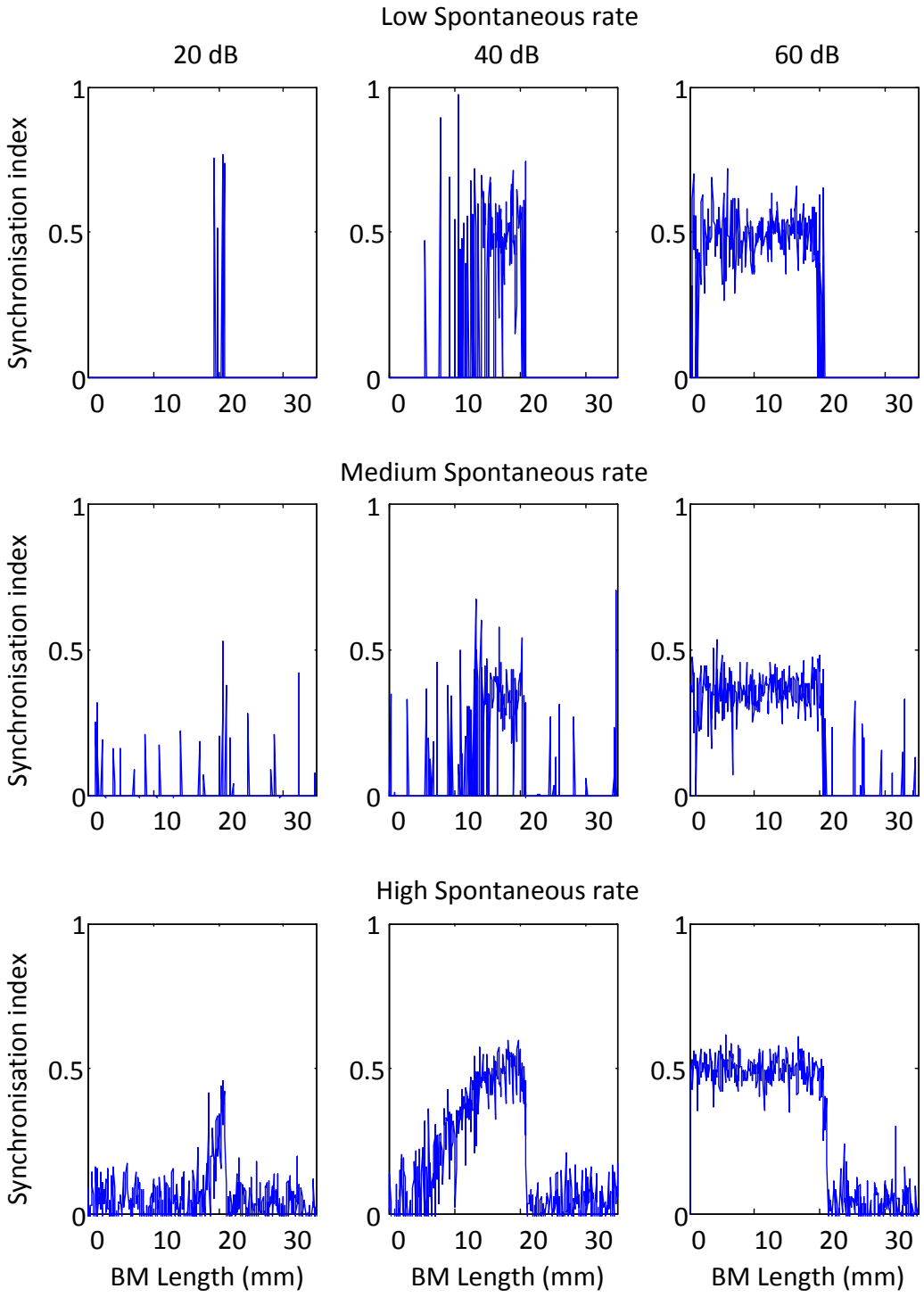

Figure 3.24. Synchronisation Index for low, medium, and high spontaneous rate nerve fibres for a pure tone stimulation of 1000 Hz at different intensity levels.

Figure 3.25 demonstrates the relation to the CF of the nerve fibre synchronisation along the length of the basilar membrane. This relation illustrates that not only fibres stimulated at

their CF are synchronised to the pure tone, but fibres with higher CF are also synchronised, especially at higher stimulation levels. Synchronisation of nerve fibres at higher stimulation may contain information about the pure tone stimulus, which could correspond to the temporal encoding employed by the auditory system at higher intensity levels.

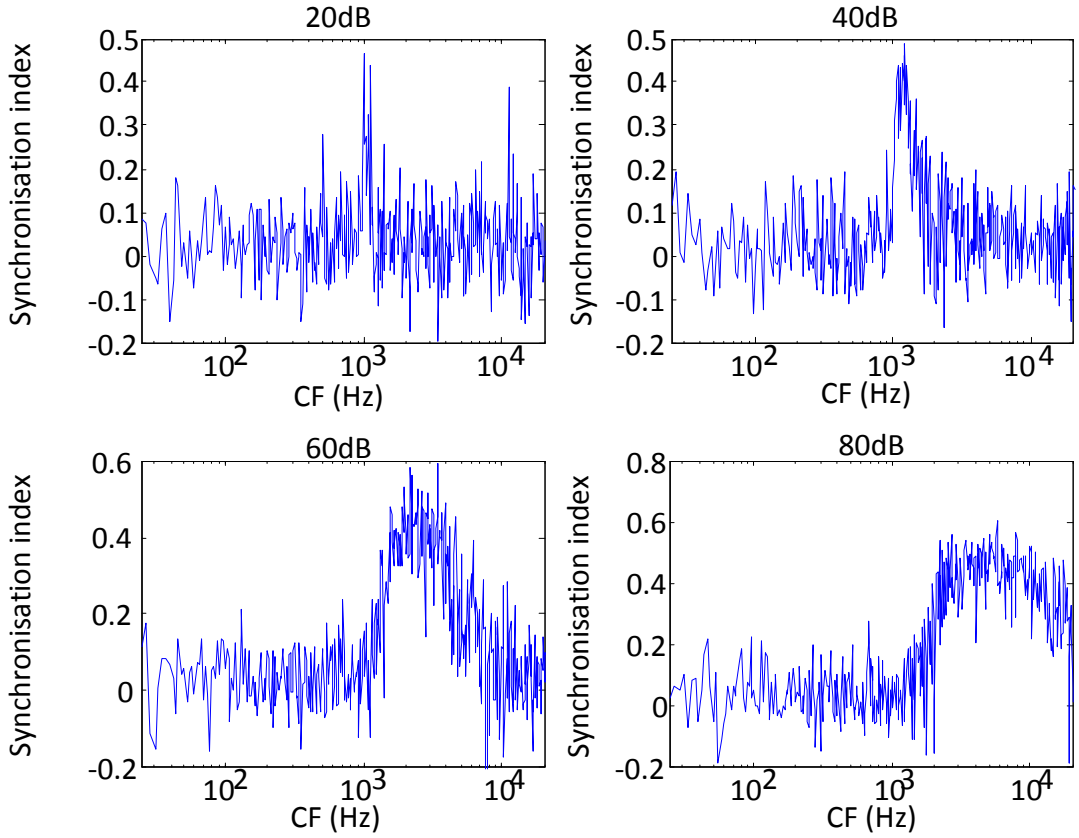

Figure 3.25. Synchronisation index as a function of characteristic frequency of nerve fibres along the length of the basilar membrane for 1000Hz pure tone input stimulation at multiple intensity levels.

Rate place information at low level of stimulation perhaps proves to be more reliable, demonstrated by Figure 3.23, Figure 3.24 and Figure 3.25. The increase in stimulation level affects the basilar membrane displacement width and the spatial synchronisation of nerve fibre increase towards the base of the cochlea, causing the peak of synchronisation to lack clear definition. At 80 dB stimulation levels, the place information is clearly not defined, suggesting that the temporal encoding may be more reliable (Chatterjee and Zwislocki, 1998). Predictions in Figure 3.23 to Figure 3.25 can be compared to data. Figure 3.26 below shows the discharge rate of nerve fibres along the length of the cochlea for pure tone stimulus from Kim, Parham, Sirianni and Chang (1991). The measured results from Kim et al. (1991) show an increase in spatial width of firing with an increase

in intensity level. This correlates with the predicted nerve fibres response seen in the synchronisation response (Figure 3.23, Figure 3.24, and Figure 3.25) and the nerve fibre spiking patterns (Figure 3.16, Figure 3.17, and Figure 3.18) of the model.

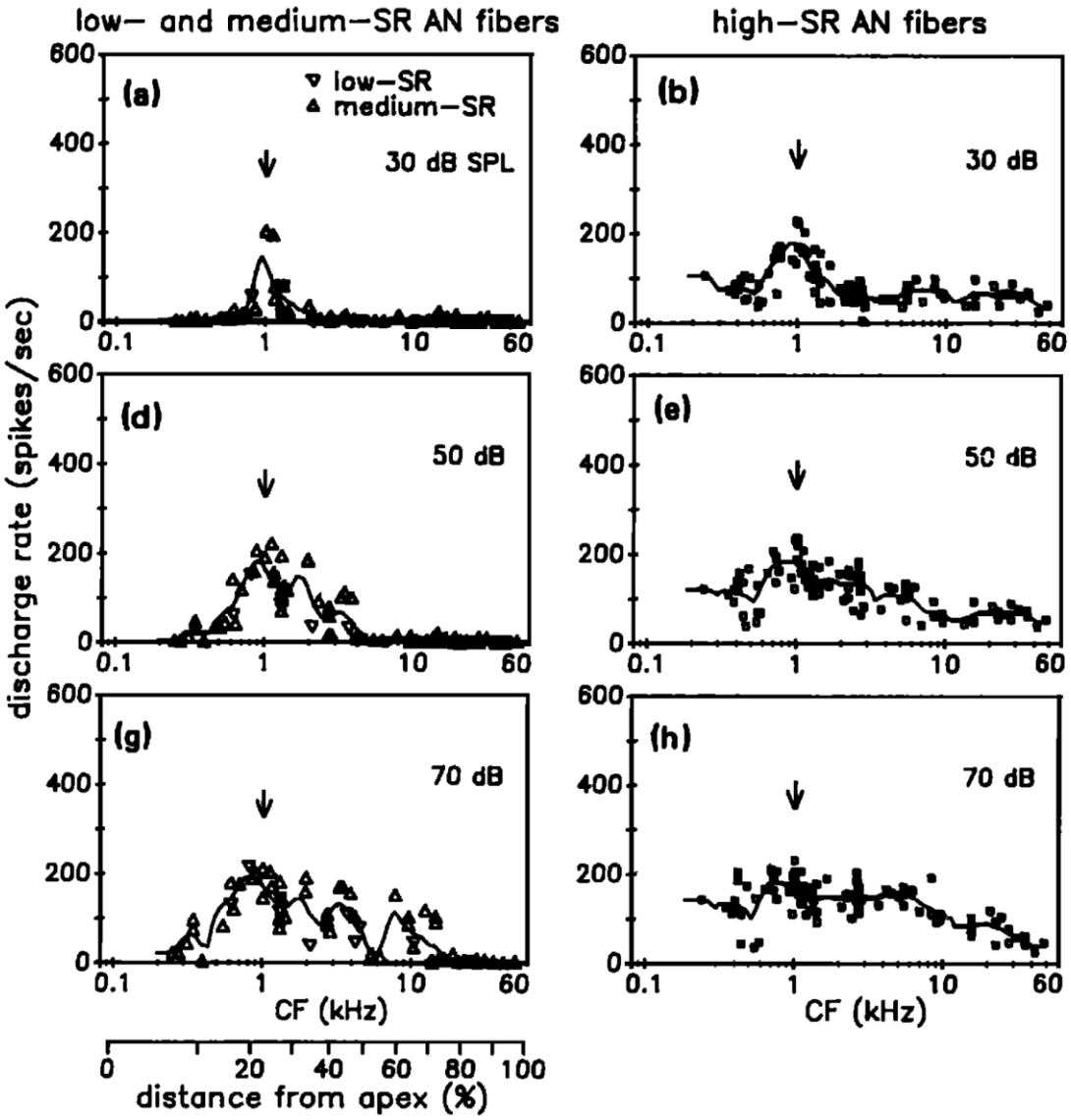

Figure 3.26. Spatial profiles of absolute discharge-rate response of two neural groups in unanesthetized decerebrate cats (Kim et al., 1991). The input stimulation is 1000 Hz pure tone. From Kim et al. (1991) with permission.

Figure 3.27 is the average synchronisation of nerve fibres along the length of the cochlea for a pure tone stimulus. The results show an increase in synchronisation index from 50 Hz to 100 Hz, which could be attributed to the fact that the stimulus duration was too short to cause a higher average synchronisation at 50 to 100 Hz. The average trend is then for the synchronisation index to decrease almost linearly as frequency increases up to around 1400 Hz. At higher frequencies, the average synchronisation is close to zero. A synchronisation

rate close to zero could mean that temporal encoding does not occur at these higher frequencies. The lack of temporal encoding at higher frequencies will be discussed in later chapters.

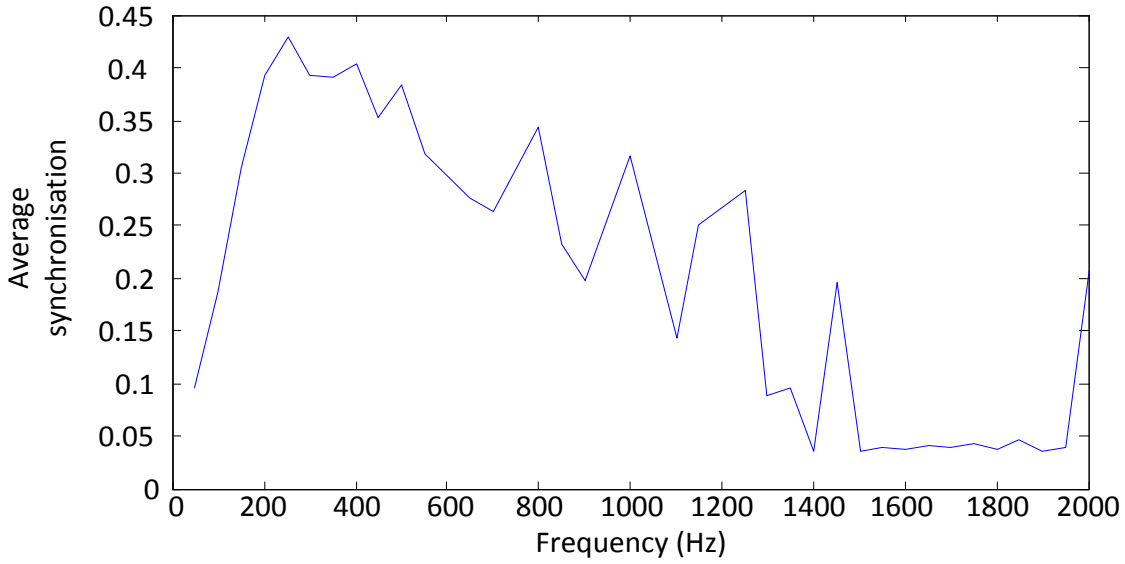

Figure 3.27. Average synchronisation predictions of 400 nerve fibres for stimulus frequencies across the length of the basilar membrane.

3.3.5 Predicted neural tuning curves

This section considers the predicted neural tuning curves of a single auditory nerve fibre in the cochlea at different positions along the length of the cochlea. The neural tuning curve is the nerve fibre's threshold at which the nerve fibre firing rate is above the spontaneous rate of the nerve fibre. In fact, what is shown here are not tuning curves per definition, but tuning characteristics that reflect the tuning curve. However, these tuning characteristics will simply be referred to as tuning curves in this work. They are obtained by selecting a nerve fibre along the length of the basilar membrane. Multiple stimulations at different frequencies are applied to the nerve fibre via the normal hearing travelling wave model and the predicted firing rate of the nerve fibre is determined. The step size for intensity level was 2 dB steps from 0 dB to 80 dB and for frequency, 100 steps per decade from 0 Hz to 20 000 Hz.

Figure 3.28 and Figure 3.29 show the predicted tuning curves for two high spontaneous rate nerve fibres. The fibre in Figure 3.28 should have a CF of 1715 Hz, based on its position in the cochlea (Greenwood, 1990). Note that the selected fibre is no different from any of the other fibres in this model – it does not have a specific pre-programmed tuning

curve. The firing rate has a peak with minimum threshold of around 15 dB SPL. An increase in stimulation frequency shows the threshold increasing rapidly until the nerve fibre fires at the spontaneous rate (this nerve fibre was a high spontaneous rate nerve fibre and had a spontaneous rate of 100 spikes/second), while a decrease in stimulation frequency shows a slower drop-off of threshold.

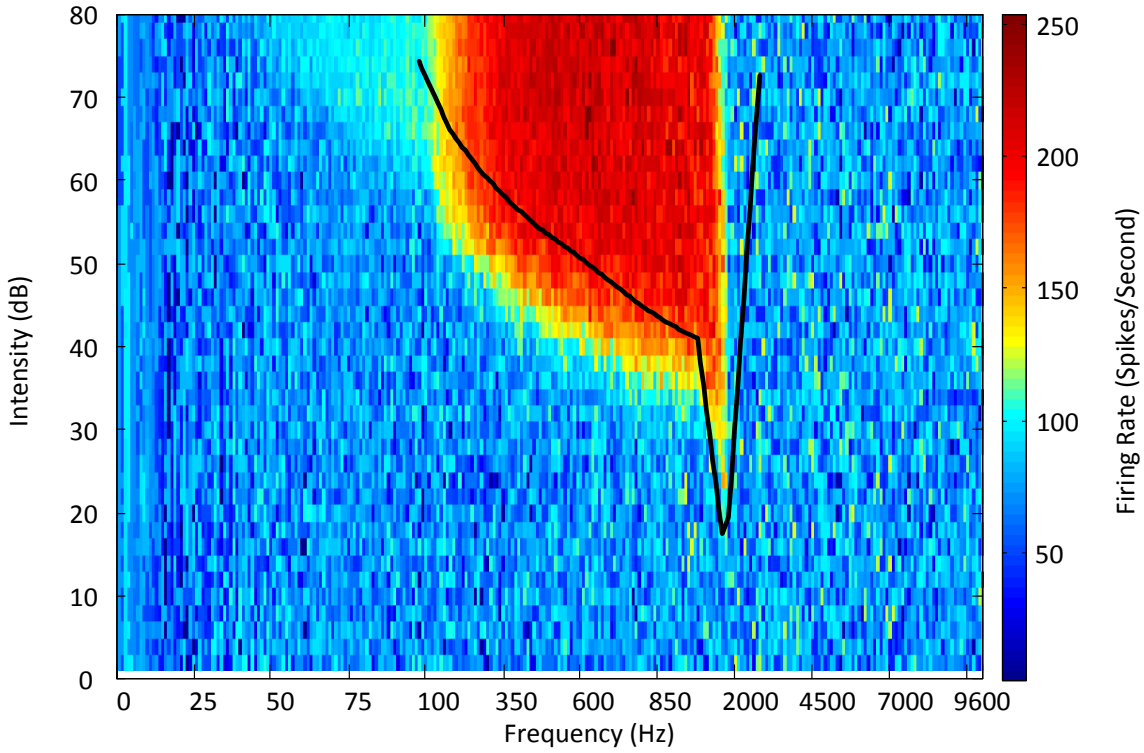

Figure 3.28. Nerve fibre tuning curve with a CF of 1715 Hz. The solid line represents the nerve fibre tuning curve for measured data (Ruggero et al., 1997, Temchin, Rich and Ruggero, 2008).

The neuron of Figure 3.29, has a CF of 6091 Hz. The tuning curve shows similar characteristics to those of the nerve fibre discussed above, including a rapid increase in threshold with increase in stimulation frequency above CF, and a slow increase in threshold with decrease in stimulation frequency below CF.

The slower roll-off towards the apex (lower frequencies) may be related to the cochlea travelling wave in which an increase in stimulation intensity causes an increase in the width of the area of the basilar membrane deflection towards the base of the cochlea (as seen in Figure 3.12); this would contribute to lower stimulation frequencies causing the stimulation of nerve fibres with higher CF to be stimulated above spontaneous rate.

It is important to note that tuning curves shown here are not for nerve fibres that were designed to have specific tuning curves. All nerve fibres are generic fibres that differ only in spontaneous rate and threshold. The observed tuning curves result from the travelling wave displacement pattern, where nerve fibres are stimulated across a large extent in the cochlea as the travelling wave travels towards its point of peak displacement.

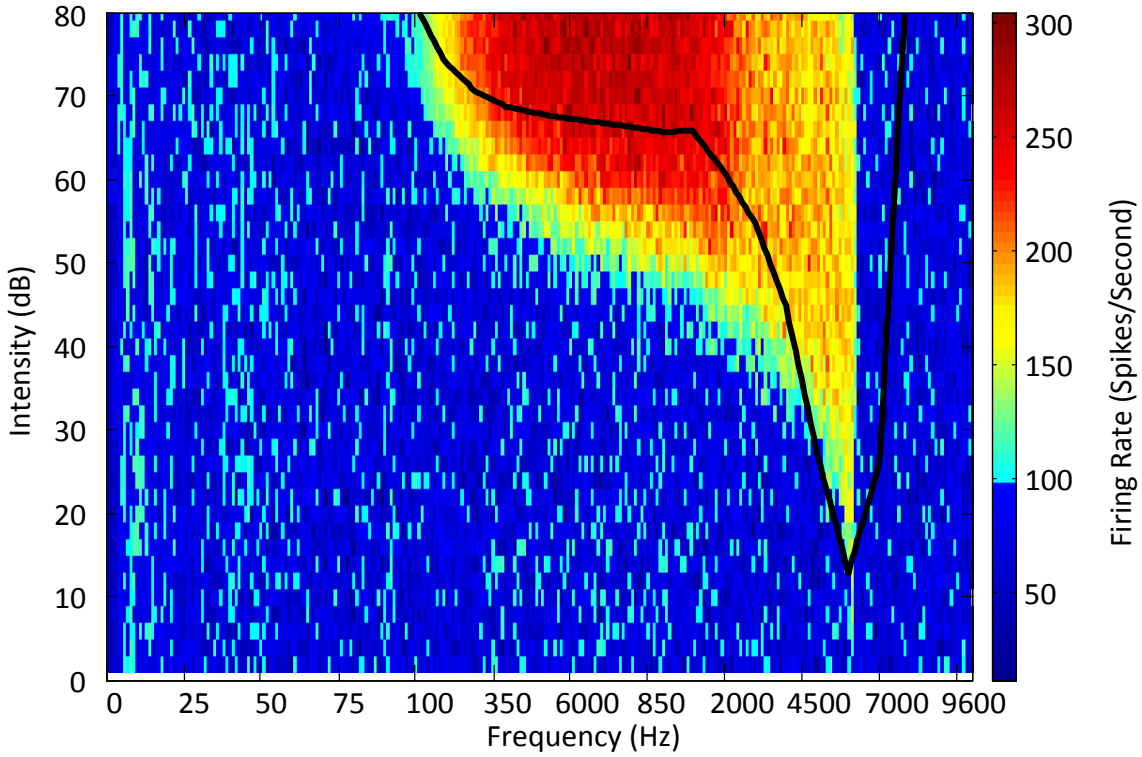

Figure 3.29. Nerve fibre tuning curve with a CF of 6091 Hz. The solid line represents the nerve fibre tuning curve measure data (Temchin et al., 2008).

In these figures, the predicted nerve fibres are compared to measured chinchilla auditory nerve data (Temchin et al., 2008). Measured chinchilla tuning curves are less steep in the roll-off towards the higher frequencies of stimulus. Two stages of roll-off in tuning curve threshold can be observed for both the chinchilla data and the model predictions below CF: a steep increase in threshold as stimulus frequency decreases, followed a more gradual increase in threshold beyond that.

3.3.6 Nerve fibre spiking patterns for missing fundamental signals

As a further test of the model, this section shows the travelling wave, nerve fibre spiking diagrams, and ISI histograms for a more complex signal. Of course, any signal can be represented (through Fourier analysis) as a sum of sinusoids. The purposes of the present

study called for the use of a sum of harmonically-related pure tones missing the fundamental frequency. A signal with a fundamental frequency (f0) is constructed using multiple harmonics of the fundamental frequency, but excludes the fundamental. Figure 3.30 shows the construction of the signal used as the stimulation signal. The multiples of the fundamental frequency are added together to get a cosine missing fundamental signal. The period of the stimulation signal still reflects the fundamental frequency.

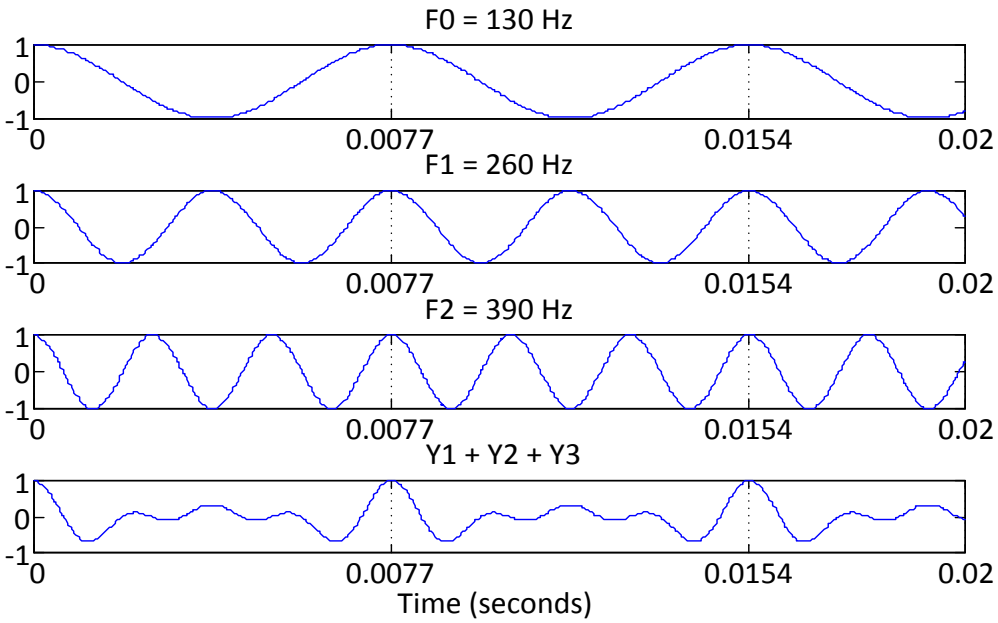

Figure 3.30. Construction of the missing fundamental signal.

Figure 3.31 shows the prediction of the basilar membrane displacement due to the missing fundamental stimulation signal and compares that stimulation with the fundamental. The basilar membrane has two maximum peaks during one period of the fundamental frequency as seen in Figure 3.33. However, the amplitude of the basilar membrane displacement is lower for the missing fundamental stimulation signal for the same sound pressure level.

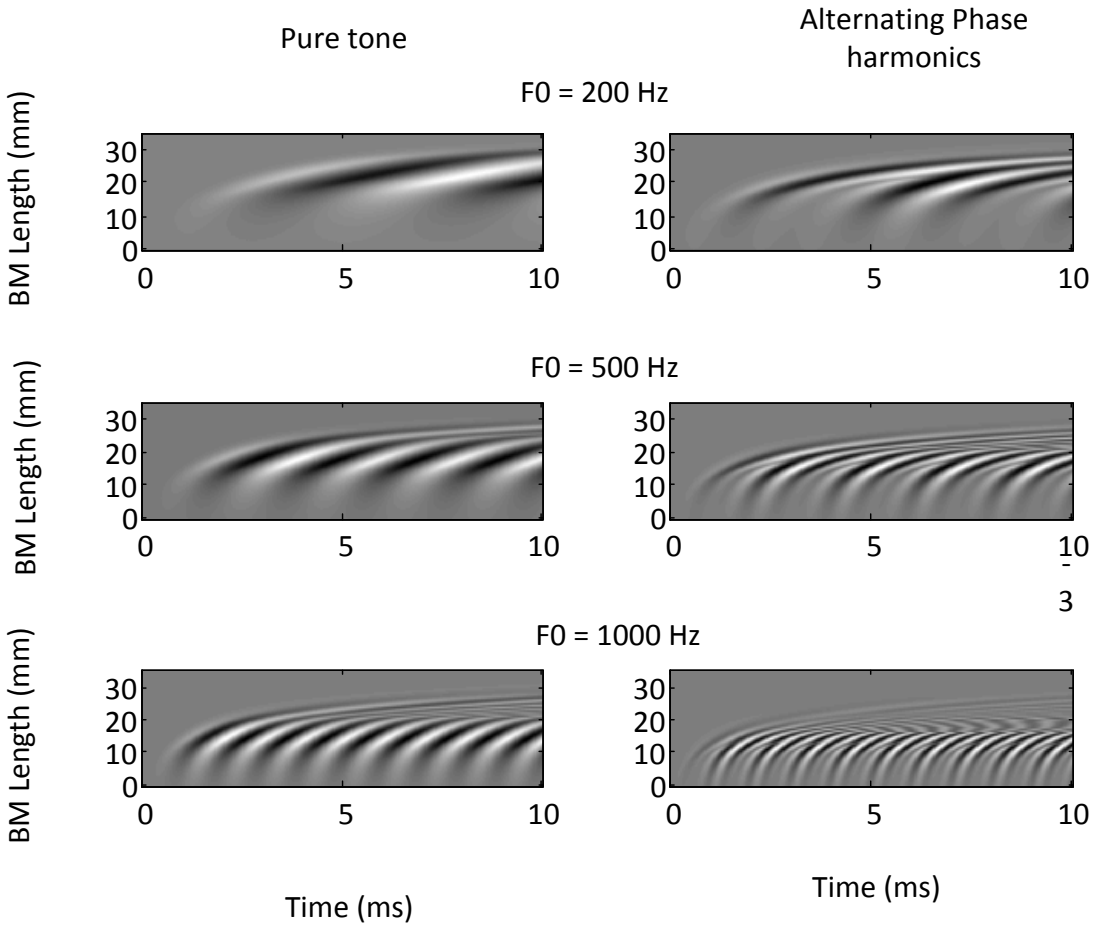

Figure 3.31. Travelling wave of an alternating phase missing fundamental stimulation compare to a travelling wave of a pure tone. Alternating phase and consine phase are compared in Figure 3.32.

Figure 3.32 shows the comparison between the pure tone, alternating phase missing fundamental, and cosine phase missing fundamental for a 500 Hz signal. The basilar membrane displacement is lower for the two missing fundamental stimulation signals. The difference between the alternating phase and the cosine phase is the spacing in time (similar to the phase shift) of the travelling waves moving from the base to the apex. The spacing is seen in multiples of the signal which construct the missing fundamental signal.

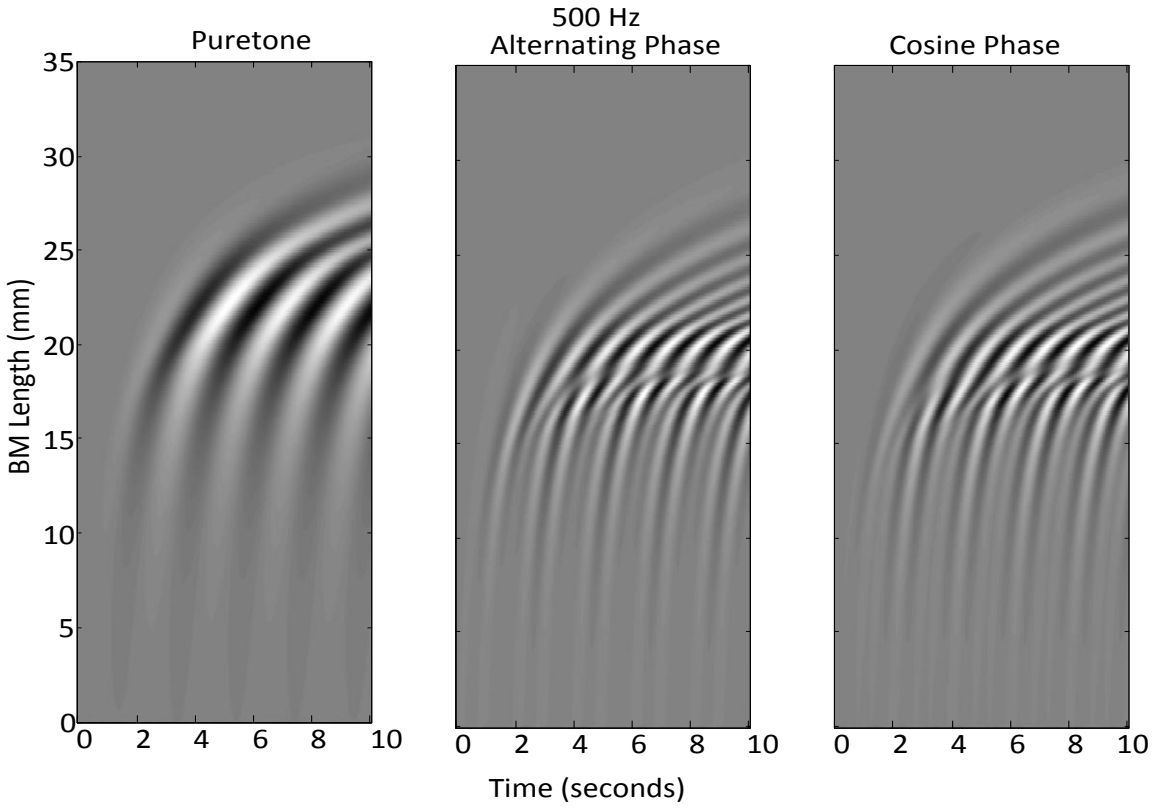

Figure 3.32. Basilar membrane response from pure tone and missing fundamental response of alternating phase and Cosine phase for 500 Hz fundamental signal.

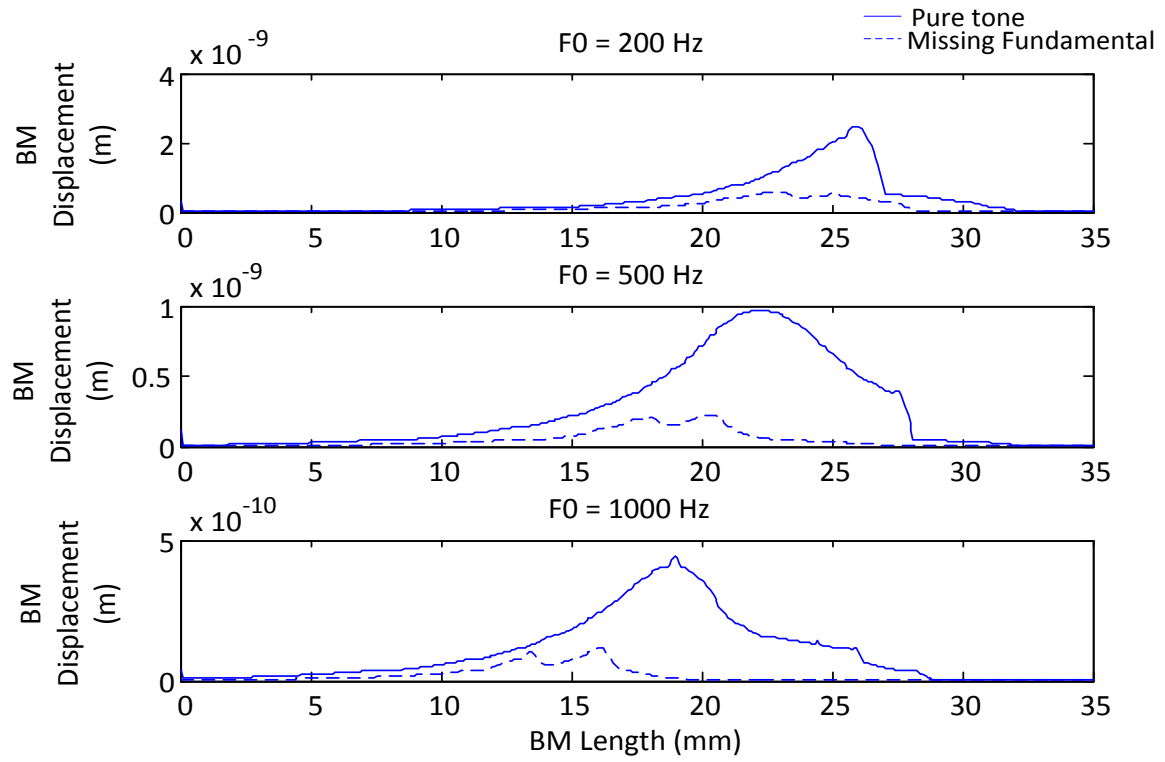

Figure 3.33. Pure tone basilar membrane maximum displacement versus missing fundamental maximum displacement.

The ISI histograms produced by the missing fundamental signal are shown in Figure 3.34 and Figure 3.35. The tallest peaks of the ISI histogram have a period of the fundamental frequency and secondary lower peaks are around half the period of the fundamental frequency. The cosine phase missing fundamental stimulation signal is compared to data from literature (Cedolin and Delgutte, 2005) in Figure 3.35. The predictions and the data from literature show similar characteristics. According to these ISI histograms, the auditory system has enough information available to extract the missing fundamental pitch. As shown in Chapter 5, a pitch extraction algorithm will also extract the pitch of the missing fundamental.

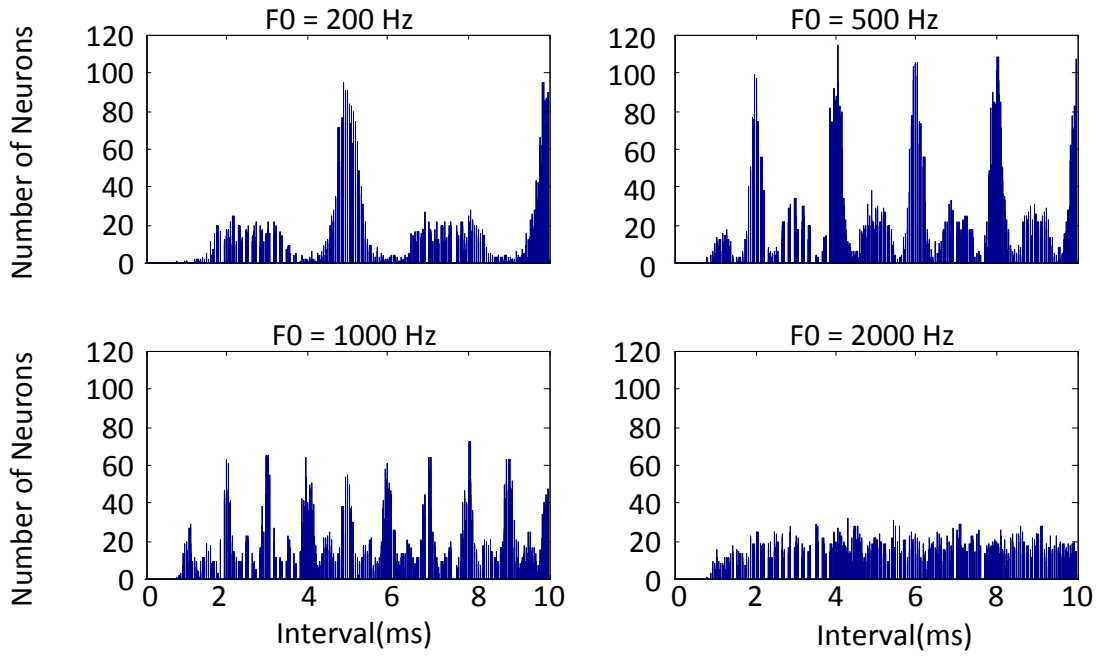

Figure 3.34. ISI histograms for alternating phase missing fundamental frequencies signals.

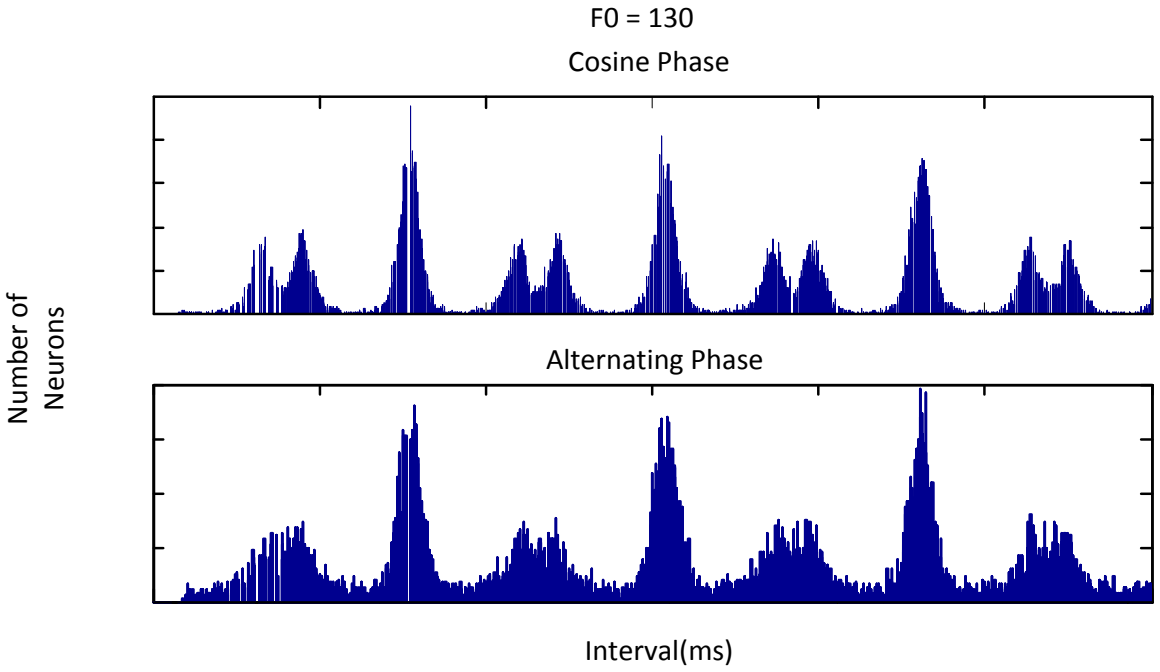

Figure 3.35. Comparison between cosine phase missing fundamental stimulation and an alternating phase missing fundamental stimulation for a fundamental frequency of 130 Hz. Model predictions are shown these measurements are comparable to those from Cedolin and Delgutte (2005).

3.4 DISCUSSION

3.4.1 Travelling wave mechanics

The implemented model of the travelling wave based on the Duifhuis (2012) and Duke and Jülicher (2003) nonlinear, active, and one-dimensional models, is able to reproduce characteristics of measured data found in literature (Rhode, 1971, Rhode et al., 2010). These characteristics include the active gain of the model seen during low level stimulation (Figure 3.10, and Figure 3.11), the increase in linearity and decrease of gain with the increase of sound pressure level (Figure 3.10 and Figure 3.11), the travelling wave delay from the base to the apex of the cochlea (Figure 3.9), the tonotopical arrangement of the frequency spectrum along the length of the cochlea (Figure 3.8), the broadening of the travelling wave with increase in stimulation loudness level (Figure 3.11 and Figure 3.12), and the shift in peak excitation towards the base of the cochlea with increase in stimulation level (Figure 3.12).

This chapter's contribution is not in the development of original models for the travelling wave or neural excitation, but in (1) the careful selection and adaptation of existing models, and (2) the combination of the two. Specifically, a unique viewpoint of the present model lies in the use of a travelling wave model as input to an auditory nerve fibre model. Auditory models often focus on the implementation of cochlea filtering characteristics using filter functions (Carney, 1993, Zhang et al., 2001, Moore and Glasberg, 2004), ignoring the role of the displacement of the basilar membrane resulting from travelling waves. However, the mechanical response of the basilar membrane could be responsible for many of the characteristics of nerve fibre firing (Rhode, 2007, Cheatham, 2008, Temchin, Recio-Spinoso and Ruggero, 2011). A particularly important finding from the results shown in this chapter is that a cochlea model does not require explicit filtering to be able to predict nerve fibre firing patterns that show filtering characteristics. Specifically, observed tuning curves are probably a direct consequence of the travelling wave.

The one dimensional travelling wave model has shown good reproduction of travelling wave characteristics. However, it has been previously mentioned that a one dimensional travelling wave model is mainly adequate for qualitative purposes (Viergever and Diependaal, 1983). It will not accurately enough predict the BM displacement of the travelling wave for an input stimulation. However, the present research focus is on the delivery of spatiotemporal information to auditory nerve fibres using travelling wave

characteristics, which the model is able to represent (Duifhuis, 2012), and application of this to cochlear implant stimulation (considered in chapter 4).

3.4.2 Neural spike train patterns

The implemented nerve fibre model is able to model low, medium, and high spontaneous rate nerve fibres, which correlates to nerve fibres seen in measured data. In the spike train patterns, during pure tone stimulation, it can be observed that nerve fibres lock onto the basilar membrane displacement characteristics: following of the travelling wave delay from the base to the apex to the cochlea, and the broadening of the nerve fibre locking onto the stimulation period of the pure tones (as seen with the ISI histogram in the results).

The synchronisation of nerve fibres increase with pure tone stimulation intensity level. Increase in the stimulation level causes nerve fibres at the base of the cochlea to synchronise to the pure tone stimulation (Johnson, 1980); this correlates to the basilar membrane displacement increasing in width of displacement area with increased loudness level, seen in Figure 3.15, which causes a broader range of nerve fibres to fire.

The increase in synchronisation of nerve fibres along the length of the cochlea may enable temporal encoding to be dominating at higher loudness levels, in contrast to lower loudness levels, where place encoding is prominent. Loudness information may also correlate to the number of synchronised nerve fibres along the length of the basilar membrane and not the increase in firing rate of the nerve fibres (Chatterjee and Zwislocki, 1998). The temporal information available can be seen by the ISI histograms. The predicted ISI histograms show a strong correlation to the temporal pitch of the pure tone by locking onto the period of the pure tone. An increase in intensity level of the pure tone results in an increase in the peak height and peak sharpness around the period of the pure tone. This is all further explored in chapter 5 with the use of temporal pitch prediction algorithm.

An important observation from this chapter is that tuning curves result from travelling wave characteristics. This relationship between the basilar membrane displacement and the auditory nerve fibre firing patterns suggests that encoding of sound information in the nerve fibre spiking patterns mostly reflects the travelling wave of the basilar membrane (i.e. while the travelling wave is often ignored, the present work suggests that its role in cochlea filtering has been underestimated). This suggestion is supported by the observed role of octopus cells in the cochlea nucleus, which are sensitive to the travelling wave characteristics during stimulation (McGinley et al., 2012). This suggestion also supports

the viewpoint of Greenberg (1997) who noted the importance of the travelling wave in pitch perception. As will be seen in later chapters, it appears possible to incorporate travelling wave characteristics in CI speech processors to obtain more natural nerve firing patterns with CIs.

3.5 SUMMARY

This chapter describes the development of a normal hearing model. The normal hearing model consists of a middle ear, travelling wave hydrodynamic model, IHC model, IHC-AN synapse model, and spike generation model. The auditory model developed benefits from a travelling wave model, which is closer to the physical reality of basilar membrane displacement than filter based models. This development allows for the effects of the travelling wave characteristics on nerve fibre spiking patterns to be studied and to be applied. The results show that the model can predict many characteristics of basilar membrane displacement, as well as nerve fibre data. This model, developed for normal hearing, will now be applied to cochlear implants in the chapters to follow.

CHAPTER 4 IMPLEMENTATION OF A TRAVELLING WAVE MODEL FOR COCHLEAR ELECTRICAL STIMULATION

4.1 CHAPTER OBJECTIVES

The previous chapter used a normal hearing model to gain insight into the delivery of spatial-temporal information to the nerve fibres. Chapter 3 laid the foundation for Chapter 4 to focus on the delivery of spatial temporal information in cochlear implants.

The chapter focuses on the development of a travelling wave electrical stimulation method that may be used in cochlear implant speech processors and the plausibility of a travelling wave stimulation method for cochlear implants by investigating the travelling wave electrical stimulation method in a computational model. The computer model predicts and then decodes the nerve fibre action potential patterns to see if useful spatial-temporal information will potentially be available to the cochlear implant user.

The approach of a computational model is used to gain an understanding of how electrical stimulation will affect nerve fibre firing patterns without having to measure nerve fibre responses. This will allow for the validation of the potential advantages of travelling wave electrical stimulation before moving forward to animal model experiments.

4.2 OVERVIEW OF MODELS

4.2.1 Vocoder model overview

The block diagram in Figure 4.1, describes a model for a typical vocoder speech processor (e.g., the ACE speech processing strategy used in the Nucleus speech processor) that may be used to gain insight into the nerve response characteristics expected during vocoder-type cochlear implant stimulation. The development of a vocoder model allows the investigation of the spike train patterns produced by the electrical stimulation firing patterns. These results from the electrical stimulation model will be compared to the normal hearing results and electrical stimulation results to understand the delivery of speech cues in current cochlear implants.

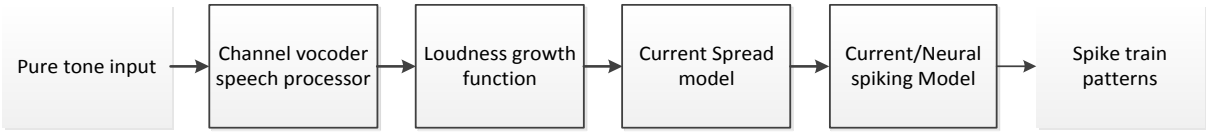

Figure 4.1. Model of a vocoder-type CI speech processor used to predict spike trains which are expected in cochlear implant users.

The vocoder speech processor model shown in Figure 4.1 above is used to predict the spike train patterns measured in cochlear implant users, specifically for pure tone and tone complexes with missing fundamental stimuli in the work described here. The pure tone predictions were used to explore if cochlear implants can potentially transfer pitch using temporal encoding of nerve fibres. Missing fundamental signals were used to observe if pitch could be encoded using signals with a more complex nature. Complex tones are a combination of pure tones combined and can be decoded into the fundamental pure tones using Fourier analysis. This study may then utilize pure tones, assuming that if nerve fibres can decode pure tones, the nerve fibre decoding for complex signals would hold. This assumption was then validated used missing fundamental signals. The model uses a channel vocoder speech processor to generate the electrical stimulation pattern across the electrodes of the electrode array which is implanted in the human ear. The electrical stimulation pulses current amplitudes that are calculated from acoustic intensity (in dB SPL) by a loudness growth function (LGF) (Nogueira et al., 2005) based on just audible threshold and the user's comfortable current level (measured in CL or current level units), which compensates for the logarithmic increase in perceived loudness as a result of the increased stimulation current.

With the electrical current levels of the stimulation pulses known for each electrode (generated in response to a sinusoidal input signal or tone complex for this model), current spread is predicted. The current spreads away from each electrode, so that interaction occurs on the nerve fibre plane (i.e. current pulses from multiple electrodes may interact at a specific neural population to give an effective stimulus at this point). This interaction takes place in space (as current spreads along the cochlea canals), but also in time (as pulses are non-simultaneous, at least in Cochlear devices' processing algorithms). A simple model of this interaction was used to predict the stimulation of the individual nerve

fibres along the length of the electrode array inside of the cochlea (Strydom and Hanekom, 2011).

A fibre activation model, in response to electrical stimulation (Bruce et al., 1999a, Bruce et al., 1999b), then predicts nerve fibre spiking patterns from the stimulation pulses and the current spread along the length of the cochlea.

4.2.2 Travelling wave model overview

This section gives an overview of the travelling wave model, developed and adjusted from the TW model of chapter 3 to be suitable for application in cochlear implant speech processors. The block diagram in Figure 4.2 below, summarises the travelling wave model implementation. The model is used to predict the spike train patterns that are expected to result from a TW-based speech processor.

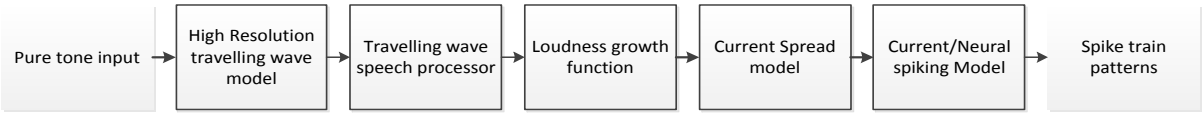

Figure 4.2. Cochlear implant travelling wave stimulation model diagram.

A pure tone input is used to predict the expected output spike train patterns resulting from an electrical stimulation. A high resolution travelling wave model (first processing step) predicts the displacement of the basilar membrane. This hydrodynamic travelling wave model from chapter 3 (non-linear, active, and one-dimensional) is used as the input into a travelling wave speech processor which determines the stimulation pulses that should be applied on each electrode along the length of the electrode array. The stimulation mechanics will be discussed further on in this chapter in more detail.

The rest of the steps are the same as for the vocoder-based model. Stimulation pulses, generated from the travelling wave speech processor, are then used to calculate the current level amplitude by a loudness growth function (LGF) based on just audible threshold and the user's comfortable current level (measured in CL or current level units). As before, current spread is predicted using the model of Strydom and Hanekom (2011) and the electrical stimulation model (i.e. the neural spiking model of (Bruce et al., 1999a, Bruce et al., 1999b), final processing step in Fig 28) predicts the nerve fibre spiking patterns resulting from electrical stimulation.

4.3 DETAILS OF MODELS

This sections gives details of the processing steps of the vocoder model (Figure 4.1) and the TW model (Figure 4.2). Where these models use identical processing blocks, these are described only once.

4.3.1 Channel vocoder speech processor

The channel vocoder speech processor uses the ACE processing strategy as implemented on a Nucleus-24 device (Loizou, 2006). The ACE processing strategy estimates the input spectrum of the stimulus signal using a Fast Fourier Transform (FFT). The FFT is then filtered into n (12 – 22) filter bands. The filter bands are spaced linearly from 188 Hz to 1312 Hz and then logarithmically spaced up to 7938 Hz. The n envelopes are estimated by summing the power of the FFT bins within the filter band. The amplitude (m) bands are then selected in each stimulation time frame. A maximum of 8 – 12 maximum amplitudes are chosen for stimulation.

The stimulation rate of the electrodes can be in the range from 250 pulses per a second (pps) to 2400 pps and is limited by the combined maximum stimulation rate of 14400 pps across all channels. For this study the jitter of the stimulation pulses was chosen to be constant.

4.3.2 High resolution travelling wave model

The displacement of the basilar membrane is predicted using the 1D, nonlinear, active travelling wave model from Duke and Jülicher (2003). This included the middle ear model from O'Connor and Puria (2008). The travelling wave model is implemented in the way described in Chapter 3 for normal hearing. The implementation is exactly as described in Chapter 3 and none of the details are repeated here. However, an example of the output is given in Figure 4.3.

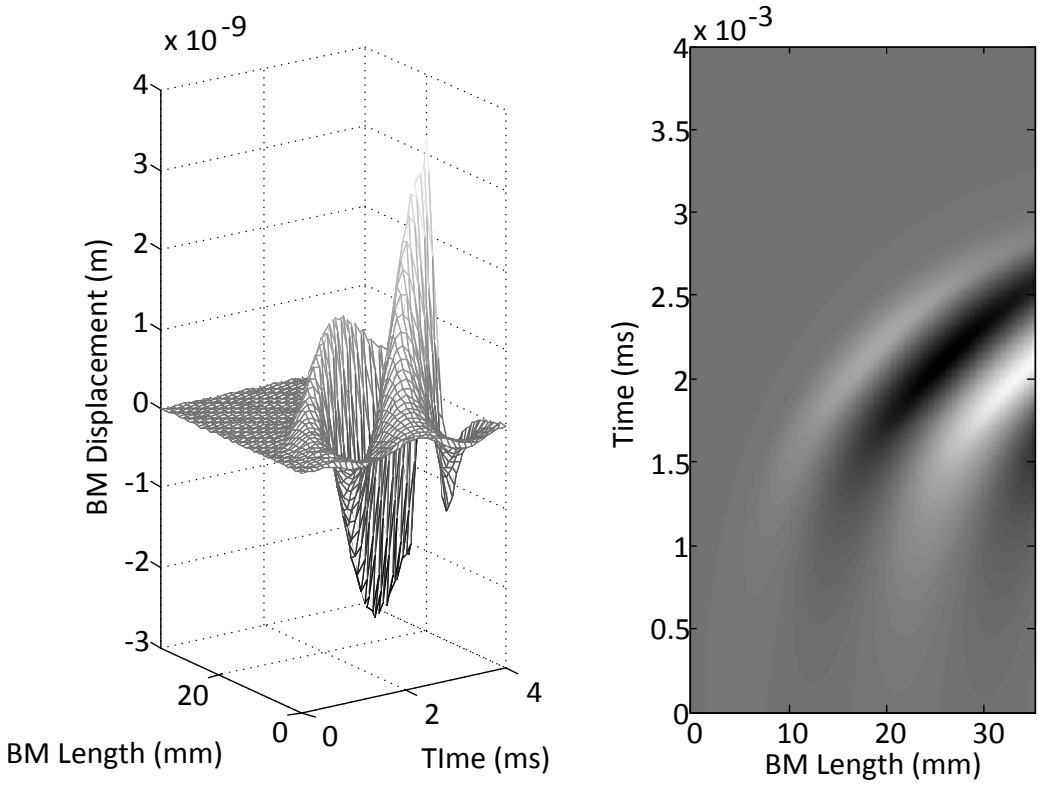

Figure 4.3. Output of the travelling wave model. More results are illustrated in the previous chapter.

4.3.3 Travelling wave speech processor

To implement a travelling wave speech processor requires a lower resolution (in space and in time) than that of the high resolution hydrodynamic travelling wave model of the previous processing step. An earlier model developed by Du Preez (2011) for TW stimulation used a decimation method in space and time to match the available resolution of a particular CI speech processor (Nucleus) to generate the stimulation profile for each electrode.

A trade-off between spatial and temporal resolution exists. The spatial sampling of the high resolution TW model was down sampled to 22 equally spaced positions, reflecting electrode contacts that are on average 0.75 mm apart (simulating a Nucleus CI24 cochlear implant, (Clark, 2003) for a distance of 25 mm electrode length. The basilar membrane displacement, as predicted at these discrete positions in the high resolution TW model, would then be mapped to the relevant electrode, resulting in stimulation amplitudes on the electrode that are related to the basilar membrane displacement.

The time sampling was down sampled to 14400 pps, and this sampling rate was spread across the electrodes (i.e. the maximum sampling rate per electrode was dependent on the number of electrodes used for stimulation). If a fixed number of 8 electrodes were in use, the maximum sampling pulse rate was 1800 pps (Figure 4.4 a) on each electrode; and when 5 electrodes were in use, maximum sampling pulse rate was 2880 pps (Figure 4.4 b). Clearly, a trade-off exists here: if fewer electrodes are used, the spatial resolution is lower while the temporal resolution is higher.

Finally, the infinite resolution amplitude of the high resolution TW model is mapped to stimulation pulse amplitudes with a resolution of 8 bits (256 current levels) to reflect the Nucleus device's resolution.

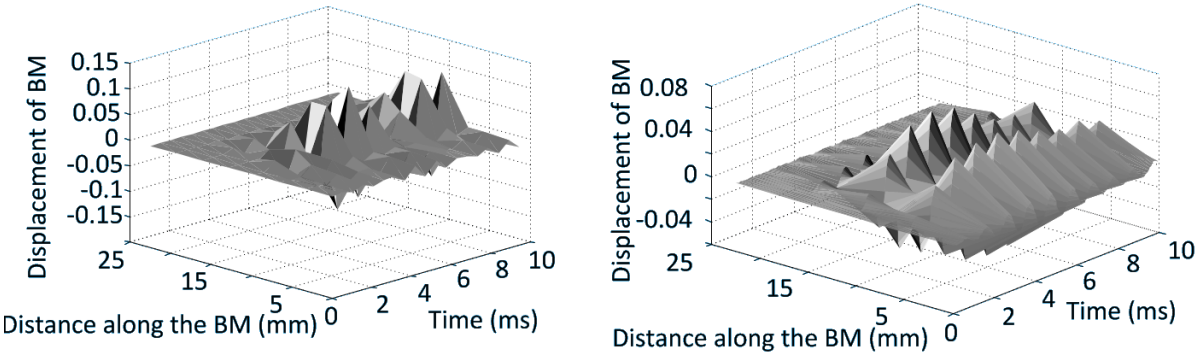

Figure 4.4. Downsampling of the travelling wave of 1000 Hz pure tone which has been down sampled to (a) 8 electrodes, 1800 pps/e and (b) 5 electrodes, 2880 pps/e. The x-axis is the distance from the base of the cochlea.

Figure 4.5 and Figure 4.6 show the spatial and time resolution trade-off for multiple stimulation pulse rates (Figure 4.5) and different electrode spatial resolutions. Decreasing the pulse rates causes a distortion of the travelling wave peaks and ridges along the length of the basilar membrane. Decreasing the number of electrodes (Figure 4.6) affects amplitude greatly, until the travelling wave is not distinguishable when decreased to 5 electrodes.

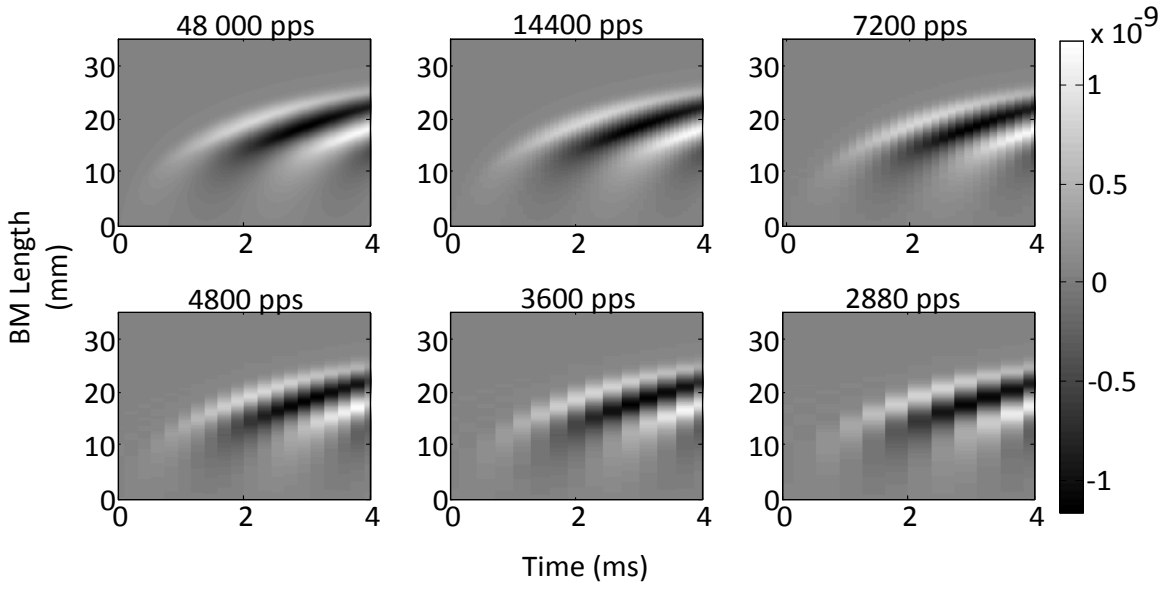

Figure 4.5. Basilar membrane displacement predictions from the travelling wave model and the effects of down sampling temporal resolution.

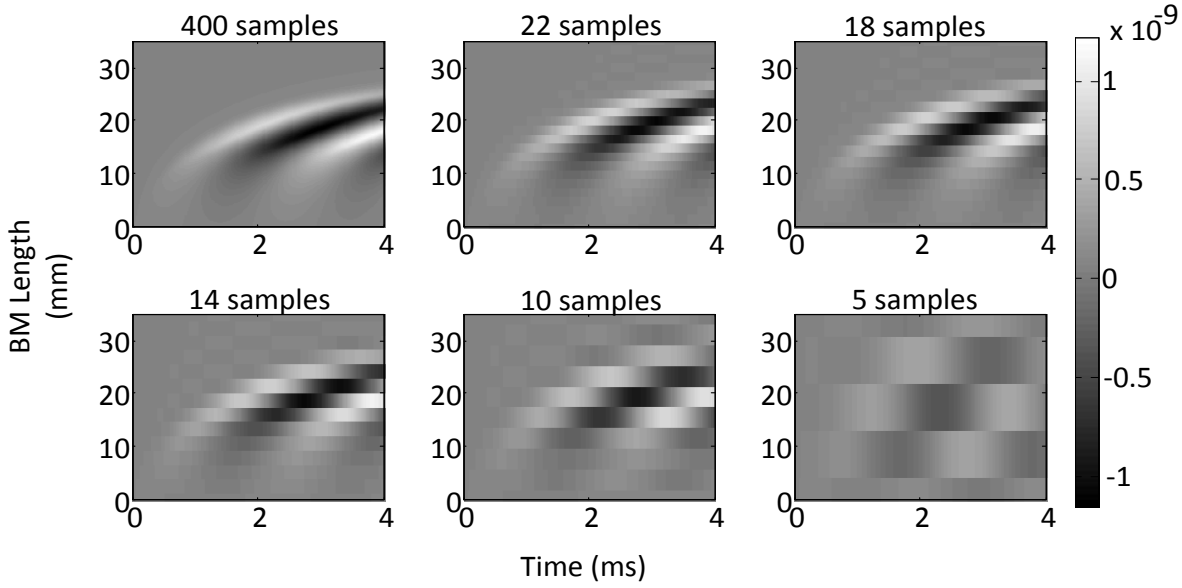

Figure 4.6. Basilar membrane displacement predictions from the travelling wave model and the effects of down sampling spatial resolution.

Because of the trade-off mentioned above, the representation of the travelling wave is highly dependent on the choice of the number of electrodes and the sampling rate per electrode. An incorrect choice of the spatial and time sampling could distort the travelling wave representation. This is explored further in the results section.

4.3.4 Alternative travelling wave peak stimulation model

Because of the space-time tradeoff resulting from resolution limitations in a CI speech processor, an alternative approach is proposed here. Rather than to attempt to recreate the travelling wave's complete spatial extent in the cochlea, the proposed alternative is to let stimuli track only the peak displacement of the basilar membrane as the travelling wave propagates along the BM. The peak of the travelling wave during stimulation moves from the basal end of the cochlea to the maximum point of deflection (which correlates to the place frequency map of the basilar membrane) and quickly decreases in amplitude towards the apex of the cochlea.

In this alternative, the high resolution travelling wave model (400 kHz sampling rate), is still down sampled in both temporal and spatial resolution. Only one electrode will be stimulated and the position of this active electrode will depend on where the basilar membrane displacement is greatest at a specific time. Amplitude of stimulation will depend on amplitude of the TW at the particular place and instant of sampling. Spatial resolution is limited to 22 places (22 electrodes), but time resolution is as high as is allowed by a particular cochlear implant speech processor's highest pulse rate. In this work, this was selected to be 14400 pps, corresponding with the Nucleus cochlear implant's maximum rate.

This alternative travelling wave peak stimulation method of following the peak displacement of the basilar membrane proposes elimination of the sampling of the entire surface of the travelling wave (spatial and temporal), resulting in low temporal and spatial resolution. The method will down sample the travelling wave into 22 samples along the length of basilar membrane and then compress this to fit the total length of the electrode array. The compression of the travelling wave is done to include the entire travelling wave in the stimulation profile, as the electrode array is only 25 mm long. This work focuses on the delivery of temporal information to the auditory nerve fibres; by compressing the travelling wave along the length of the basilar membrane, nerve fibres will be exposed to the entire travelling wave stimulation and possibly more temporal information.

4.3.5 Loudness growth function model

The loudness growth function maps the stimulation amplitude to the dynamic range of the cochlear implant user. The dynamic range is measured between the electrical threshold

level and comfort level of the cochlear implant user (Loizou, 2006). The ACE strategy uses a logarithmic shaped function to map the acoustic stimulation amplitude from the filter, or travelling wave amplitude ($a(z_i)$), to electrical stimulation (Nogueira et al., 2005), and this function was used in the present implementation (Du Preez, 2011). The function is described by:

$$q(z_i) = \begin{cases} \frac{\log(1+\rho(\frac{a(z_i)-s}{m-s}))}{\log(1+\rho)}, & s \leq a(z_i) \leq m, \\ 0, & a(z_i) < s, \\ 1, & a(z_i) \geq m \end{cases} \tag{4.1}$$

where q is the output in the range between 0 and 1, corresponding to the current level values between threshold level (T) and the comfort level (C) of the cochlear implant user, z_i is the channel number, and a is the input simulated BM response amplitude. Inputs lower than the base level (s) are mapped to the T level and inputs higher than the saturation level (m) are mapped to the comfort level (C). The parameter ρ is the steepness factor, which controls the steepness of the curve of the LGF which is described by Swanson (2005) in the Nucleus MATLAB toolbox. The ρ is calculated for each, from the Nucleus MATLAB toolbox, minimum and maximum input current values used in the current map of a cochlear implant user.

The output of the LGF function ($q(z_i)$) in equation 4.1 is a linearly mapped to current level units (CL) from 0 to 255 in the Nucleus 24 processor. The loudness growth function compensates for the logarithmic increase in perceived loudness with electrical stimulation current. The current on each channel (I_{CL}) is described by equation 4.2:

$$I_{CL} = 10\ x\ 175^{\frac{CL}{255}}, \tag{4.2}$$

where I_{CL} is a minimum of 10 µA and a maximum of 1.75 mA. Equation 4.2 was implemented in the third generation cochlear implants. The range was from 0 for T level to 255 for C level for this study.

4.3.6 Current spread model

Current spread refers to the spread of current from the stimulation electrode to the surrounding nerve fibres. Current spread is described by the decay of electrical stimulation as the distance increases from the electrode's surface. The decay in current due to monopolar electrode stimulation was found to have a mean of around 2.8 dB/mm, whereas a bipolar electrode array will have a mean decay in current of around 7.4 dB/mm (Bingabr et al., 2008).

The current as a result of current decay is described in the following equation (Smith, 2011, Strydom and Hanekom, 2011) :

$$I_{eff} = I10^{\left(\frac{-\Delta d\delta}{20}\right)}, \tag{4.3}$$

where I_{eff} is the effective stimulation current at the distance Δd (in mm) from the stimulation electrode that has stimulation current I. δ is the current decay rate in dB/mm, and a value of 2.8 dB/mm was used in this study, corresponding to monopolar stimulation (Bingabr et al., 2008). An overlap in stimulation current from different electrodes reaching each nerve fibre is observed due to current spread. To calculate the total effective stimulation current at a distance from the stimulation electrodes, the total current arriving at a particular neural position is the sum of the current from the all the electrodes:

$$I_{eff}(c) = \sum_{z=1}^{Z} I(z)\, 10^{\left(\frac{-\Delta d_c(z)\delta}{20}\right)}, \tag{4.4}$$

where I_{eff} is the total effective current due to stimulation current from all electrodes, I(z), at the different distances, Δd. Index z enumerates the electrodes, and there are Z electrodes contributing to the current at position c. Δd_c is the distance between electrode z and position c on the neural plane.

4.3.7 Current/Neural spiking Model

Relating the electrical stimulation current to nerve fibre spikes uses a stochastic computational model, which was implemented by Smith (2011). Smith (2011) modified the model from Bruce et al. (1999b) to use currents instead of potentials to describe the stimulation of the nerve fibres. The details of the model used in this work can be found in Smith (2011), which is the implementation of the current to neural spiking model. The parameter used in the model is found in Table 3.

The stimulation current output from the CI processors is used to predict the nerve spiking patterns. The model compares the stimulus current I_{eff}, resulting from the input pulse train, with the summation of the threshold, noise, and refractory currents ($I_{thr}+I_{noise}+I_{refr}$). If the stimulus current is greater than or equal to the summed threshold, noise, and refractory currents, then neural firing occurs. This comparison is performed at every time sample in the cathodic phase of the biphasic stimulus pulse. I_{refr} is the refractory period where nerve fibres are unable to fire unless the threshold current is met. The refractory period is separated into two parts. The first is 0.7 ms where the refractory current is at infinity and the second is the exponential decay of the current which has a time constant of 1.32 ms. The probability of firing occurring will increase during the second period. The noise current I_{noise} affects the threshold voltage of the nerve fibre. The change in threshold voltage ($V_{Threshold}$) is caused by the noise current which increases the stochastic nature of the firing times for the nerve spiking times.

Table 3. Parameter table showing the refractory period values used in Current/Neural spiking Model which is the parameters from Smith (2011).

Parameter	Value
Refractory period absolute (I_{refr})	0.7 ms
Refractory period exponential decay	1.32 ms
Refractory noise variance (V_{noise})	7.4141 uV

4.4 RESULTS

This section documents predictions from the cochlea electrical stimulation models, i.e. the CI vocoder speech processor, and the two alternative TW speech processors. The nerve fibre space-time firing characteristics from the electrical stimulation model will be compared to the results of the normal hearing model obtained in Chapter 3.

The electrode placement for the simulations run in this chapter are given in Table 4 below.

Table 4. Electrode placement inside of the cochlea. Measurement from the base of the cochlea to the electrode position.

Electrode Number	Electrode position (mm)
1	10
2	10.75
3	11.5
4	12.25
5	13
6	13.75
7	14.5
8	15.25
9	16
10	16.75
11	17.5
12	18.25
13	19
14	19.75
15	20.5
16	21.25
17	22
18	22.75
19	23.5
20	24.25
21	25
22	25.75

4.4.1 Comparison between normal hearing and CI stimulation

The two proposed travelling wave speech processor models are compared to the ACE cochlear implant speech processing (by Cochlear) and the normal hearing model. The ACE CI strategy was implemented in the model using different pulse rates of 600 pps, 1200 pps, and 3600 pps. The model output is a space-time spike train pattern, as shown in figures that

follow. From these spiking diagrams, it is seen that the model predicts that the nerve fibres fire in a specific area along the length of the cochlea for a specific frequency, as expected. An increase in frequency causes the place of stimulation to move towards the base of the cochlea, as expected. This is shown in Figure 4.7. The shift in place of stimulation is relatively small as the spacing between electrodes is only 0.75 mm. For example, the electrode closest to the 200 Hz cochlea position is electrode 19, while the electrode closest the 500 Hz cochlea place is electrode 16, which is a shift of 2.25 mm.

The travelling wave speech processors are subject to the previously mentioned space-time resolution trade-off. When allowing for 22 electrodes to be active and the sampling rate to be 654 pps of each electrode, it is observed that the spike train pattern attempts to follow the travelling wave. However, it's evident that in Figure 4.7 for the TW speech processor (TW), there is under sampling of the travelling wave, seen by the nerve fibres firing patterns not being continuous along the length of the basilar membrane as would one would expect to accurately reflect basilar membrane displacement.

The nerve fibre spiking diagrams for the alternative travelling wave (TW A) processor with 22 electrodes active (Figure 4.7) show that the predicted spike train patterns follow almost the entire travelling wave of the first period of the stimulus. However, as the stimulus continues, the travelling wave is only tracked for a shorter length of the basilar membrane and as the frequency increases, this tracking distances decreases.

In all three CI speech processors shown in Figure 4.7, the maximum distance away from the base of the cochlea which can be stimulated is at electrode 22 (25 mm in this model).

Insight into the delivery of spatio-temporal information can be obtained from ISI histograms, produced from the nerve fibre firing diagrams (Figure 4.8) by collapsing the space dimension of the spiking diagrams in Figure 4.7. These ISI histograms essentially ignore any place information and consider the information available in temporal aspects of neural firing patterns. The ISI histogram for NH shows that spike firings lock onto the period of the stimulus, i.e. the period of the pure tone signal which was 5 ms, 2 ms and 1 ms respectively in the predictions shown in this figure. However, note that ISI histogram peaks for the ACE processing strategy (a vocoder strategy) cluster around multiples of the electrical stimulation period, which was 0.83 ms (stimulation rate of 1200 pps) in this example. This characteristic is also observed in Figure 4.10, where the stimulation rate is

600 pps, and the peaks are clusters around 1.67 ms and the multiples of the stimulation pulse rate.

The results suggests that vocoder-type (or ACE-like) processing strategies do not capture periodicity information of the acoustic signal or do not elicit neural phase locking that is related to the characteristics of the acoustic signal. This is an important conclusion that will be explored further in the discussion.

Similarly, the TW strategy as originally used by Du Preez (2011), does not result in ISI peaks clustered around the pure tone period for any of the frequencies used in this simulation when all 22 electrodes are active (Figure 4.8 third column (TW)). Due to the highest spatial resolution (22 electrodes) being selected , the temporal resolution was probably too low to allow pulsatile stimuli to be placed accurately enough in time to reflect periodicity information that was captured in the travelling wave. However (Figure 4.9 and Figure 4.10), when fewer electrodes are active (a five-electrode stimulus using electrodes 2, 6, 10, 14, and 18), which permits an increases the maximum temporal sampling rate on each electrode and decreases the spatial resolution, ISI peaks (Figure 4.10) can be observed around the periods of 200 Hz (5 ms) and multiples of 500 Hz (4 ms, 6 ms) stimuli. There are no apparent ISI peaks around multiples of the period (1 ms) of the 1000 Hz stimulus, and stimulus peaks that do appear are generally spaced closer than 1ms apart.

The alternative travelling wave speech processors' predicted spike train patterns in Figure 4.7 show scattering smaller distribution of spikes along the length of the basilar membrane. The nerve fibres follow the travelling wave from base to the point of maximum stimulation. However, after the first travelling wave, it's observed that the nerve fibre firing range is shortened along the length of the basilar membrane. It's shortened because the first travelling wave peak displacement is larger than the displacement of the start of the second travelling wave and the basilar membrane deflection is much lower at the base in comparison to deflection closer to the maximum point of stimulation for a pure tone at the same instant of time.

Shortening of the activation range can be seen in Figure 4.7 and Figure 4.9 as stimulation frequency increases. This is due to the quicker intervals between travelling waves and the shift in the peak basilar membrane displacement towards the base of the cochlea. The travelling wave, as seen in Chapter 3, is shorter for higher frequencies as the maximum peak of stimulation is closer to the base where the point of origin exists.

Figure 4.9 shows predictions from the alternative travelling wave speech processor with a decreased stimulation rate of 2880 pps with fewer electrodes activated for stimulation (electrodes 2, 6, 10, 14, and 18). This affects the continuous firing of nerve fibres along the length of the basilar membrane, causing gaps in the firing pattern.

The ISI histograms for the alternative travelling wave speech processor in Figure 4.8 and Figure 4.10 show grouping around the period of the stimulation signal and multiples thereof. However, for 1000 Hz pure tone, the 1 ms peak is small compared to the peaks of the multiples. The decrease in the number of electrodes and stimulation rate to 2880 pps causes a greater “noise floor” in the ISI histograms in Figure 4.8. Noise floor here refers to intervals not relating to the period of the signal and its multiples.

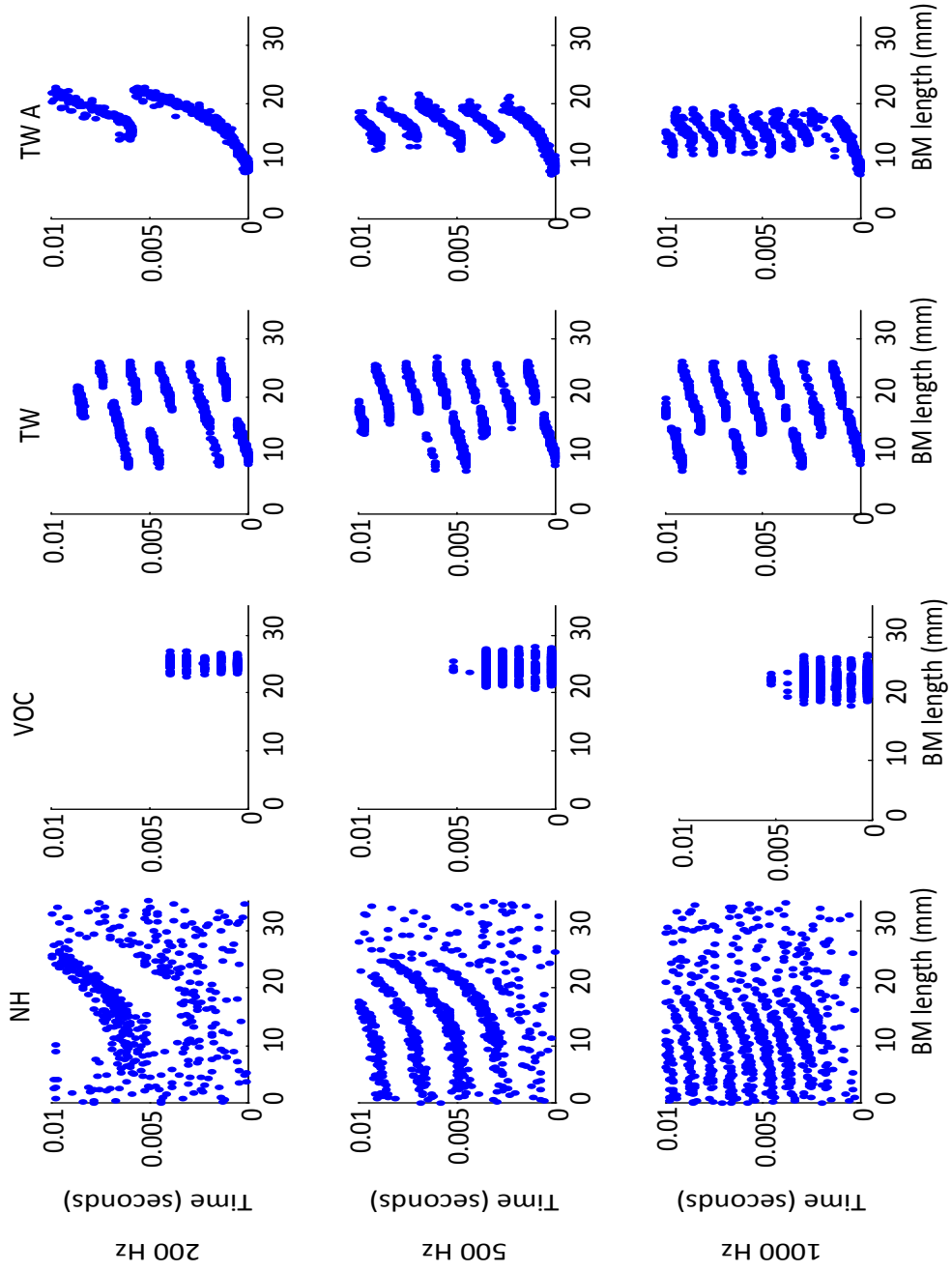

Figure 4.7. Nerve fibre spiking diagram for normal hearing compared to CI stimulation profiles, for pure tones of 200 Hz, 500 Hz, and 1000 Hz. The stimulation pulse rate for the CI ACE vocoder speech processing strategy was 1200 pps. The travelling wave speech processor has all 22 electrodes active in this example, and a sampling rate of 654 pps per electrode . The alternative travelling wave processor has a sampling rate of 14400 pps.

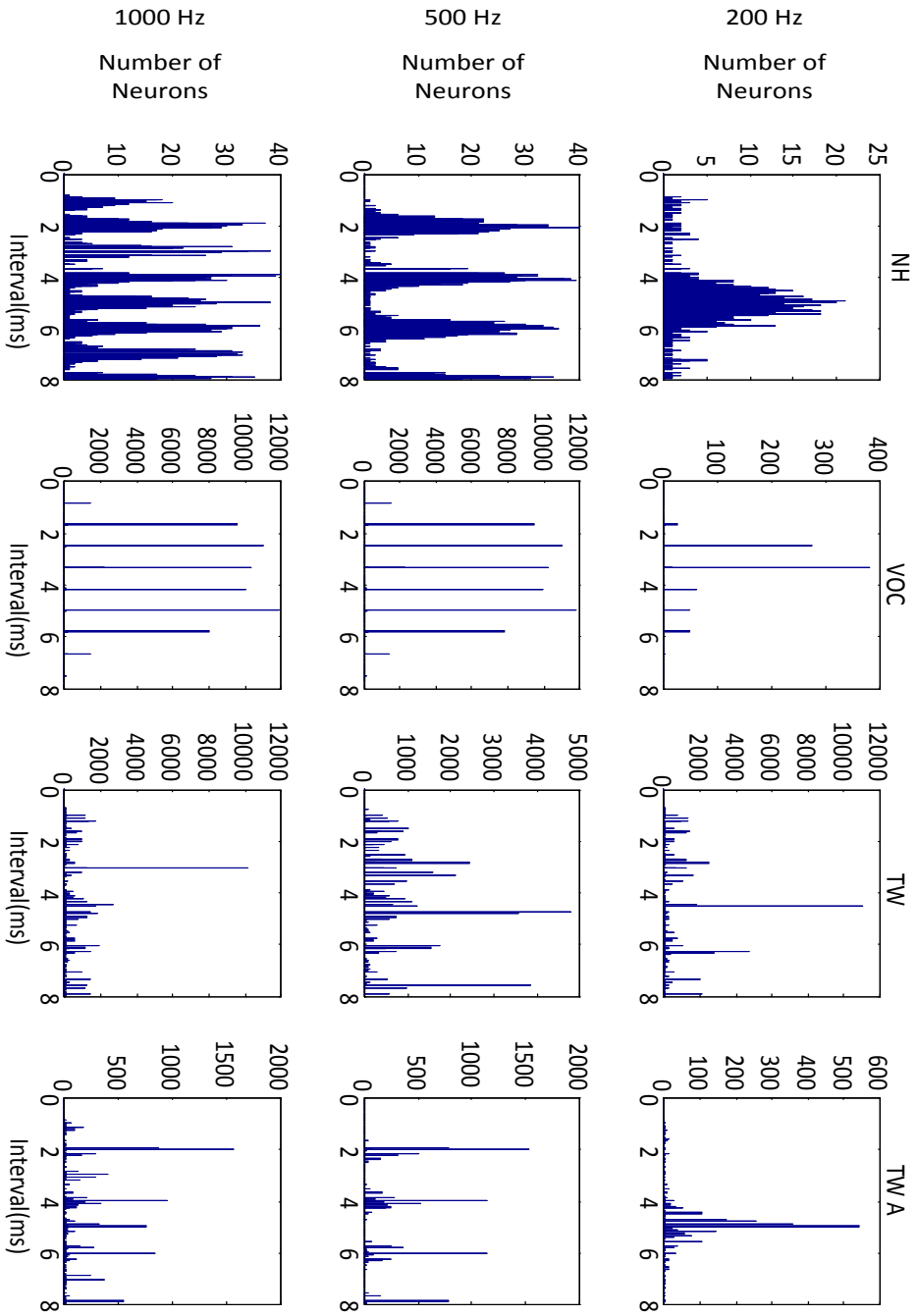

Figure 4.8. ISI histograms for normal hearing compared to CI stimulation profiles, for the pure tones of 200 Hz, 500 Hz, and 1000 Hz. The stimulation pulse rate for the CI ACE vocoder speech processing strategy was 1200 pps. The travelling wave speech processor had all 22 electrodes active, and a sampling rate of 654 pps. The alternative travelling wave processor has a sampling rate of 14400 pps.

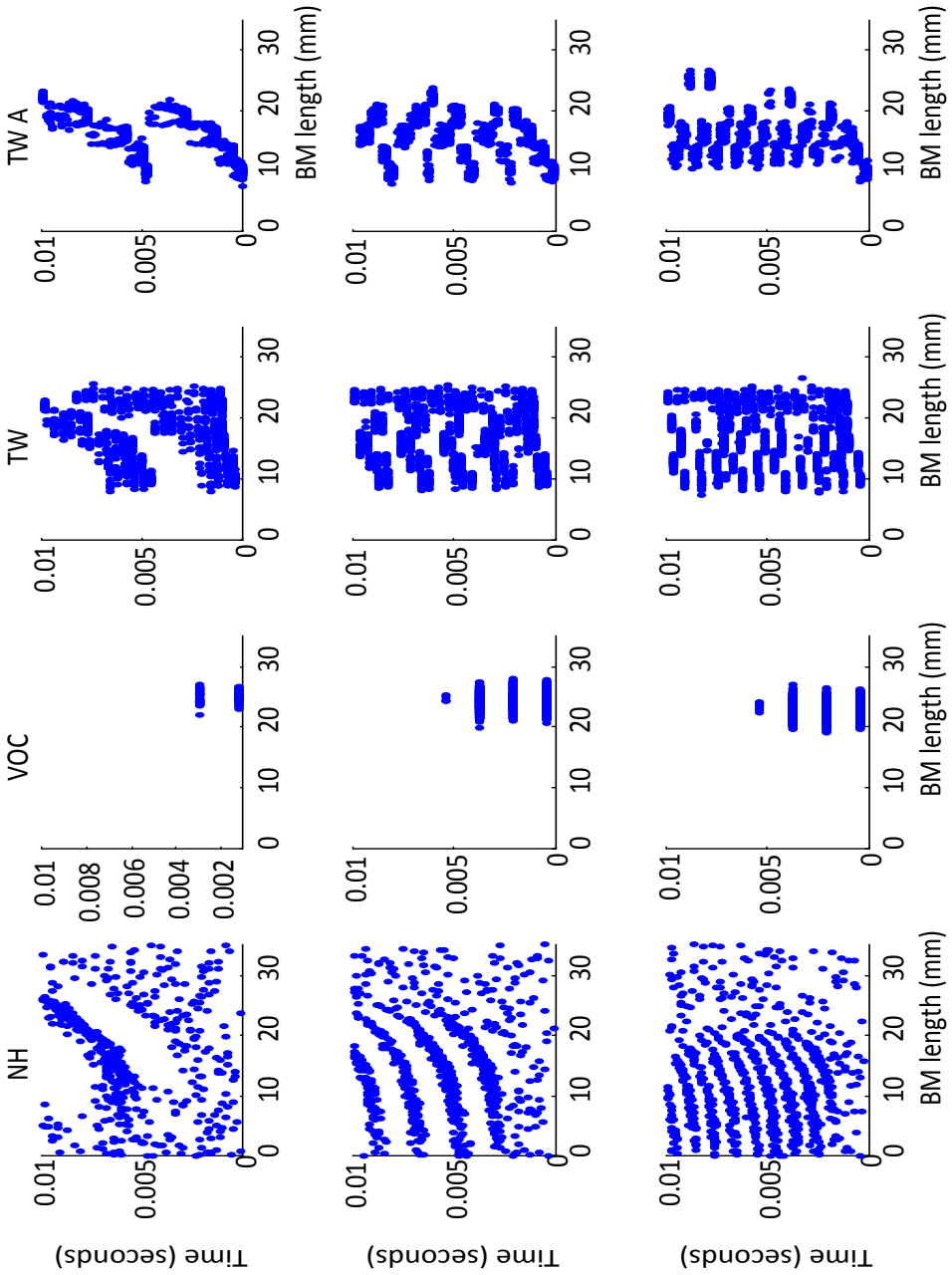

Figure 4.9. Nerve fibre spiking diagram for normal hearing compared to CI stimulation profiles, for the pure tones of 200 Hz, 500 Hz, and 1000 Hz. The stimulation pulse rate for the CI ACE vocoder speech processing strategy was 600 pps. Electrodes 2, 6, 10, 14, and 18 were active and a sampling rate of 2880 pps per electrode was used. The alternative travelling wave processor used a sampling rate of 2880 pps.

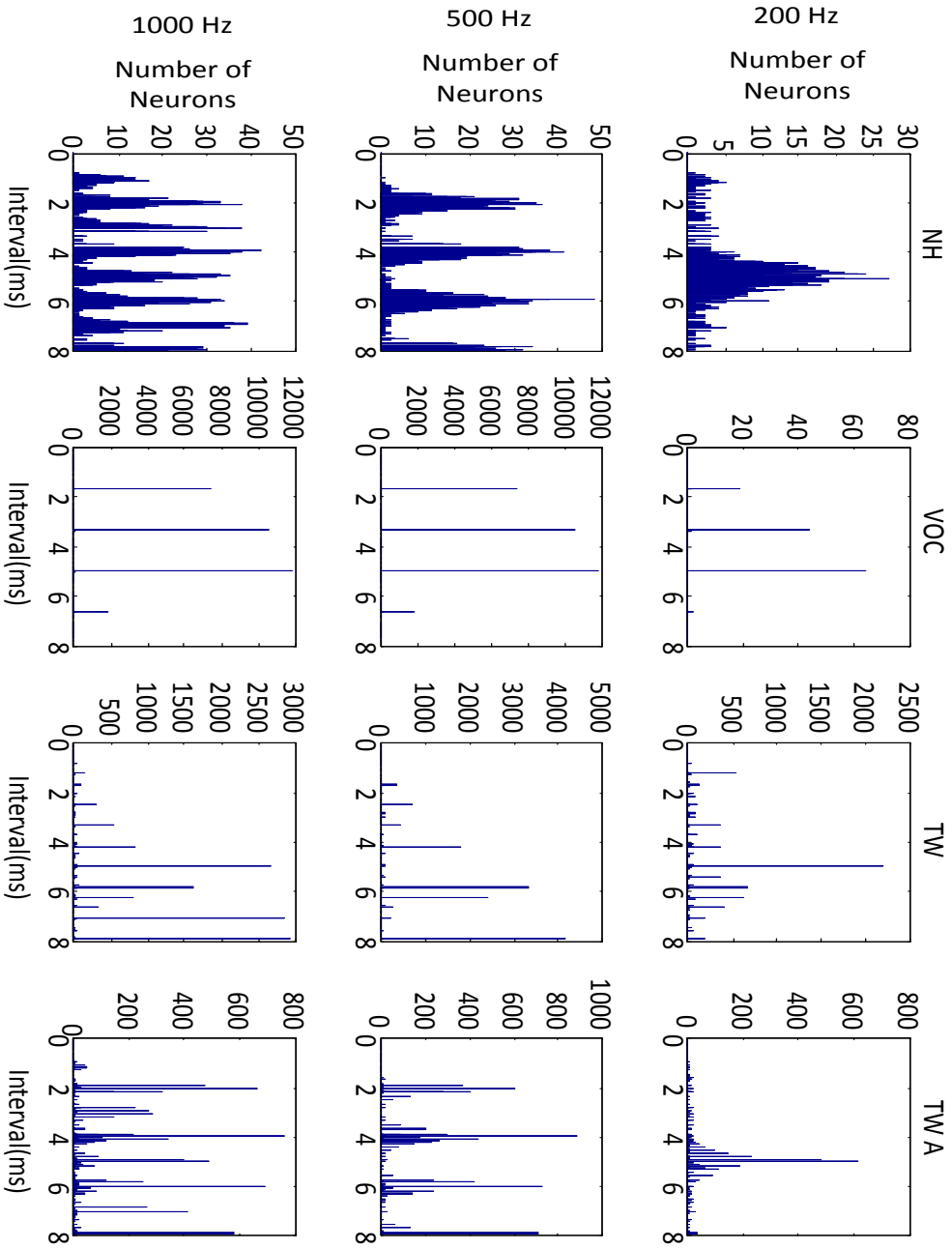

Figure 4.10. ISI histrogram for normal hearing compared to CI stimulation profiles, for pure tones of 200 Hz, 500 Hz, and 1000 Hz. The stimulation pulse rate for the CI ACE vocoder speech processing strategy was 600 pps. Electrodes 2, 6, 10, 14, and 18 were active and the sampling rate was 2880 pps per channel in the TW processor. The alternative travelling wave processor used a sampling rate of 2880 pps.

Figure 4.11 and Figure 4.12 show the results for the travelling wave speech processor and the space-time trade-off for a pure tone stimulus of 500 Hz at 70 dB SPL. The electrodes are evenly spaced across the length of the basilar membrane, and as the number of electrodes decrease the sampling rate is increased. With a high number of electrodes active at a low sampling rate (panel b), the spike firing diagram appears blurred and the pattern of the travelling wave cannot be seen clearly. This is similar for a low number of electrodes with a high firing rate, equally spaced along the electrode array (panel g). Five electrodes spaced closely together, and placed around the peak of the basilar membrane displacement during a pure tone stimulus showed the pattern of the travelling wave the best as seen in Figure 4.11 panel h. However, the spike pattern range is shortened markedly, as expected.

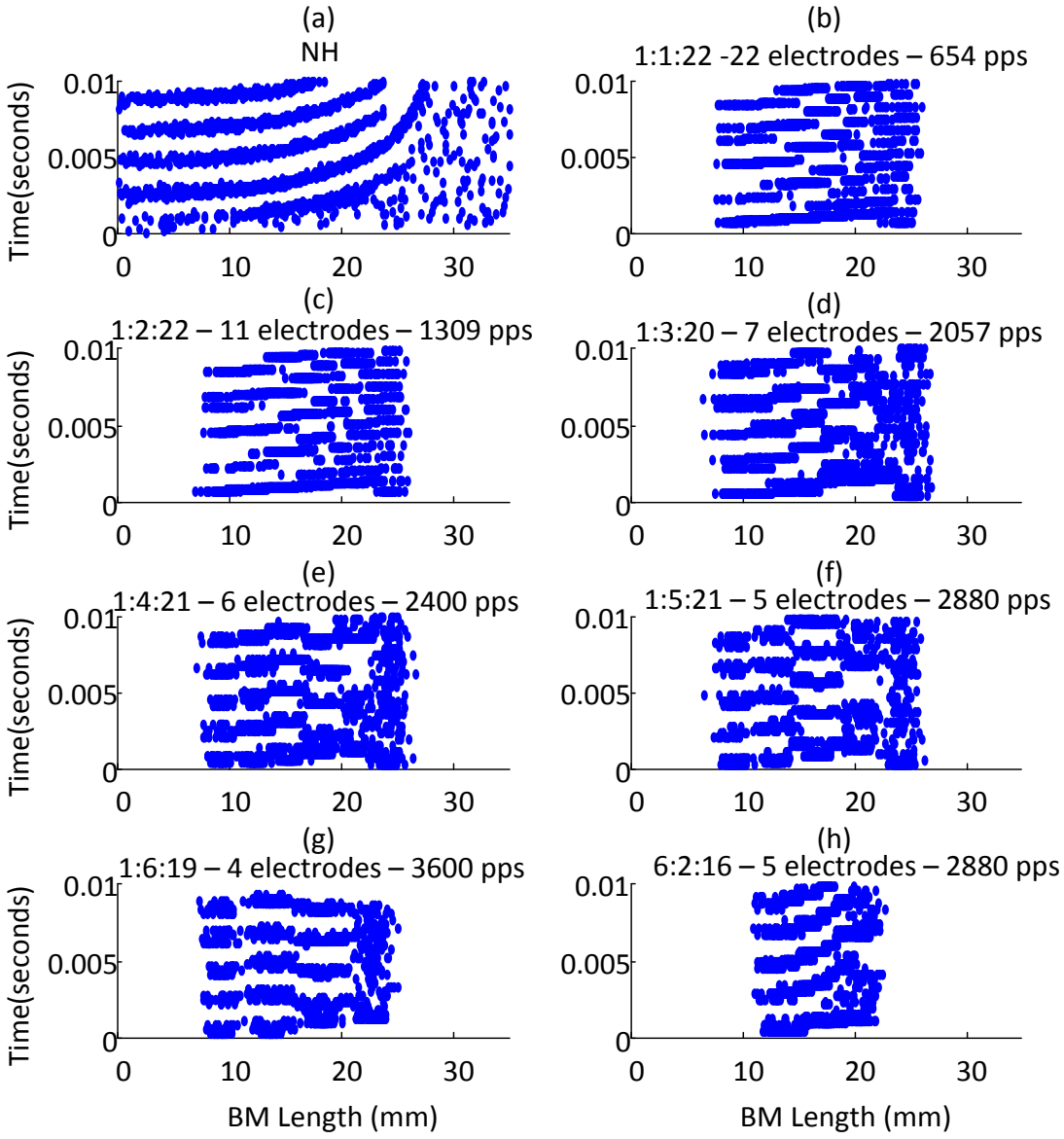

Figure 4.11. Spike train patterns for the travelling wave speech processor for a decrease in the number of electrodes used. Decreasing the number of electrodes permits an increase in the sampling rate (pulses per second). The notation first number:second number:third number above each panel means "start electrode, spacing between electrodes, last electrode," and the pulse rate per electrode is indicated after this.

The space-time trade-off effects are shown on the ISI histogram in Figure 4.12. Only when a lower number of electrodes is active do the intervals cluster around the period of the pure tone stimulus (panels f, g, and h). The ISI histograms clusters are limited in spread and do not have a distribution around the peak of the cluster. The only ISI histogram with the first

peak at the period of the stimulus signal is the simulation with 4 active electrodes and a pulse rate of 3600 pps (panel g).

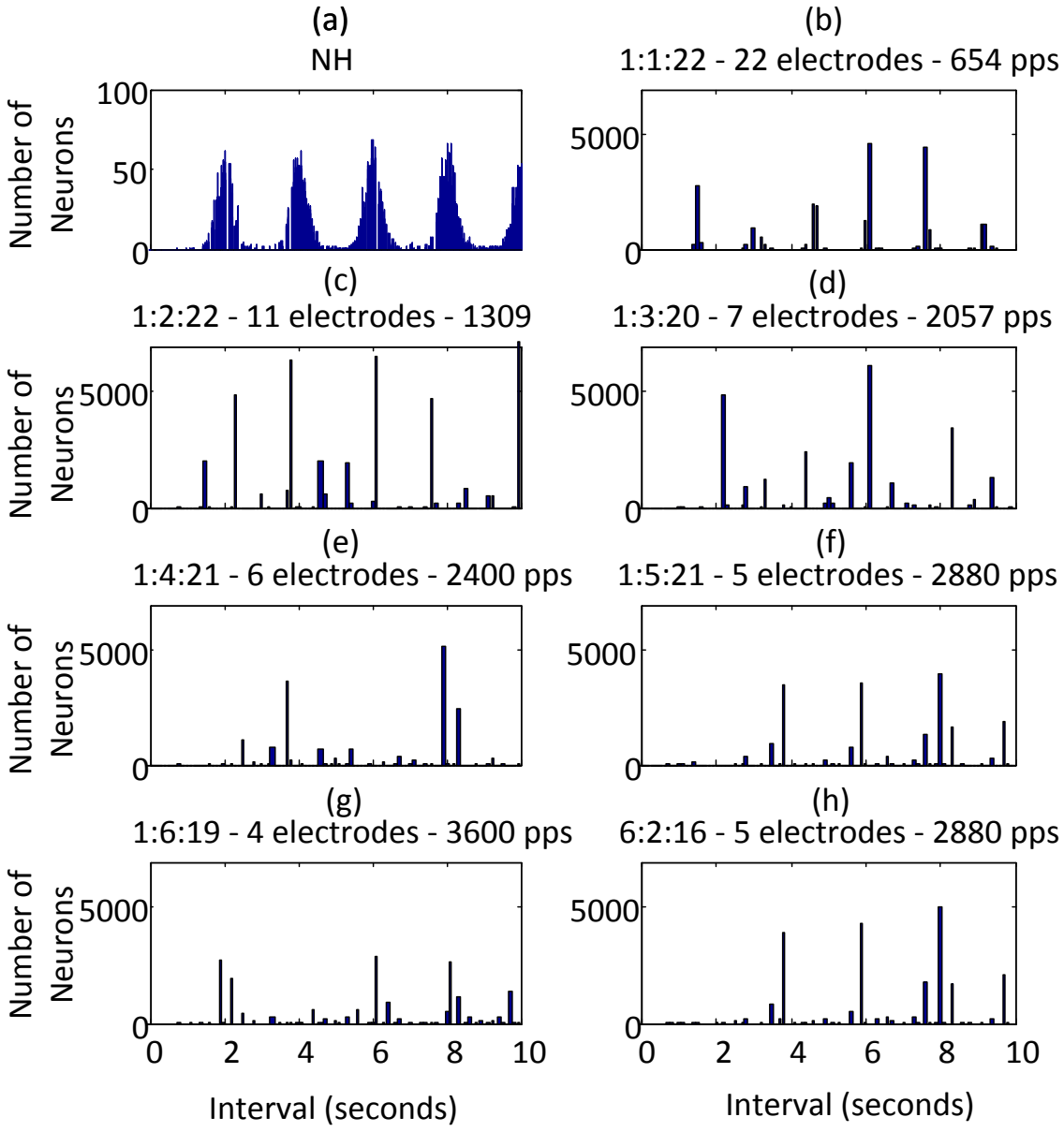

Figure 4.12. ISI histograms for the travelling wave speech processor for different numbers of electrodes used. The notation above each panel has the same meaning as on the previous figure. Decreasing the number of electrode results in an increase in the sampling rate (pulses per second).

As seen in Figure 4.7 to Figure 4.10, the alternative TW strategy is predicted to be better at preserving temporal information captured in the travelling wave. The paragraphs that follow consider the predicted spike train patterns for the alternative TW strategy in more detail. Figure 4.13 illustrates the nerve fibre spiking patterns for 1000 Hz pure tone

stimulation at 40, 60, and 80 dB SPL intensity levels. The electrical stimulation firing diagram does not show the random firings seen in the normal hearing model at low levels of stimulation. Rather, the travelling wave is tracked along the cochlea for the initial travelling wave generated by the model. The electrically-elicited travelling wave length is shortened because of the electrode range used and the place of stimulation is shifted towards the base (the peak basilar membrane displacement for a 20 dB 1000 Hz stimulus signal is at around 21 mm for normal hearing (this can be seen in chapter 3). Second and further travelling waves are not clearly reflected in the spike train patterns, and the extent of activation is shortened further for these. The travelling wave is tracked more clearly at higher stimulation intensities. Spike train patterns for the 80 dB SPL stimulus closely resembles those of NH over the section of the cochlea where the electrode array is situated.

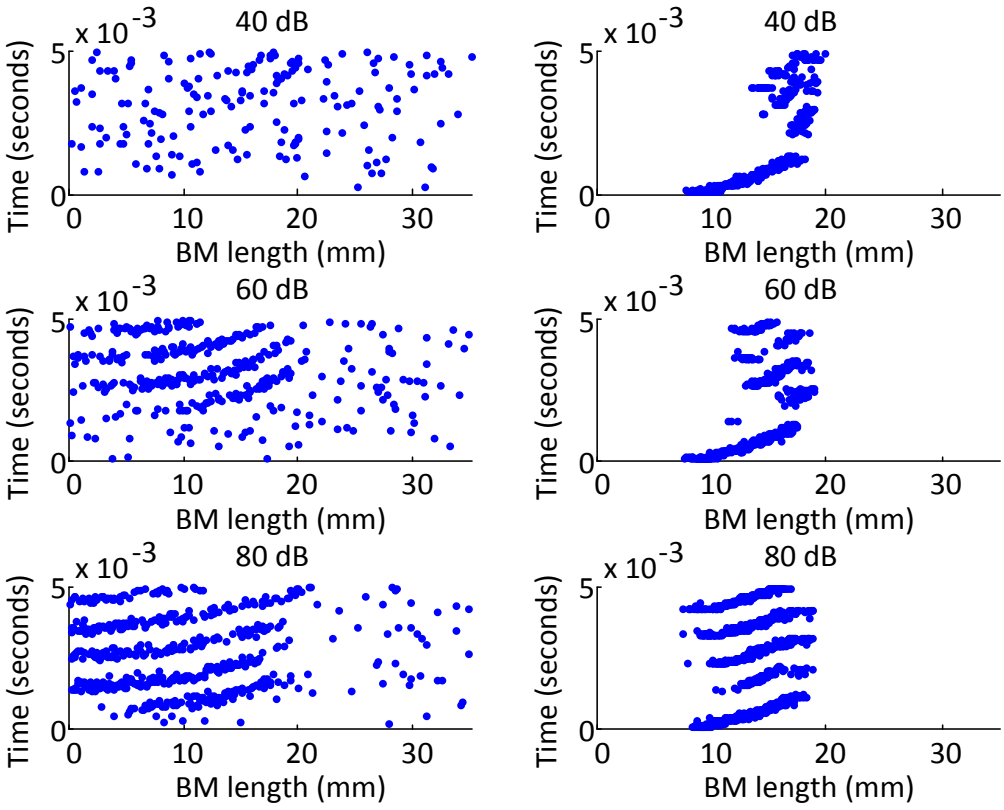

Figure 4.13. Normal hearing versus electrical stimulation spike firing diagram for 1000 Hz pure tone.

The Figure 4.14 compares the travelling wave neural activation patterns to normal hearing for 200 Hz, 500 Hz, and 1000 Hz at 60 dB SPL. Nerve fibres can be seen to lock onto the

travelling wave for all of these stimulation frequencies. However, the range is limited to the electrode extent in the cochlea, where it is shorter and random firings are far fewer.

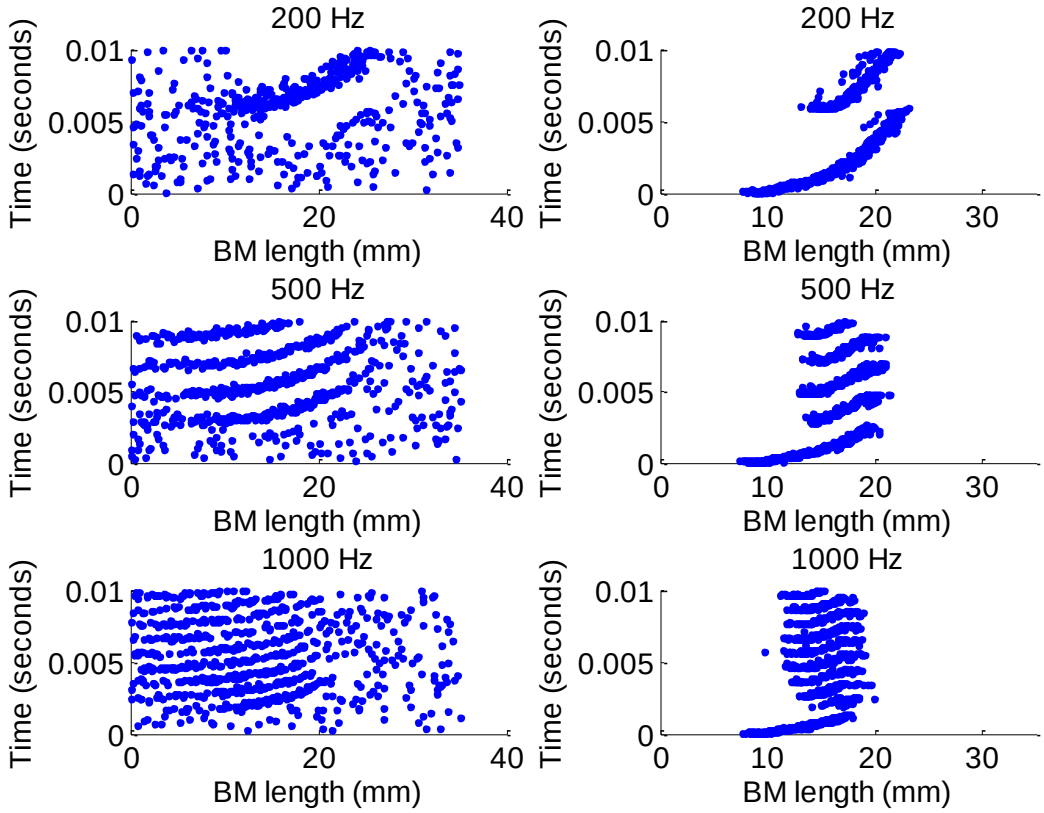

Figure 4.14. Normal hearing versus electrical stimulation spike firing diagram for 200, 500, and 1000 Hz stimulation pure tone at 50 dB SPL.

ISI histograms for different intensities of stimulation (Figure 4.15) shows phase locking to the pure tone stimulus period for all the simulated intensities for both normal hearing and electrical stimulation nerve fibre predictions.

The grouping of the intervals between spikes during travelling wave electrical stimulation is observed around the 1 ms time interval for the 1000 Hz pure tone stimulus used in these simulations. However, the histogram peaks are far sharper than those of normal hearing (i.e. phase locking is more precise). This spread decreases as the intensity increases from 40 dB to 80 dB SPL.

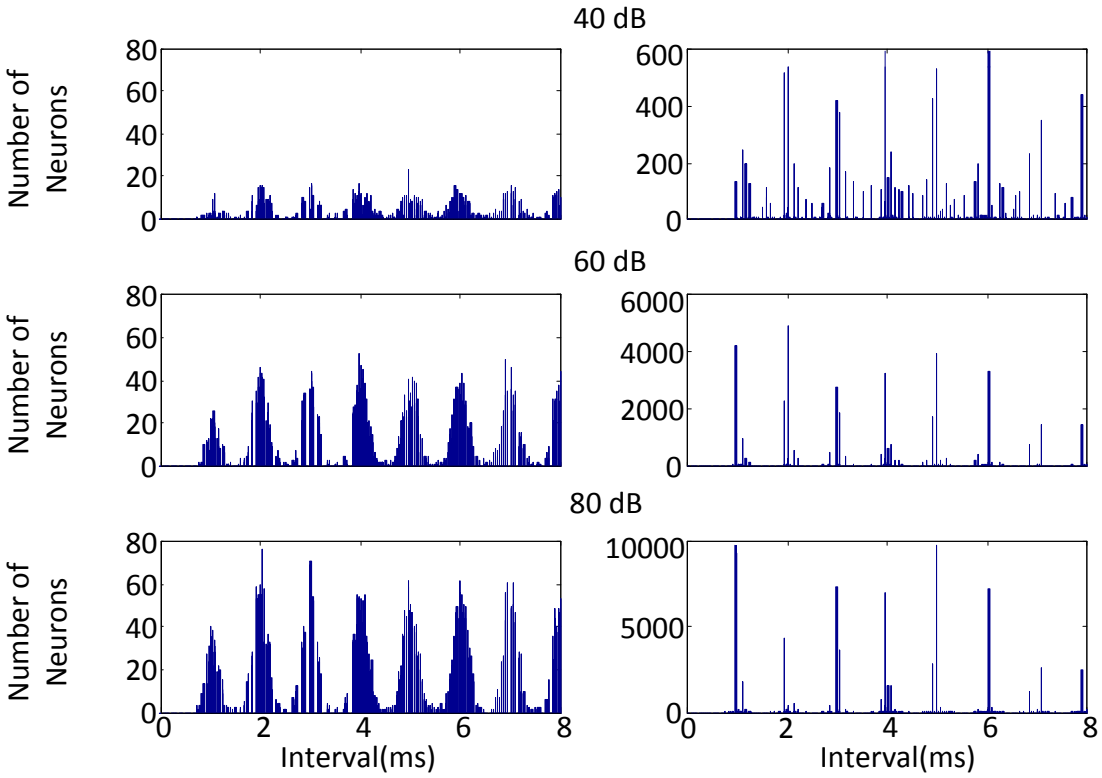

Figure 4.15. Normal hearing versus TW electrical stimulation for 1000 Hz pure tone stimulation at different stimulation intensity levels. The ISI histograms' y-axes have different ranges in each panel in the figure. This is to clearly show the shape of the ISI histogram.

The grouping of the nerve fibre spiking intervals around the stimulus period is shown for different pure tone frequencies in Figure 4.16. At low frequencies, the spread of the clusters around the stimulus period is similar to the spread seen in NH. Phase locking is more precise at higher frequencies and the spread of ISI histogram clusters is far smaller than those of the normal hearing. At 1000 Hz, the trend also appears to be for clusters to appear at twice the stimulus period, four times the stimulus period and so on. The cluster at 1 ms is small, possibly because of refractory effects that are still seen at 1 ms.

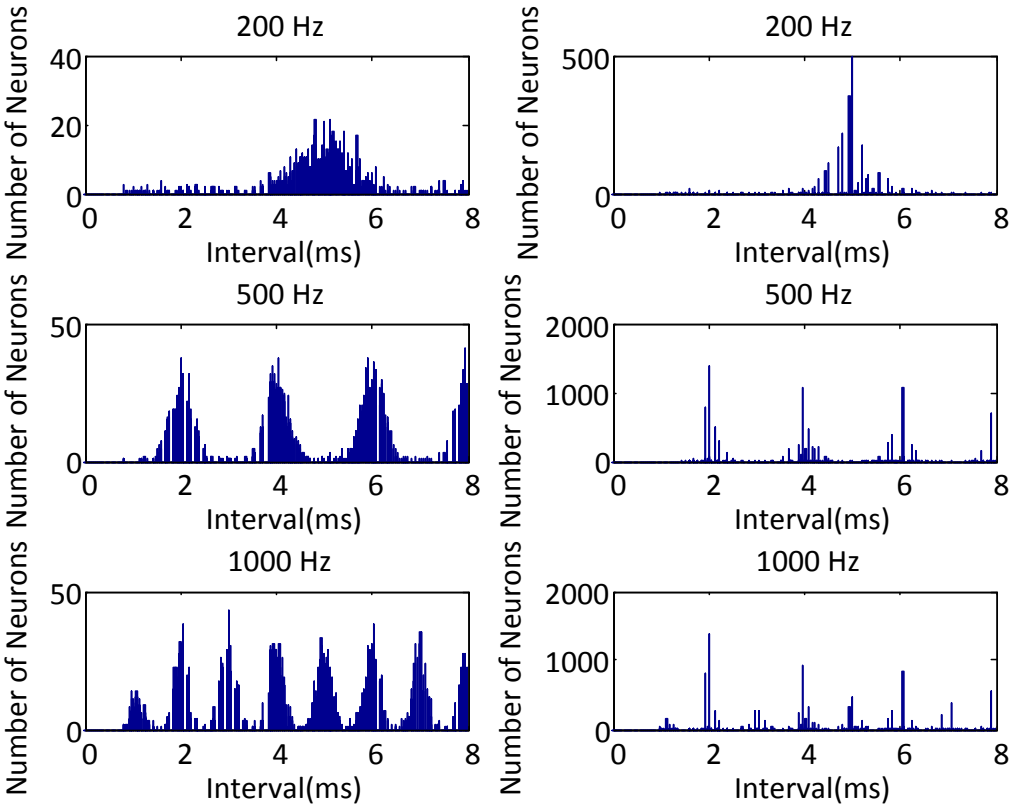

Figure 4.16. Normal hearing versus TW electrical stimulation for 200Hz, 500 Hz, and 1000 Hz at 60 dB SPL. Note that the ISI histograms ranges are different for different panels.

4.4.2 Synchronisation Index

The synchronisation index may be calculated from neural spike train patterns, as done in Chapter 3. The synchronisation index is a measure of the nerve fibre firing clustering around the pure tone signal stimulation peak, when observing a period histogram at one full period of the stimulation signal. Synchronisation is calculated by finding the period histogram from spike trains and then calculating the number of fibres locking onto the period of the pure tone as done in equation 3.35. The synchronization index was calculated for normal hearing and the alternative TW processor. The outcome is plotted in Figure 4.17 and Figure 4.18.

A comparison of synchronisation in normal hearing to the proposed TW processor, alternative TW processor, and VOC processor is shown in Figure 4.17, for a 200 Hz pure tone at 50 SPL. A clear difference between normal hearing and the CI TW stimulation is the shortening of the length of synchronisation along the basilar membrane, as expected.

In the area covered by the electrode, the alternative TW processing has a synchronisation index that is very similar to that of NH. The same is not true for vocoder stimulation or for TW stimulation (different trade-offs between number of electrodes and rate shown). This suggests that the alternative TW processing could be of value to encode the rate pitch information in low-frequency signals. The range of the synchronised nerve fibres is longer for the TW processer reflecting that the entire electrode array is used for the stimulation of the nerve fibres.

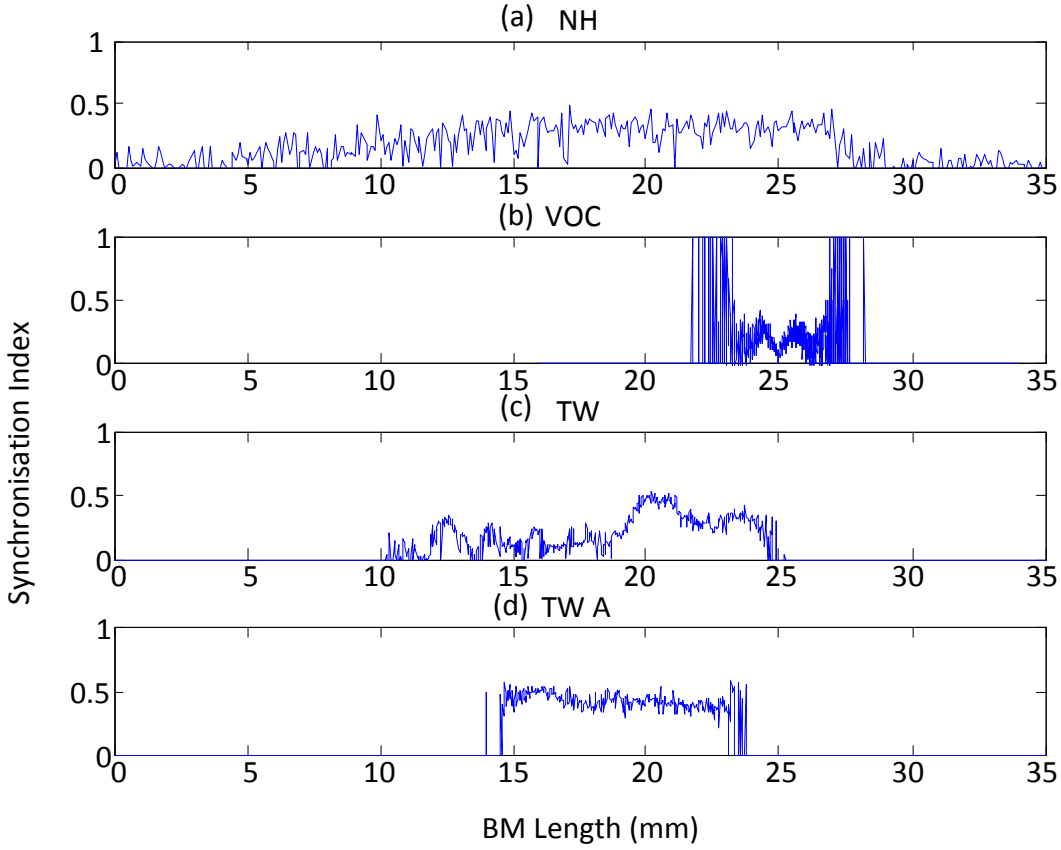

Figure 4.17. Synchronisation index comparison between normal hearing and the CI implant processing strategies for a 200 Hz pure tone at 50 dB. The synchronisation index was non-zero for the extent of the electrode array for TW stimulation, but zero elsewhere.

Figure 4.18 shows the synchronisation index for normal hearing nerve fibres compared to vocoder and CI TW electrical stimulation for a 1000 Hz pure tone. The NH synchronisation curve clearly shows a cut-off at around 21 mm, which is the frequency-place for 1000 Hz. Synchronisation for the alternative TW electrical stimulation remains high and very similar to that of acoustic stimulation. This suggests that this processing may be of value to extract and encode rate pitch in CI up to at least 1000 Hz.

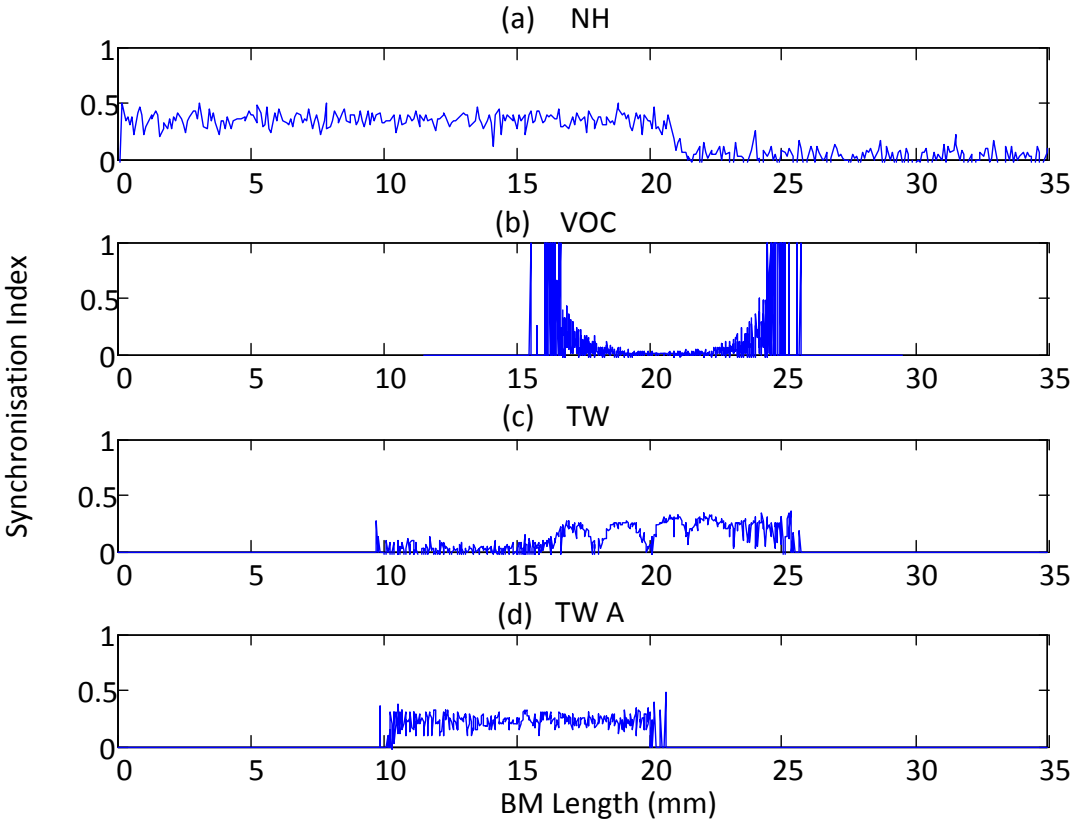

Figure 4.18. Synchronisation Index for nerve fibres along the length of the cochlea for a 1000 Hz pure tone at 60 dB.

Figure 4.17 and Figure 4.18 show apparent random spikes, with synchronization index reaching minimums of 0 and maximums 1 along the length of the basilar membrane. This is most likely due to a low number of nerve fibres firing on the nerve fibre measured which causes an inaccurate measurement. This phenomenon decreases as the stimulation time increases, allowing for more measured data; this is seen in Figure 4.19.

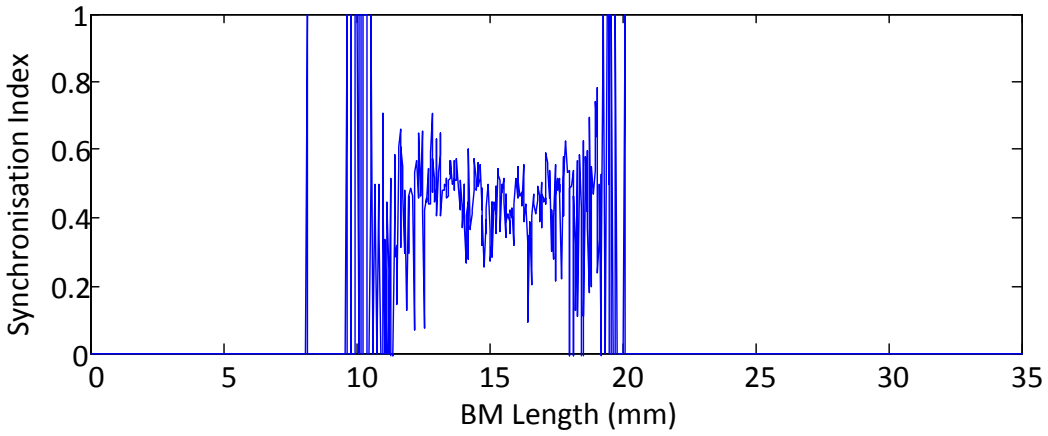

Figure 4.19. Synchronisation index for nerve along the length of the cochlea for 1000 Hz pure tone at 60 dB stimulation. The stimulation duration was 100 ms.

Figure 4.20 below shows the average synchronisation index of nerve fibres along the length of the basilar membrane for pure tone stimuli from 50 Hz to 2000 Hz in 50 Hz steps. The average synchronisation index is calculated by adding the synchronisation rate of each nerve fibre along the length of the basilar membrane and dividing the total by the total number of nerve fibres; therefore, collapsing the length of the cochlea synchronisation rate. The average synchronisation index for the alternative travelling wave processor is compared to those of the normal hearing model in Chapter 3 and the measure measurements from Javel and Mott (1988). The measurements of Javel and Mott (1988) are the maximum synchronisation of the nerve fibres while the models are compared using average synchronisation rate. However, the trend of the nerve fibres is comparable. The trend shows a slow decrease in synchronisation rate as frequency decreases, then a sharp drop off in synchronisation as the frequency of the pure tone stimuli approaches 1300 Hz. The increase in stimulus frequency shows a decrease in the average synchronisation of the nerve fibres. However, the trend of decreasing synchronisation with increase in frequency is comparable with those results of Chapter 3 predictions from the normal hearing auditory model.

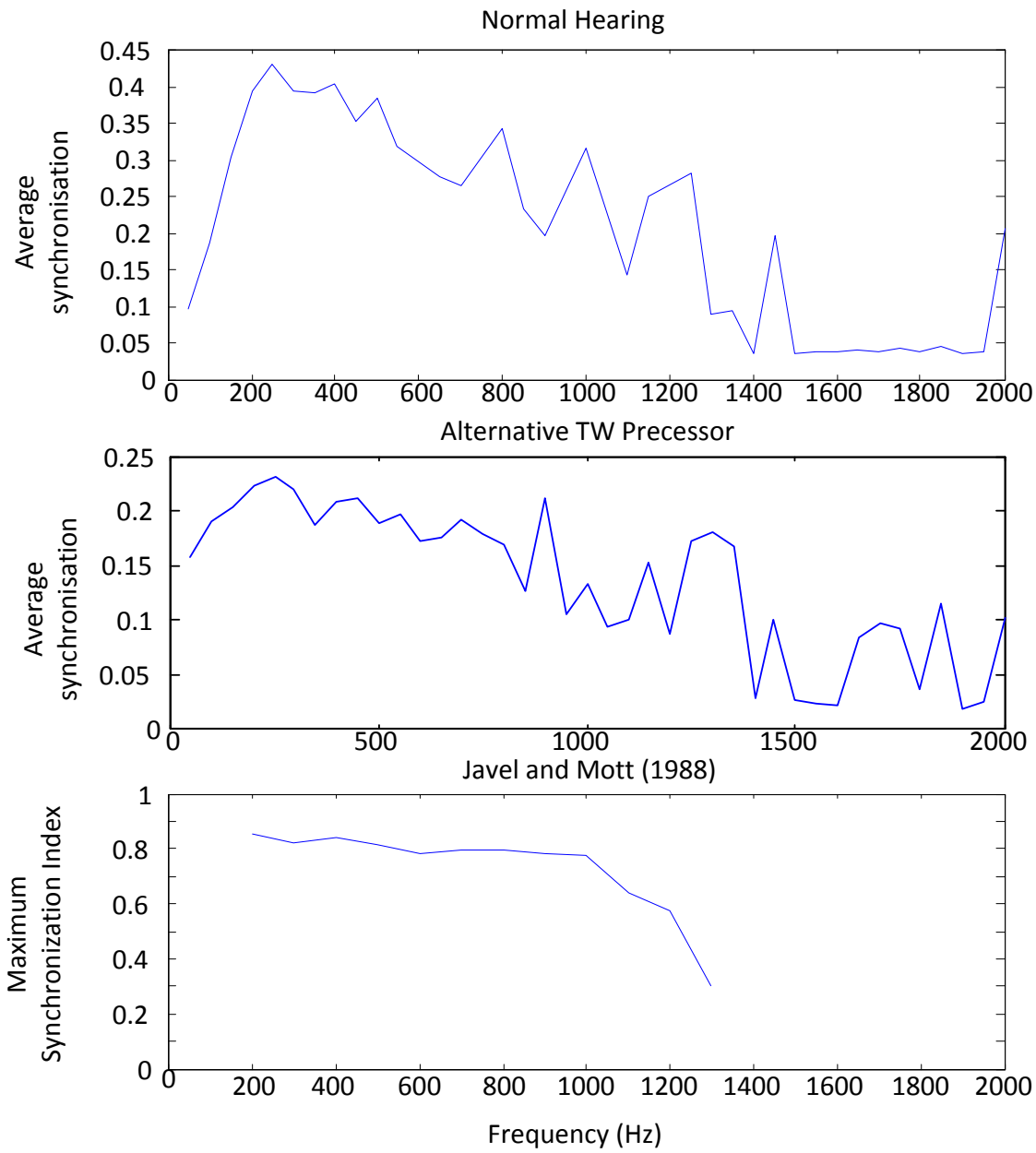

Figure 4.20. Average synchronisation of nerve fibres for pure tone stimuli from 50 Hz to 2000 Hz for the alternative TW speech processor and for normal hearing model from Chapter 3. The models for the alternative TW speech processor are compared to Javel and Mott (1988).

4.4.3 Missing fundamental stimulation results

The model predictions in the previous section suggest that the alternative TW processor may be able to induce some of the important characteristics of TW activated nerve fibre patterns. A crucial question is whether these electrically-elicited predicted spike train patterns, that resemble those of NH, imply that encoding of temporal pitch would be better with the proposed alternative TW processing than with a vocoder processor. This can only

be answered through perceptual testing. However, to be of value in encoding of temporal pitch information, a processor may be expected to induce spike train patterns from which pitch can be extracted in different conditions that simulate perceptual pitch experiments. Specifically, one of the most well-known perceptual characteristics of pitch is that the fundamental can be heard in harmonic tone complexes where the fundamental is removed (i.e. where there is no spectral component at the fundamental frequency). Figure 4.21 shows the nerve fibre spiking diagrams for the prediction of stimulation signals with a missing fundamental for the alternative TW speech processor; these results are compared to those from the normal hearing auditory model of Chapter 3. The missing stimulation signal was constructed using five overtones of the fundamental frequency. The nerve fibres can be seen spiking during both multiples of the stimulus signals. However, for the alternative TW speech processor the spike train patterns are not as wide in range as those of the normal hearing model (this is as expected), and the multiples of the stimulus period are harder to identify.

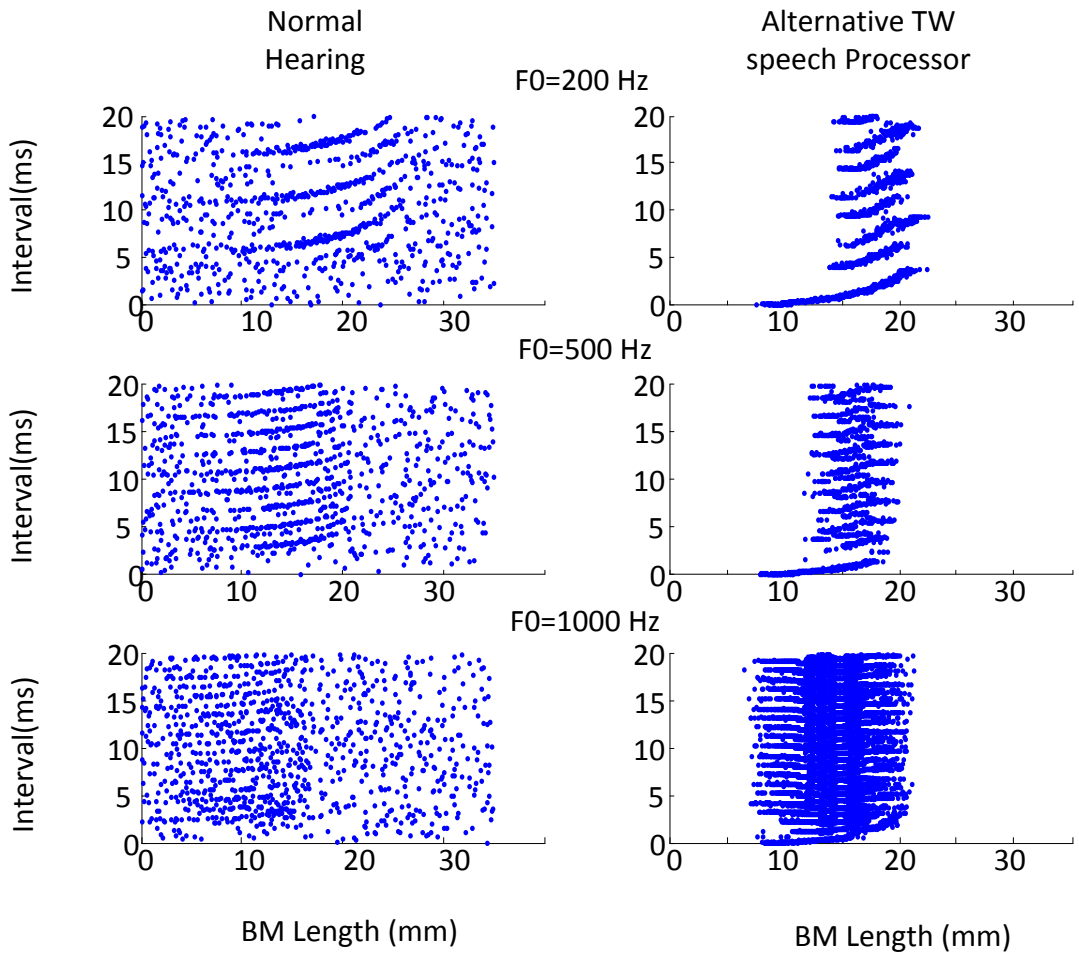

Figure 4.21. Nerve fibre spiking patterns for normal hearing and electrical stimulation by the alternative travelling wave speech processor for stimulation signals with missing fundamental.

The ISI histograms for the missing fundamental stimulus signal are shown in Figure 4.22, where the alternative travelling wave speech processor is compared to the normal hearing model of Chapter 3. A strong correlation is seen between the ISI histograms of the alternative travelling wave speech processor and that of the normal hearing model for 200 Hz. As the frequency of the missing fundamental is increased, the alternative travelling wave speech processors' ISI histograms become more random in spike interval times and cluster less around the missing fundamental frequency's period. A central processor should be able to extract the missing fundamental pitch from the spike trains at 200 Hz and 500 Hz. Whether or not the missing fundamental can still be decoded from the alternative travelling wave speech processor's spike train pattern at higher frequencies will be explored in the next chapter.

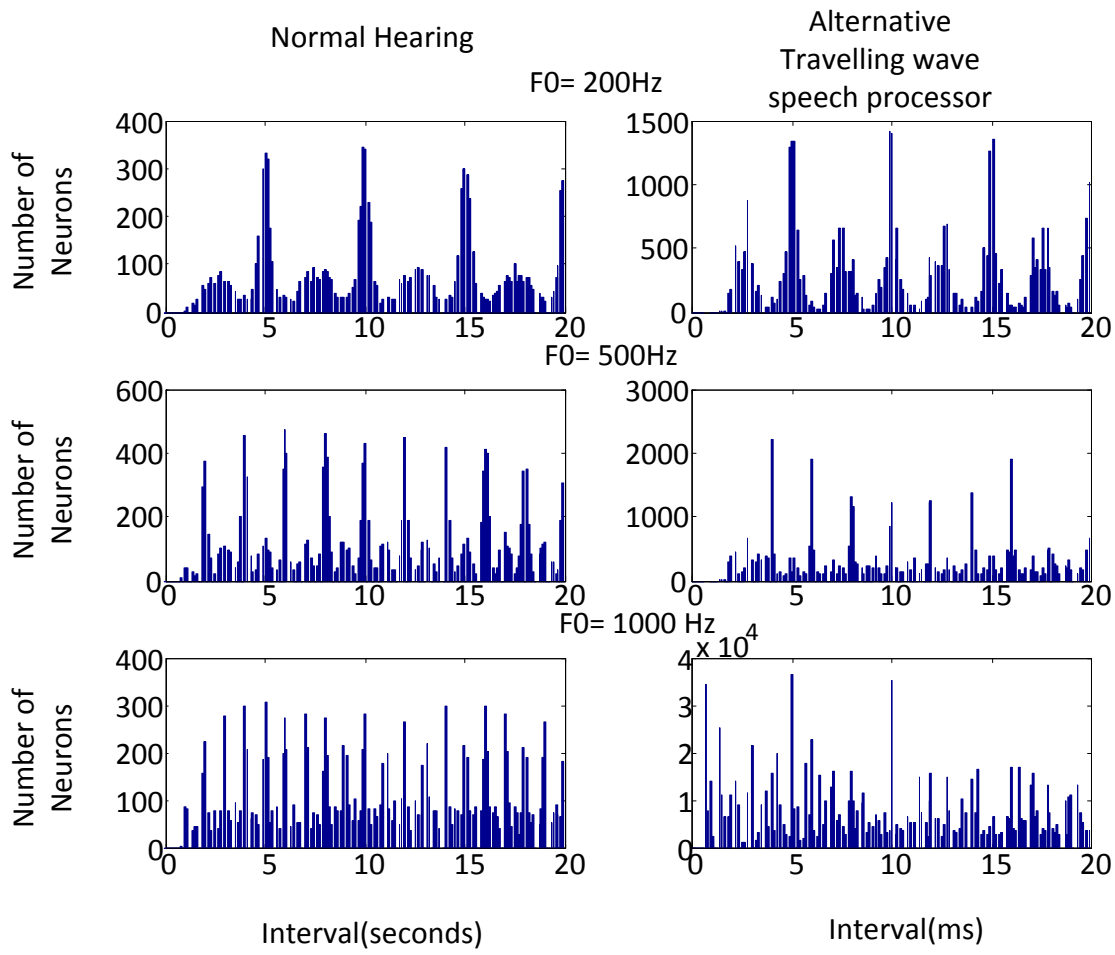

Figure 4.22. ISI histograms for the normal hearing and the alternative travelling wave speech processor for missing fundamental stimulation signals.

4.5 DISCUSSION

4.5.1 Implementation of a travelling wave processor

The model predictions in this chapter suggest that travelling wave speech processors may be of value. Travelling wave speech processors may be able to elicit neural excitation patterns for cochlear implant users that may deliver improved temporal pitch information. However, it is necessary to consider whether this is practical.

The generation of the travelling wave on a processing platform using a hydrodynamic model is time consuming, potentially making real time implementation unfeasible. Real-time implementation of either the TW speech processor or the alternative TW speech processor will probably require faster processing than that available in present cochlear implant speech processors. Processing time on the PC platform used from stimulus onset to

electrode stimulation is slow. An input stimulus of 10 ms will take 1.348 seconds to process up to the point where an electrode firing pattern is obtained. This is due to the high resolution travelling wave model, sampled at 400 kHz, that causes a bottle neck in the processing of the signal.

Using a simpler linear travelling wave model will decrease the processing time as the sampling rate of the travelling wave model can be greatly decreased. However, a linear, non-active travelling wave model will not contain amplification, which is important in determining the dynamic range properties of the auditory nerve fibres, and will not contain the distortion products caused by the intermodulation between two stimulus tones in the cochlea, which may be important to convey temporal properties of complex signals (Kemp, 2002). These two properties need to be observed and understood in a future study to determine the contribution to perception of sound.

The use of an analog travelling wave model to predict the displacement of the basilar membrane during stimulation (Wang et al., 2015, Yang et al., 2015) may be an alternative. This will allow for real time processing without computational delay.

Whichever route is taken, a travelling wave model, which could be implemented with little delay - from speech to electrode stimulation patterns, needs to be explored. More importantly, however, is the ability to show, in an animal model, that the predicted nerve fibre patterns are in fact elicited.

4.5.2 Travelling wave stimulation methods

Two alternatives have been considered here for presenting travelling wave information to the electrically stimulated cochlea. The method of Du Preez (2011) is subject to a temporal-spatial trade-off. The sampling of the travelling wave was done by allocating the number of electrodes that would be used for stimulation and by dividing the maximum total pulse rate between the number of electrodes. Each electrode will then be sampled sequentially one after another according to the amplitude of the travelling wave. As an example, if four electrodes where chosen for sampling, the sampling rate would be the maximum sampling rate (14400 pps in the cochlear device modelled) divided by 4 (3600 Hz) and the stimulation amplitude would be controlled by the displacement of the basilar membrane at the specific place (four cochlea places) and at the specific time.

This method of place and time sampling naturally has a trade-off between spatial and temporal resolution, resulting in instances in time in which the active electrode has stimulation amplitude of zero. This time slot could possibly be used to stimulate a different electrode (where the basilar membrane displacement is greater than zero).

The trade-off between temporal and spatial resolution presents a problem that will need to be overcome in practical TW speech processors.

The problem may be mitigated in newer speech processors, or speech processors that allow higher rates of stimulation. The PULSARci^{100} from Med-El and the Clarion CII from Advanced Bionics Corporation, with overall maximum stimulation rates of 50 704 pps and 82 000 pps (Laneau and Wouters, 2004), as opposed to the 14400 pps rates in the processor modelled for this work, may allow for a greater spatial resolution whilst still obtaining a higher stimulation rate. The high stimulation rate across multiple electrodes that increases the spatial resolution, is necessary for obtaining the ISI histograms, which are similar to those of temporal characteristics seen in measured results (Rhode, 1971). The problem is not the stimulation rate as such, but the decrease in spatial resolution in order to have the stimulation rate over a required level to sample the travelling wave in the temporal domain.

The proposed alternative TW stimulation method, that tracks the point of maximum amplitude in the travelling wave, mitigates the problem of needing to know the optimal trade-off between the number of electrodes and rate needed to sample the travelling wave in time and space. The alternative travelling wave method uses all of the available electrodes, and stimulates these at maximum available rate of the cochlear implant processor (of 14400 pps in the presently modelled device). In the simulations shown, the alternative travelling wave processor appears to have the potential for the encoding of pitch information captured by the travelling wave. It is not known whether this will work in practice and whether this will work for any other signals (apart from the pure tones and missing fundamental stimulation signals that were tested), as no animal models or psychoacoustic tests have been carried out so far.

Implementation of travelling wave characteristics have been attempted in cochlear implants before, such as the processors suggested by Taft et al. (Taft et al., 2009, Taft et al., 2010). Taft et al. (2009) used the ACE processing strategy to implement across-filterbank frequency delays which related to the cochlea travelling wave delay. This showed an improvement in pitch ranking tests across cochlear implant users. However, the

improvement probably resulted from better processing in the sense that stimulation pulses had less temporal cross-channel interaction because the delay was greater between stimulation pulses.

Harczos et al. (2013) implemented a hydrodynamic travelling wave model with wave digital filters to implement the SAM (Stimulation based on Auditory Modelling) processing strategy for cochlear implant users. The simulated basilar membrane displacement was used to decide which electrodes should be stimulated. During each cycle, three electrodes could be stimulated sequentially, allowing for three place samples to represent the travelling wave shape. Harczos et al. (2013) did not use nerve fibre predictions, but rather showed that synchronisation should be increased using electrode firing timing. They also showed adaptation in electrode activation, but did not model nerve fibre adaptation response resulting from an increase in stimulus amplitude. The just noticeable difference of 3.9 semitones dropped to 2.2 semitones for pure tones stimuli and from 7.9 semitones to 5.1 semitones for sung vowels stimuli when switching from the ACE processing strategy to the SAM processing strategy. A shortcoming of this work was that outcomes could only be assessed experimentally (in perceptual experiments), which does not allow for insight into the underlying neural activation patterns and interpretation of perceptual outcomes through an understanding of what happens at the neural level. The present work significantly extends the work of Harczos et al. (2013) and Taft et al. (2009) by assessing processing strategies through detailed nerve fibre models and predictions of nerve fibre activation patterns.

ISI histograms for the proposed travelling wave processor suggest that temporal information captured by the travelling wave may be relayed effectively to the auditory nerve fibres. However, TW speech processor may not be effective for speech processing and may suppress information other than low frequency tones or voice pitch. The broadening of the travelling wave due to an increase in loudness, the specific sites of stimulation, and the number of nerve fibres synchronised to the stimulus may all contain critical information and are all affected by the specific travelling wave stimulation method. It is not clear yet how TW processing captures specific important aspects of the signal, for example signal intensity or spectral information (e.g. formants).

The place of the maximum stimulus is shifted towards the base of the cochlea due to the depth at which the electrode array can be implanted. The electrode array insertion depth

varies greatly across implants (Van Der Marel, Briaire, Wolterbeek, Snel-Bongers, Verbist and Frijns, 2014). This means that the travelling wave processor will also have the same frequency-place mismatch seen with vocoder speech processors. The mismatch could affect the delivery of temporal information to the brain. Possibly as nerve fibres synchronise to a pure tone's period across a large extent of the basilar membrane (see, for example, Figure 4.11 a) during normal stimulation, temporal information available from a TW processor may still be available to a CI user even if these mismatches occur, as the central auditory system may have the ability to extract temporal information from an extent of nerve fibres far beyond the CF position of a pure tone (Oxenham, Bernstein and Penagos, 2004).

4.5.3 Delivery of temporal information

This study indicates that the proposed travelling wave speech processor for cochlear implants could potentially improve the delivery of temporal information to the auditory nerve fibres through cochlear implants, as temporal information can be observed in the ISI histogram. The synchronised firing of the nerve fibres to the pure tone stimulus (Rhode et al., 2010) is an important neural fibre characteristic for encoding pitch (Larsen et al., 2008). The ISI histograms can be used to predict the perceived pitch a listener would hear as it contains the fundamental frequency of the acoustic signal. Larsen et al. (2008) showed that these predictions could be done accurately from pooled ISI histograms.

If the temporal pitch information present in the predicted ISI histograms of the simulated alternative TW processor appears in real neural activation patterns, this could lead to benefits in the delivery of pitch information beyond the observed 300 Hz limit (McDermott and McKay, 1997, Zeng, 2002), to temporal pitch perception that is closer to normal hearing (Oxenham, Micheyl, Keebler, Loper and Santurette, 2011, Moore and Ernst, 2012). Oxenham et al. (2011) found that pitch was perceived for auditory signals with a fundamental pitch, which did not have a place pitch encoding relating to the fundamental frequency, above the limits of neural phase locking up to signals with a fundamental frequency of 2 kHz (above the limits of neural phase locking), while Moore and Ernst (2012) showed that phase locking of nerve fibres play a role in pitch discrimination up to around 8 kHz.

Predictions shown here indicate that the vocoder speech processor fails to encode accurate temporal information of the acoustic signal into the stimulation pulse sequence. ISI

histograms show that nerve fibres do not phase lock to the period of the stimulus signal. Hartmann and Klinke (1990) identified nerve fibres not phase locking onto the period of the pitch in CI as a potentially limiting factor for CI pitch perception, while the availability of place pitch is also limited as spatial resolution is limited in a CI. A greater focus on the accurate delivery of the temporal information obtainable from the acoustic signal may be of value.

4.6 CONCLUSION

- The ISI histograms from the proposed new travelling wave speech processor model (alternative TW processor) and the travelling wave model from Du Preez (2011) showed spiking neurons locking onto the period of the acoustic signal and not onto the stimulation electrical signal, as is seen in vocoder speech processors.
- The travelling wave model from Du Preez (2011) requires insight into the space-time trade-off to ensure that parameters are selected that allow neural spiking to lock onto the period of the acoustic signal. The optimal choice of number of electrodes and stimulation rate is not known.
- The mimicking of the travelling wave in speech processors for cochlear implants may allow for the delivery of rate pitch cues as is seen in the ISI histograms of especially the alternative TW speech processor

CHAPTER 5 PITCH INFORMATION AVAILABLE FROM SPEECH PROCESSORS

5.1 INTRODUCTION

The auditory nerve fibres encode spatio-temporal information in the firing of nerve fibres. This chapter discusses and implements models to decode the space-time neural spike trains predicted by the models implemented in the previous chapters. The decoding of the spatial and temporal aspects of the nerve fibres is carried out to validate that the proposed travelling wave speech processor provides an improvement on the CI ACE strategy for pure tone stimuli.

The temporal decoding of the nerve fibres is executed by decoding the nerve fibre spikes using ISI histograms and subsequent processing to extract the pitch of the signal. The method will be described in this chapter.

5.2 METHODS

5.2.1 Template matching using autocorrelationgram

This section summarises the processing steps needed to arrive at pitch predictions. The sections that follow will then discuss each step in more detail. The block diagram, Figure 5.1, illustrates the decoding of the nerve fibres to extract the temporal information. A template matching approach is used here to estimate the pitch of the signal, as this approach extracts the phase locking of the nerve fibres to estimate the period of the stimulus signal (Cariani and Delgutte, 1996a, Cedolin and Delgutte, 2005).

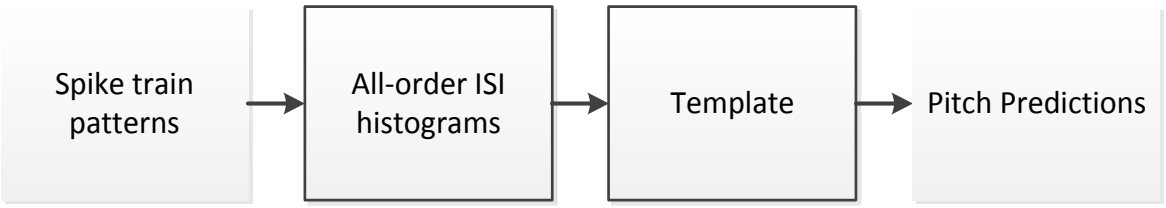

Figure 5.1. Block diagram illustrating the implementation of the decoding mechanism used for pitch detection by autocorrelation.

The nerve fibre spike train patterns from the normal hearing model and the electrical stimulation models are used to extract the temporal information encoded therein. An

autocorrelationgram is used to predict the encoded pitch of the signal from the neural excitation pattern. An autocorrelationgram describes the neural firing intervals with respect to time. Pitch prediction is done using an all-order ISI histogram instead of a first-order ISI histogram (Cariani and Delgutte, 1996a). The benefits of using an all-order ISI histogram is that it is not affected by the loudness of the stimulation signal to detect pitch (loudness level does not affect the prediction of the pitch), it predicts the upper F0 limit of pitch perception using temporal encoding, and it is phase invariant.

The pooled autocorrelationgram (Figure 5.5), which is an all-order ISI histogram of each nerve fibre firing across the length of the basilar membrane, added together to form a single all-order ISI histogram, is matched to templates of a specific pitch. The templates are constructed by using a fundamental frequency in question, F0, and multiples of the fundamental frequency which are a values of 1 in an array of zeros. The template is then correlated with the weighted ISI histogram as seen in Figure 5.2. Template matching results in a contrast ratio, from estimated pitch is then directly extracted. These steps are expanded on below.

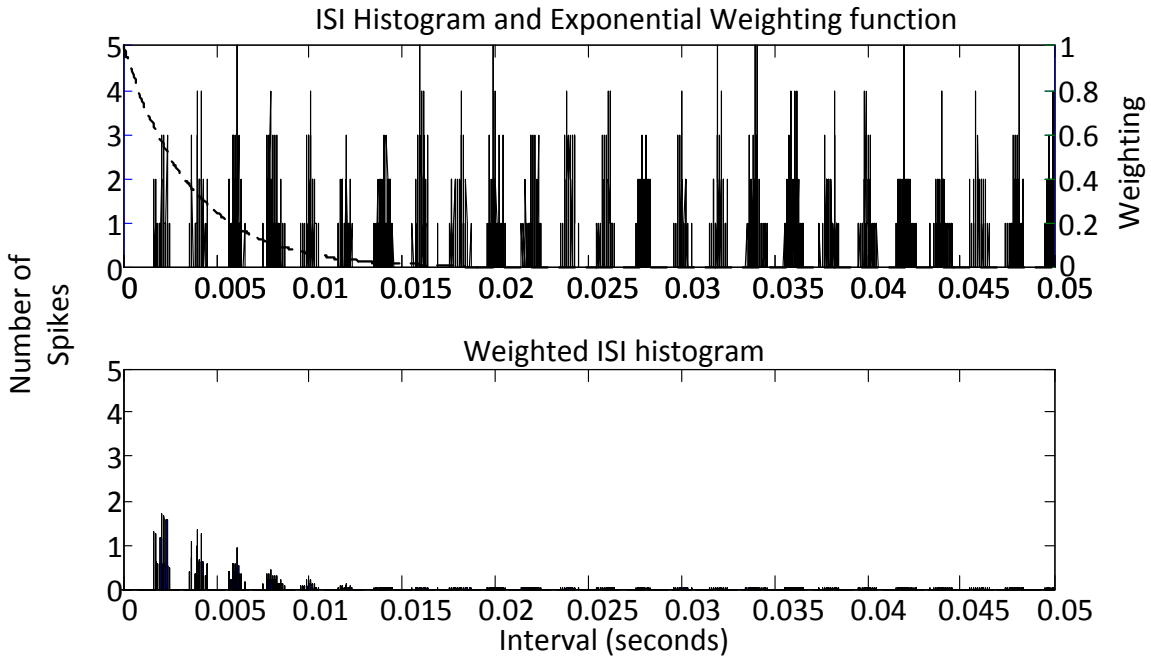

Figure 5.2. ISI histogram weighted by an exponential weight function and the resulting weighted ISI histogram.

5.2.2 Autocorrelationgram

The autocorrelationgram is used to see the running interval change in ISI information and is used in this thesis as a different illustration of the All-order ISI histograms. The autocorrelationgram is a two-dimensional histogram whose bins contain the number of ISIs of a given length ending at a particular time relative to the stimulus onset. Thus horizontal bins represent a PST (post-stimulus time) histogram of the signal, whereas vertical bins represent an ISI histogram (Cariani and Delgutte, 1996a).

Figure 5.3 below shows the autocorrelationgram for a pure tone stimulation of 1000 Hz. The autocorellationgram illustrates the nerve fibres locking onto the period of the stimulation and multiples during the entire length of the stimulus of 100 ms.

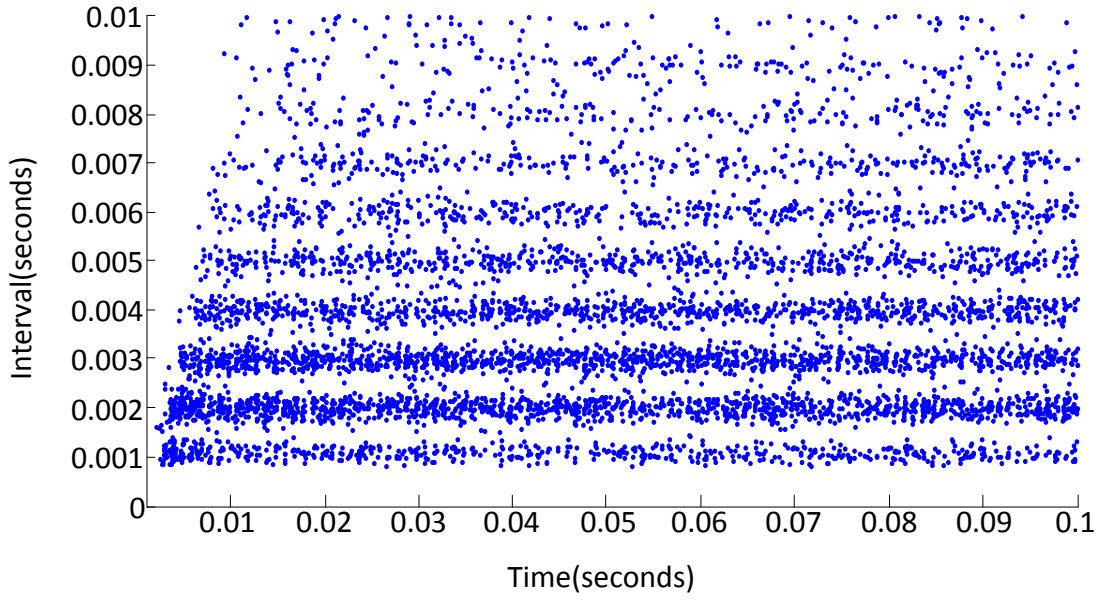

Figure 5.3. Autocorrelationgram for a pure tone of 1000 Hz. A single dot represents a single nerve fibre spike.

The autocorrelationgram can be used to illustrate the change in nerve fibre intervals relative to the stimulus onto the pitch of sound. Figure 5.4 represents the normal hearing model to a missing fundamental signal with a fundamental frequency of 300 Hz. If the stimulus pitch is changing during the duration of the stimulus, the interval will change. The autocorrelationgram can be used to see the intervals of firing of nerve fibres during different times of stimulation of the signal. This will be important during multi-tone stimulus in which the pitch perceived would change during the duration of the stimulus.

The pooled autocorrelationgram (all-order ISI histogram), which is constructed from the autocorrelationgram by summing the interval (x-axis) across the entire time period of stimulus (x-axis), shows the intervals that occurred during the entire stimulation period.

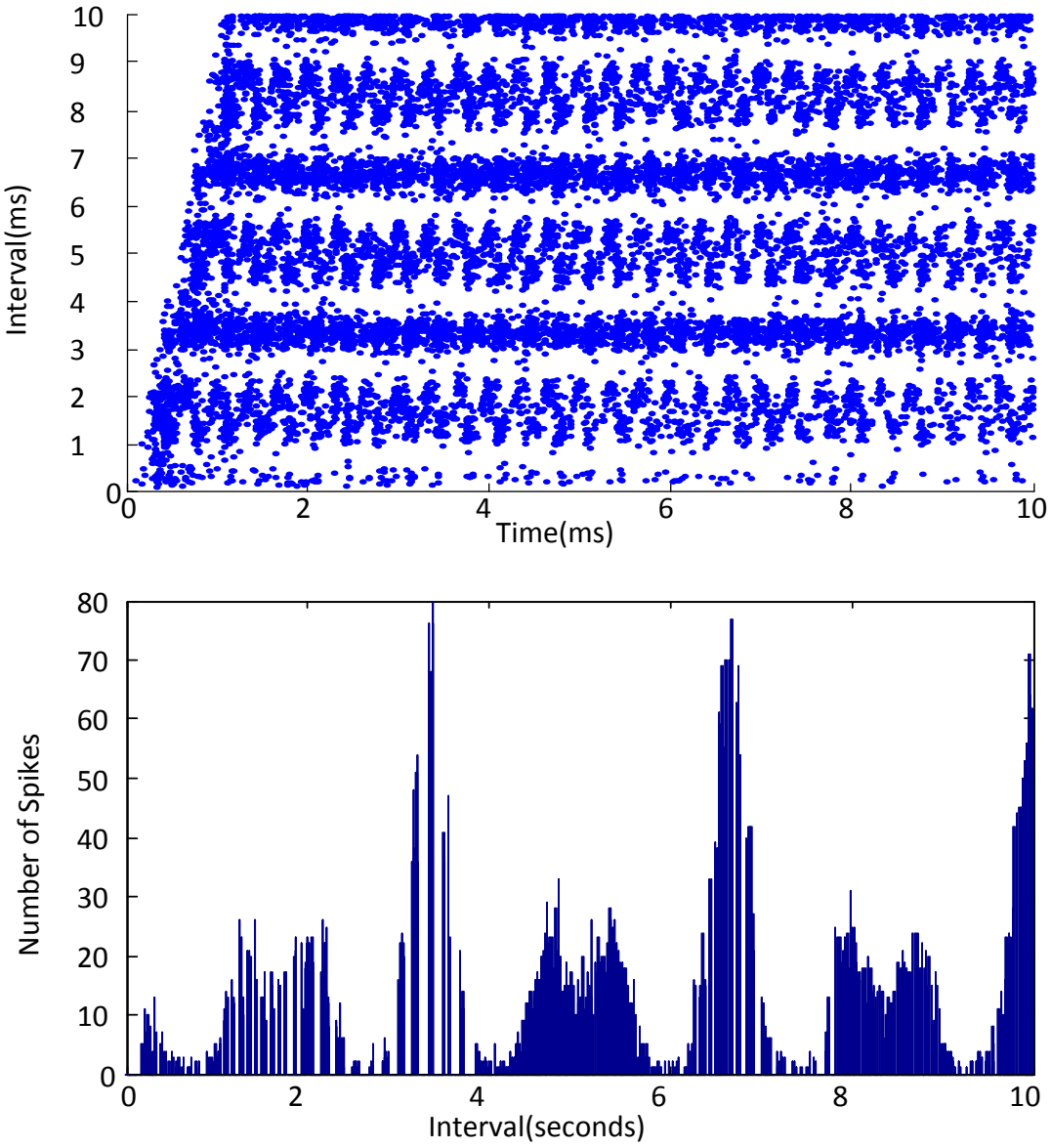

Figure 5.4. Autocorrelationgram and pooled all-order ISI histogram constructed from the autocorrelationgram for a tone complex with fundamental, with a fundamental frequency of 300 Hz.

5.2.3 All-order ISI histograms

The all-order ISI histogram is used to neutralise the effects of loudness of the stimulation signal in the pitch detection and is phase invariant to the stimulus signal while being able

to predict F0. Figure 5.5 shows the construction of all-order ISI histogram and compares this to the construction of a first-order ISI histogram.

The stimulation signal produces a nerve fibre spike train pattern of which the statistics can be shown as a PST histogram. The interval on the PST histogram between two consecutive spikes is used to produce a first-order ISI histogram. The construction of the all-order histogram shows the intervals between the consecutive spikes and several orders of non-successive spikes (Rodieck, Kiang and Gerstein, 1962, Cariani and Delgutte, 1996a). An all-order ISI histogram construction is shown in Figure 5.5(b). The interval between the two consecutive spikes is measured, then the interval between alternating spikes, then the interval between every third spike, and so on. This is continued until the time interval between all spikes are measured to produce an all-order ISI histogram.

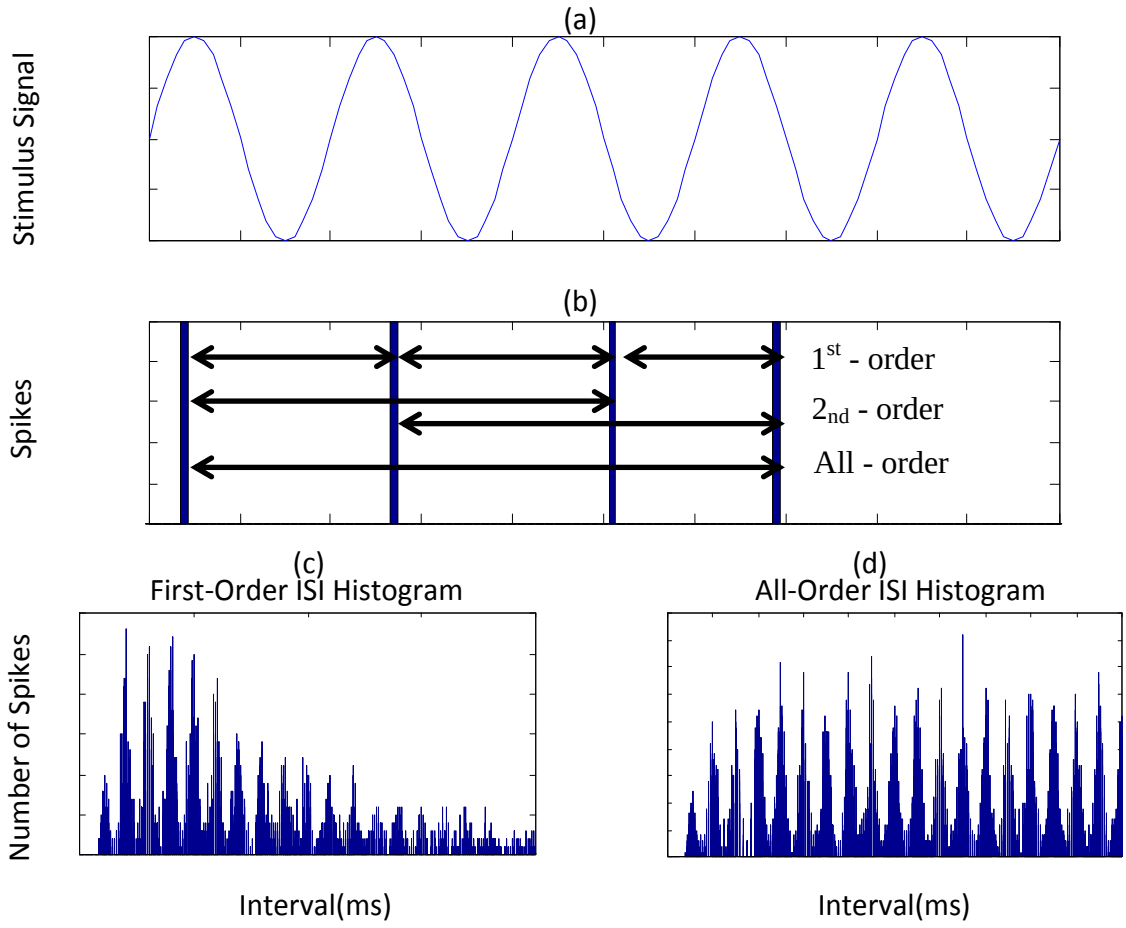

Figure 5.5. Representation of the construction of an all-order interspike interval histogram for a pure tone stimulus. (a) represents the pure tone stimulus. (b) shows the measurement of time difference between spike times for construction of multiple different order interspike interval histograms. (c) represents a first order ISI histogram and (d) represents an all-order ISI histrogram. These figures were obtained from the NH model in chapter 3.

5.2.4 Template matching

The pooled all-order ISI histogram for a pure tone stimulus is matched to a best fitting periodic template by comparing two multiple periodic templates from 0.1 ms to the length of the simulated signal (if the signal is 100 ms long, the last period template will correspond to 100 ms), to predict the F0 fundamental. The template is constructed by using a selected fundamental frequency period and multiples of the selected period. The weighted number of nerve fibre spikes correlating to the template is calculated over the weighted number of total nerve fibre spikes.

The weighting of the nerve fibre spikes is an exponential decay constant of 3.6ms for the interval histogram. Larsen et al. (2008) found this value to minimise the number of octave and sub octave errors in pitch estimation, and the auditory system is unable to detect long intervals with a lower limit of 33.3 ms (corresponding to 30 Hz) (Pressnitzer, Patterson and Krumbholz, 2001). The interval histogram bin values are weighted by being multiplied by the exponential decay function as shown in the equation below:

$$NOI_{weighted}(i) = NOI(i)e^{\frac{-1}{0.0036 \times Interval(i)}} \tag{5.1}$$

where NOI is the bin count or number of inter-spike intervals in a bin of the ISI histogram, i is the bin index, and Interval(i) is the interval time across the ISI histogram. Figure 5.6 shows the construction of the contrast ratio, which is the correlation of all the templates and the weighted ISI histogram. Figure 5.6 below shows the ISI histogram and the weighted results from equation 5.1 in panels a and b. The predicted pitch in Figure 5.6 is the template with the highest contrast ratio in panel c, which is shown below on how the contrast ratio is calculated using the periodic template. The contrast ratio is calculated from the following equation:

$$ContrastRatio(i_f) = \sum_{i=1}^{n} \frac{NOI_{weighted}\,(I_{Tu}(i_f) \times i)}{TNI}, \tag{5.2}$$

where i_f is the indicator of the frequency (predicted F0) of the current template, TNI is the total number of intervals in the ISI histogram, I_{Tu} is the bin of the ISI histogram corresponding to the period of the frequency of the current template, and n is an indicator which correlates to the length of the ISI histogram for the specified template.

The contrast ratio graph shows the contrast ratios for the possible predicted F0 values over the stimulus signal. The biggest peak has the most number of added multiples corresponding to that interval. The template corresponding to a frequency of 338 Hz has the highest contrast ratio in Figure 5.6 c. The second and third highest peaks are multiples of the predicted frequency template of 338 Hz. This is due to clustering of interval peaks in the ISI histogram which relate to the period of multiples of this frequency. The x-axis in Figure 5.6 c is the testing template frequency. For the ease of reading, the x-axis will be shown, in later contrast ratio graphs (in this chapter), as the stimulus frequency over the template frequency. This will illustrate the number of periods by which the predicted frequency, from the template matching approach, varies from the stimulus frequency.

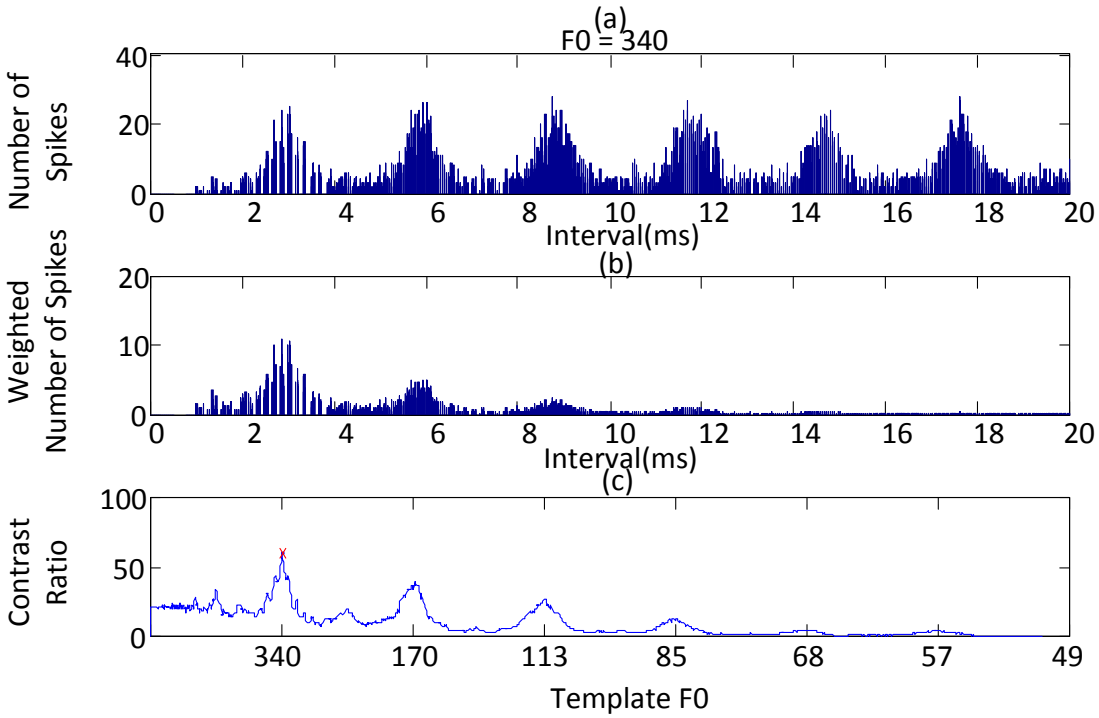

Figure 5.6. Pitch estimation based on pooled all-order inter-spike interval distributions for F0 = 340 Hz. The pooled all-order inter-spike interval distributions are weighted and then compared to the periodic templates to give the contrast ratio. The predicted pitch was 338 Hz (Cedolin and Delgutte, 2005).

5.3 RESULTS

This section looks at the pitch prediction results using the pitch prediction algorithm in addition to models from the previous chapters. The results for the normal hearing auditory model, the cochlear implant travelling wave processor models (for both the travelling wave processor and the alternative travelling wave processor), and the vocoder speech processing model will be evaluated.

5.3.1 Pitch Prediction results for normal hearing

Pitch prediction for normal hearing results are illustrated using the ISI histogram and the contrast ratio as shown in Figure 5.7. A pure tone of 700 Hz at 60 dB was used as the stimulus input to the normal hearing model. The contrast ratio predicts that the pitch heard by the listener will be 694 Hz. For a lower frequency of 110 Hz, in Figure 5.8, the results of the contrast ratio has a pitch prediction of 111.5 Hz. The contrast ratio for the 110 Hz pitch prediction has fewer multiples than those in the 700 Hz prediction; this is possibly due to the exponential decay weighting used in the pitch prediction algorithm. The frequency of 110 Hz has a longer time period; so the ISI bins, corresponding to the period and the multiples of the period of the frequency in question, are supressed more heavily by the weighting function than those of a shorter time interval as seen by equation 5.1. For example, in equation 5.1, bin with “x” number of spike would be more heavily weighted if it had an interval of 10 ms compared to a bin with an interval of 1 ms; due to equation 5.1.

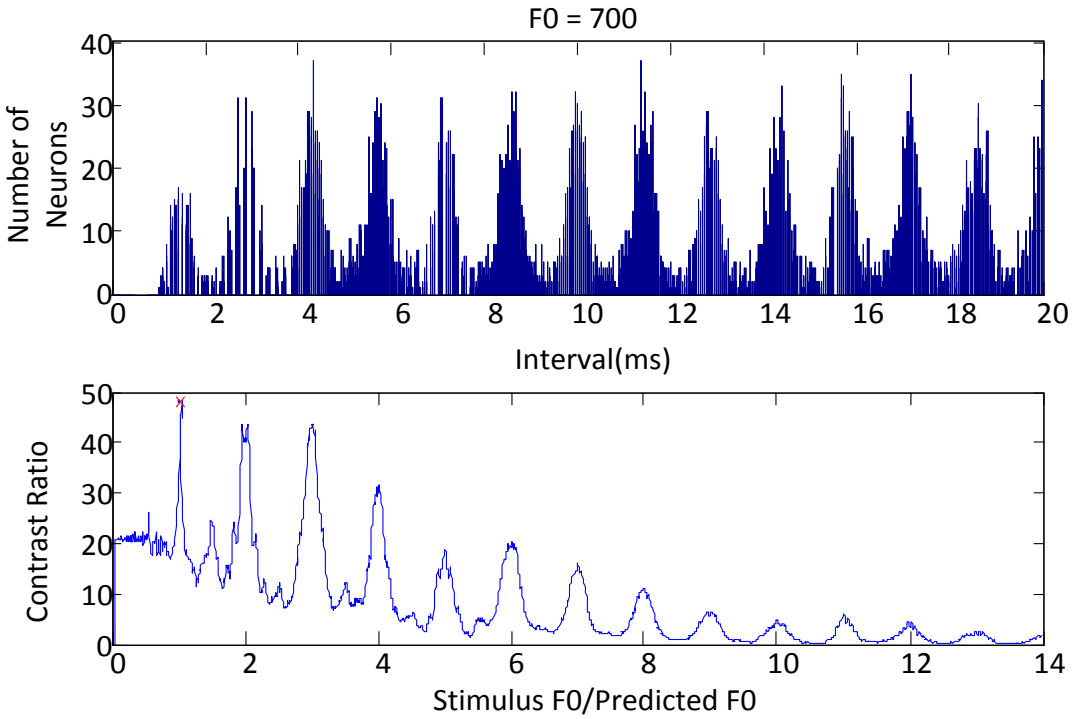

Figure 5.7. ISI histograms and pitch prediction contrast ratio for a 700 Hz pure tone using the normal hearing model from chapter 3. Pitch predicted was 694 Hz which is shown by the red cross on the contrast ratio graph.

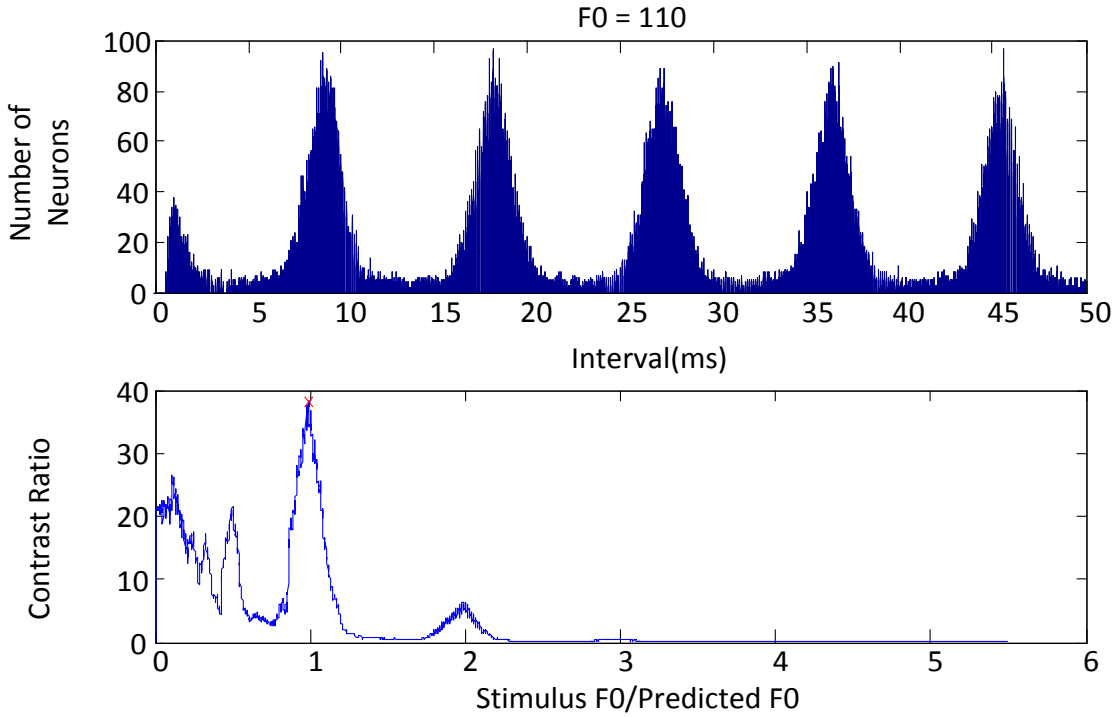

Figure 5.8. ISI histograms and pitch prediction contrast ratio for a 110 Hz pure tone using the normal hearing model from chapter 3. Pitch predicted was 111.5 Hz which is shown by the red cross on the contrast ratio graph.

Pitch estimation for normal hearing is shown in Figure 5.9 below. The pitch is estimated from 100 Hz to 1300 Hz. The pitch is accurately predicted until 1298 Hz, with maximum error of 34.8 Hz, but then rapidly drops off. Above this frequency, the pitch predicted is one third of the input pitch of the stimulus input.

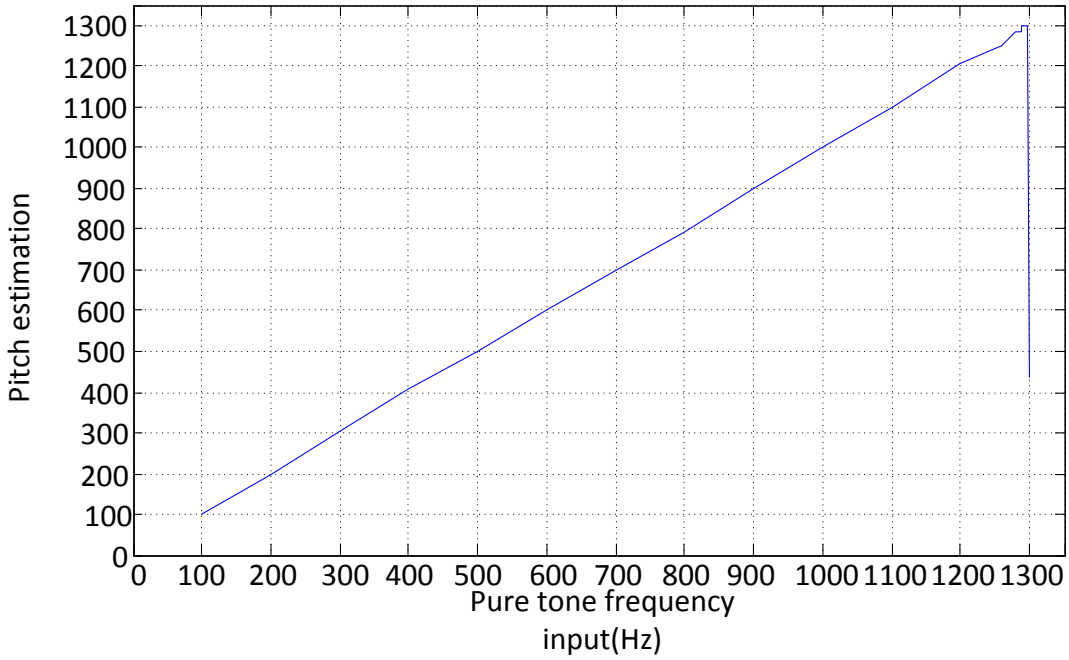

Figure 5.9. Pitch estimation for the normal hearing model.

Figure 5.10 shows what happens at higher frequencies than the apparent limit seen in Figure 5.9. The decrease in the first and second clusters in the ISI histograms are illustrated in Figure 5.10. The height of the first peak (referring to the period of the stimulus frequency around the period of 0.769 ms in Figure 5.10) and the height of the second multiple decreases as the frequency of the stimulus increases. This may be due to the stimulation period getting close to the refractory period of the nerve fibres, which would leave nerve fibres with the inability to lock onto every period of the stimulus signal.

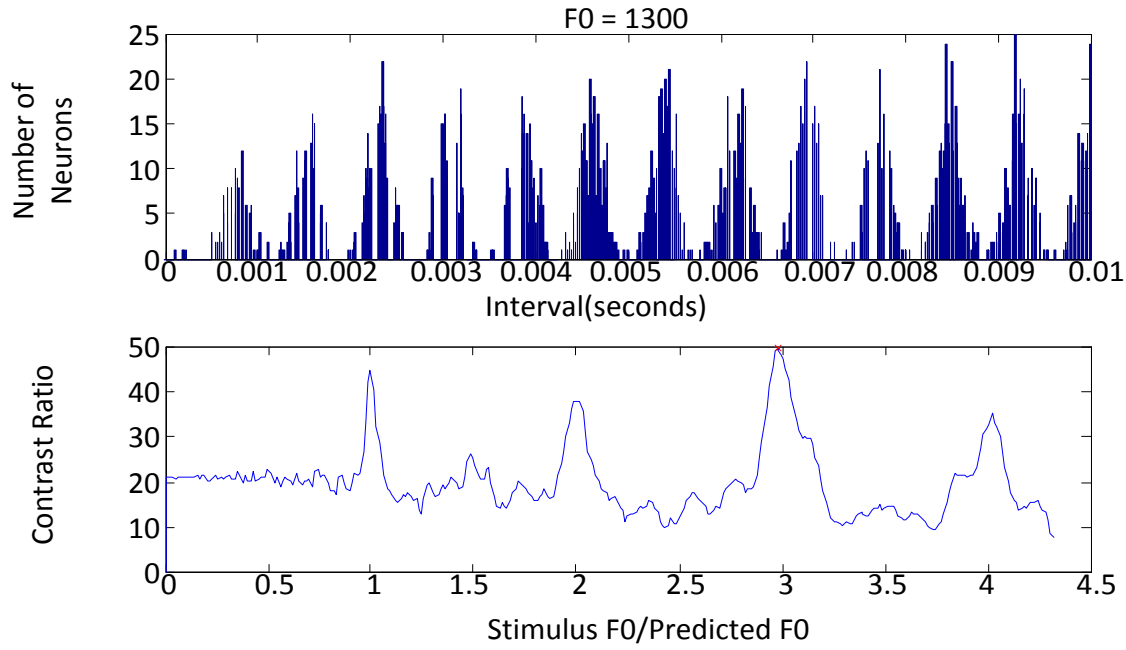

Figure 5.10. ISI histogram and prediction results for a pure tone stimulus of 1300 Hz using the normal hearing model from chapter 3.

5.3.2 Travelling wave cochlear implant processor predictions

This section examines the neural spike train patterns predicted for the TW CI processor, and the predicted pitch for pure tones. For 200 Hz and 1000 Hz, the pitch predicted is 199.6 Hz and 990.09 Hz respectively, for the electrode configuration described below.

Prediction of the 200 Hz pure tone pitch used 8 electrodes, equally spaced, placed from 12.25mm to 25.75mm along the length of the basilar membrane with a stimulation frequency of 1800 pulses/second. ISI histograms results in Figure 5.11 shows a peak around the period of the stimulus. Moreover, the multiple clusters are sharp at the interval of the stimulus frequency and there are secondary peaks corresponding to the pulse period that are clustering around the interval of the stimulus period. These clusters around the stimulus intervals are shifted by pulse period on either side of the interval. The effects can be observed in the contrast ratio graph in Figure 5.11, which shows the second and third highest values of the contrast value being shifted away from the period peak by the pulse rate of the cochlear implant. Therefore, an increase in stimulation pulse rate of the cochlear implant causes the clustering of peaks to be closer to the period of the stimulus for the travelling wave cochlear implant processor.

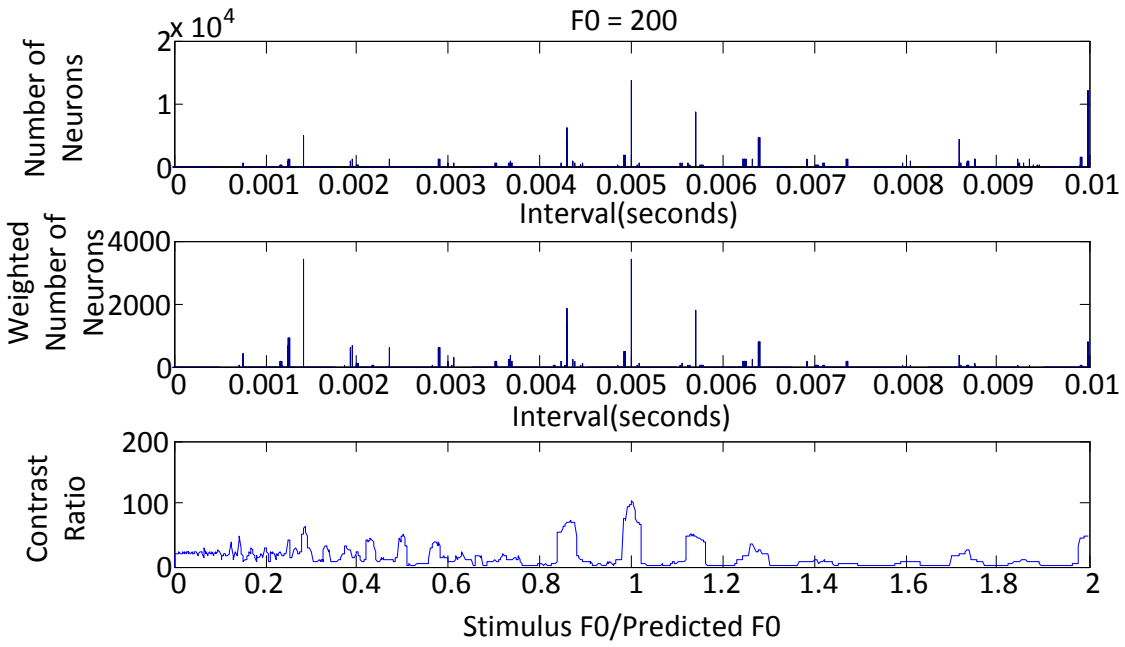

Figure 5.11. Pitch estimation of a pure tone of 200 Hz using 8 electrodes with a pulse rate of 1800 pulses/second using the **TW CI processor model**. The pitch predicted is 199.6 Hz.

Figure 5.12 and Figure 5.13 shows the pitch estimation of a pure tone of 1000 Hz using two different electrode configuration patterns: 11 electrodes equally spaced across the length of the cochlea with a pulse rate of 1309 pps and 4 electrodes equally spaced across the length of the cochlea with a pulse rate of 3600 pps. The pitches predicted are 327.86 Hz and 502.51 Hz respectively. It can be seen from Figure 5.12 and Figure 5.13 c that the ISI histograms have no intervals at the period of the stimulus signal. Figure 5.13, using 4 electrodes, has peaks at the multiples of the period of the stimulus. However, the pitch is still not predicted. Figure 5.14 will show that placement and number of pulse rate is important for correct pitch perception.

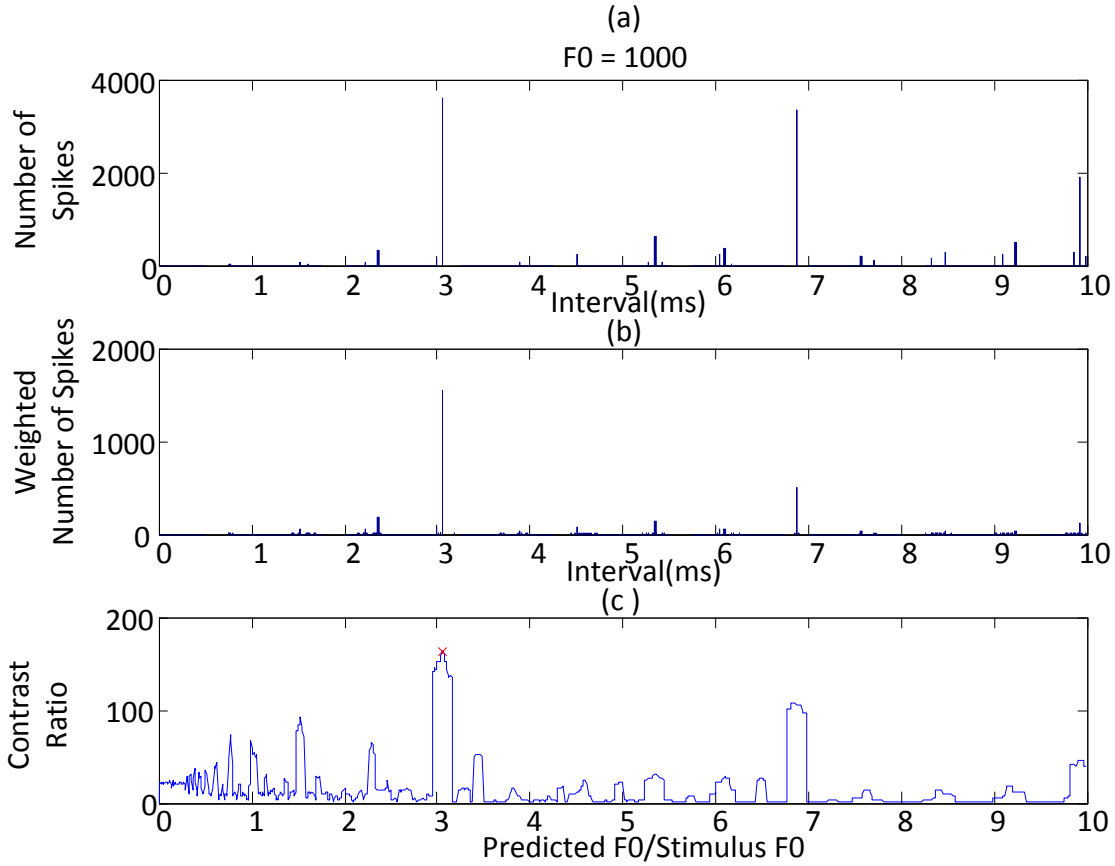

Figure 5.12. Pitch estimation of a pure tone of 1000 Hz using 11 electrodes (1:2:21) with a pulse rate of 1309 pulses/second using the **TW CI processor model**. The pitch predicted is 327.86 Hz.

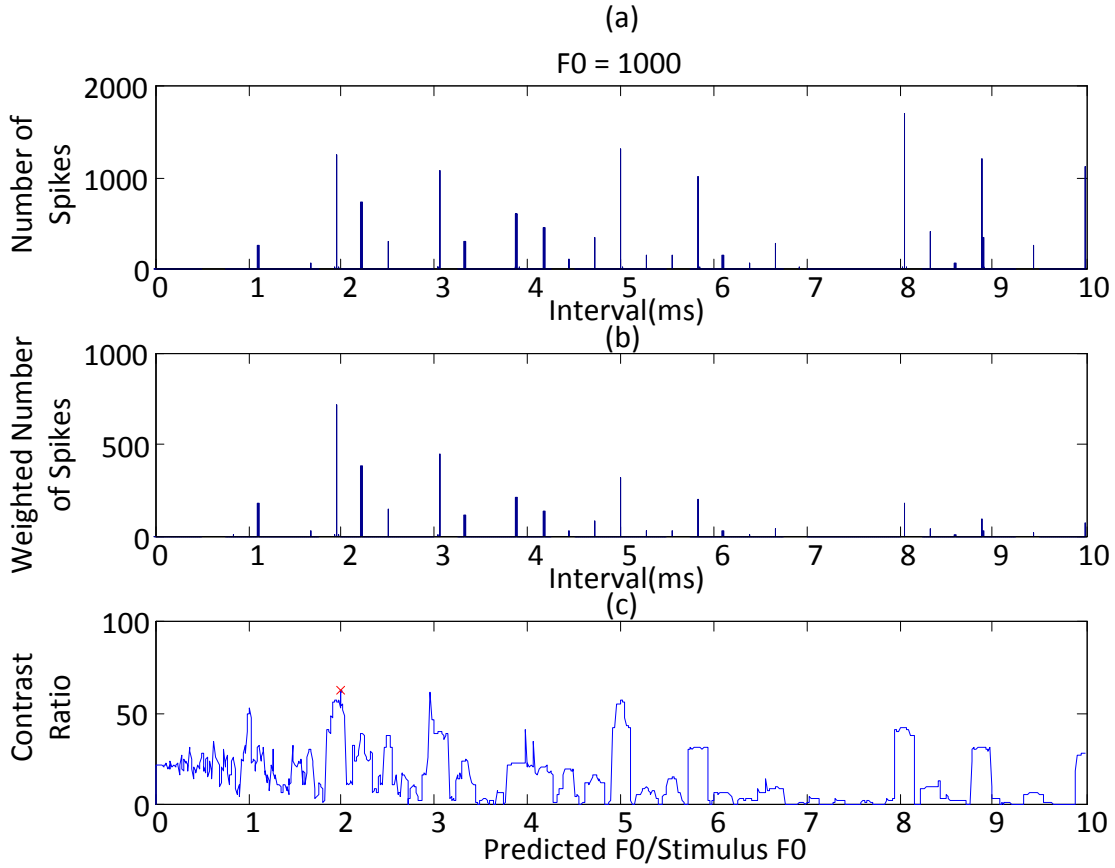

Figure 5.13. Pitch estimation of a pure tone of 1000 Hz using 4 electrodes (electrodes 2, 8, 14, and 20) with a pulse rate of 3600 pulses/second using the **TW CI processor model**. The pitch predicted is 502.51 Hz.

The 1000 Hz prediction (Figure 5.14) was done using 4 electrodes (electrodes 8, 10, 12, and 14), allowing for a maximum sampling rate of 3600 pulses/second from position 15.25 mm to position 19.75 mm, the positions closest to the peak of the travelling wave for the pure tone stimulus of 1000 Hz.

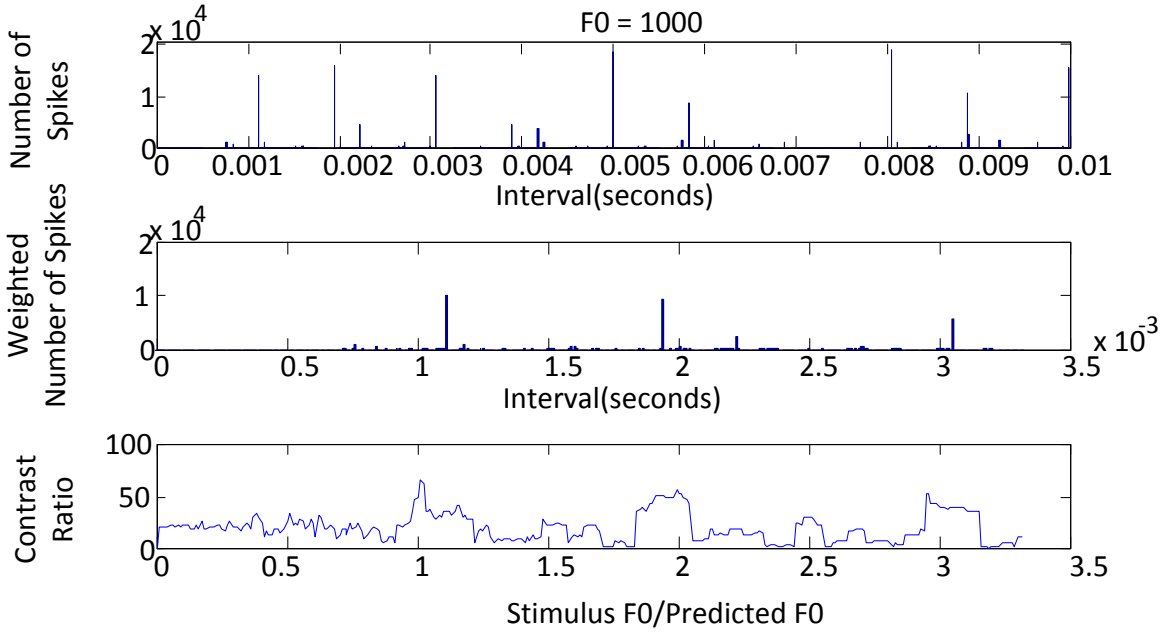

Figure 5.14. Pitch estimation of a pure tone of 1000 Hz using 4 electrodes with a pulse rate of 3600 pulses/second using the **TW CI processor model**. The pitch predicted is 990.09 Hz.

The TW CI processor requires fewer electrodes to be placed near to the peak displacement (where displacement of the basilar membrane is observed) of the travelling wave in order to enable temporal encoding to detect the pitch of the stimulus signal by causing nerve fibres to lock onto the period of the stimulus signal and convey pitch. If the temporal sampling is lower than the period of the stimulus signal, no phase locking of the nerve fibres will occur, as the temporal properties of the travelling wave is distorted by the under sampling of the stimulus signal. The energy of the travelling wave is at a maximum at peak displacement (Duifhuis, 2012) for an instant in time (as this peak shifts from base to apex of the basilar membrane). Therefore, in order to convey this temporal information and capture the information present, a higher sampling rate is required and obtained by decreasing the number of electrodes to gain a higher sampling rate.

5.3.3 VOC processor predictions

This section shows the results for the VOC speech processor with a single electrode stimulation with varying pulse rate. Figure 5.15 and Figure 5.16 show the pitch prediction of stimuli with a pulse rate of 100 pps and 300 pps respectively. The interval histogram shows a sharp clustering of intervals around the period of the pulse rate. The pitch predicted was 98 Hz and 303 Hz for the respective pulse rate.

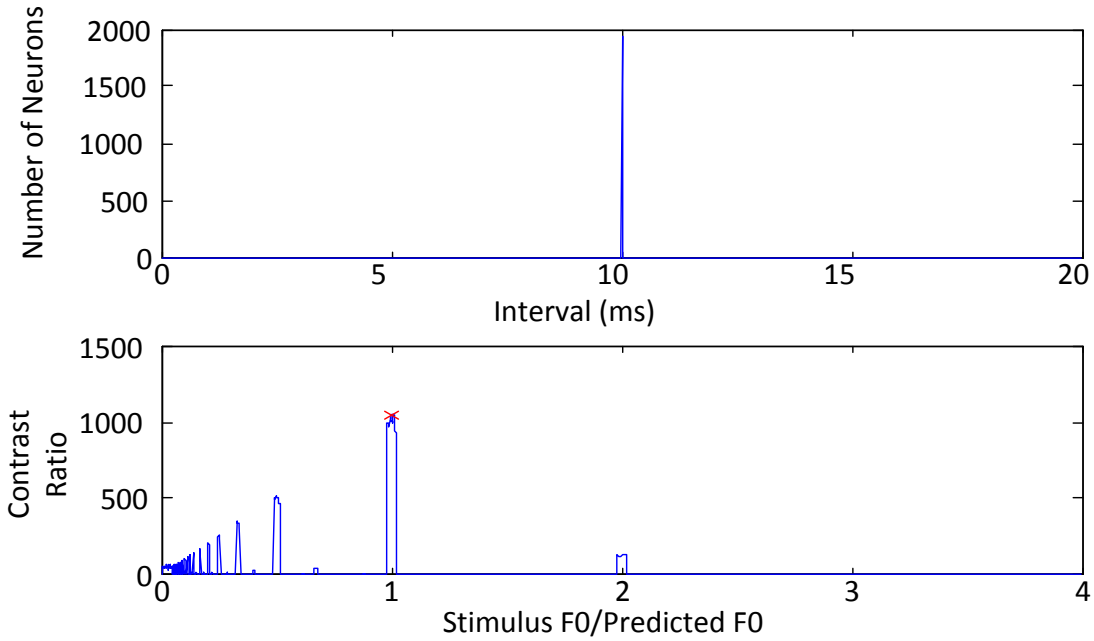

Figure 5.15. ISI histogram and pitch prediction for signal electrode stimulation with a pulse rate of 100 pps using the **VOC processor model**. The pitch predicted is 98 Hz.

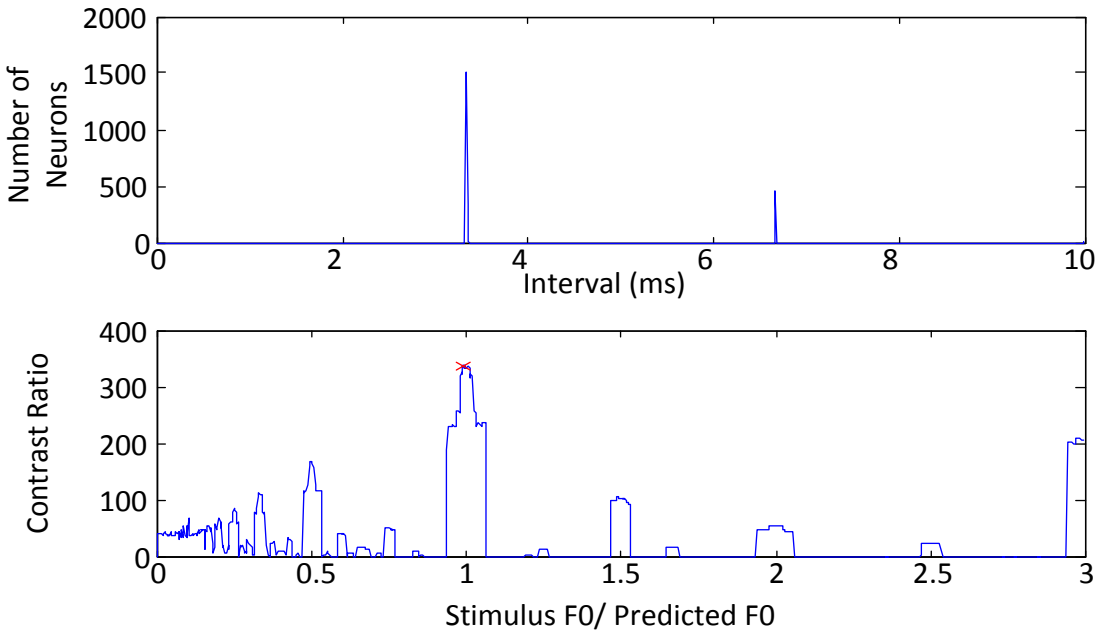

Figure 5.16. ISI histogram and pitch prediction for signal electrode stimulation with a pulse rate of 300 pps using the **VOC processor model**. The pitch predicted is 303 Hz.

Increasing the stimulus pulse rate above 300 pps causes nerve fibres to cluster at the multiples of the interval period and not the fundamental period of the pulse rate. This is observed in Figure 5.17, where a small cluster is seen at the 2 ms interval corresponding to

the period of a 500 pps stimulus. However, much larger clusters are observed at multiple periods of the stimulus (4 ms, 6 ms, and 8 ms). The pitch predicted is 249 Hz. The decrease in the number of nerve fibres locking onto the pulse of the stimulus could be due to the pulse rate approaching the refractory period of the modelled nerve fibres, causing a higher number of neurons to lock onto multiples of the period of the pulse rate.

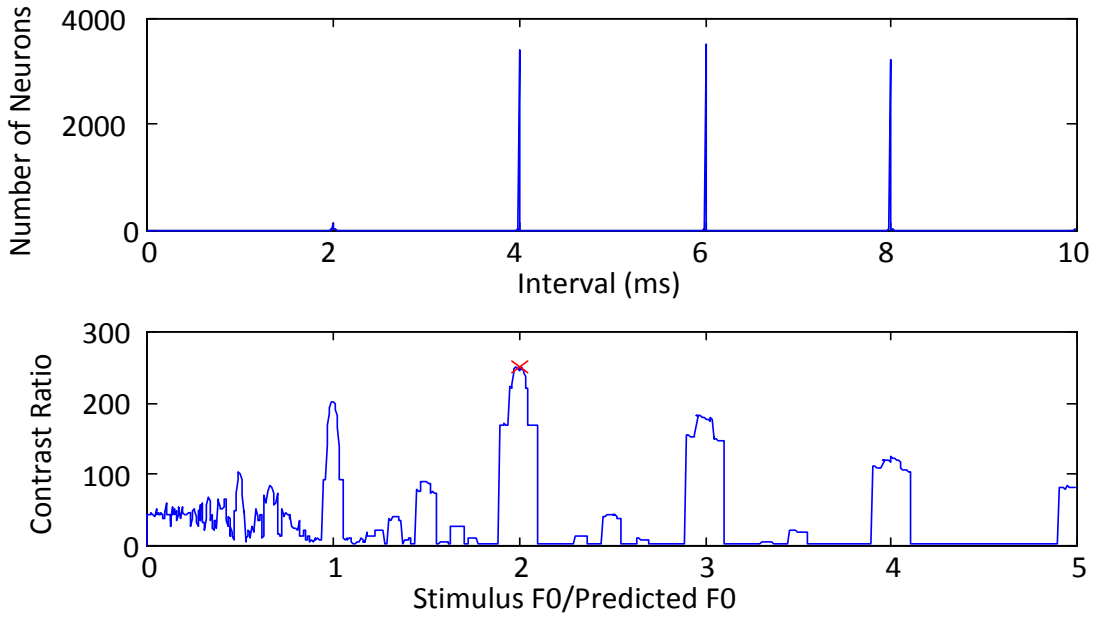

Figure 5.17. ISI histogram and pitch prediction for signal electrode stimulation with a pulse rate of 500 pps using the **VOC processor model**. The pitch predicted is 249 Hz.

In these simulations, there is no implied signal pure tone frequency. For a particular frequency of pure tone, the electrode stimulate would shift along the tonotopic axis (place pitch). As the pulse rate is tracked by the neural spike train patterns, it is worthwhile to consider whether rate pitch could be encoded by electrical stimulus pulse frequency.

The VOC processor pulse rate for a single electrode was varied from 100 pps to 700 pps in 100 pps steps and the pitch was predicted; see Figure 5.18. The pitch is predicted to be the same as the pulse rate up to 300 pps, where the pitch predicted is 309 Hz. However, for frequencies above 300 Hz, the pitch predicted is below 300 Hz. This corresponds the often-observed 300 pps limit to rate pitch (Shannon, 1983, McDermott and McKay, 1997, McKay et al., 2000, Zeng, 2002).

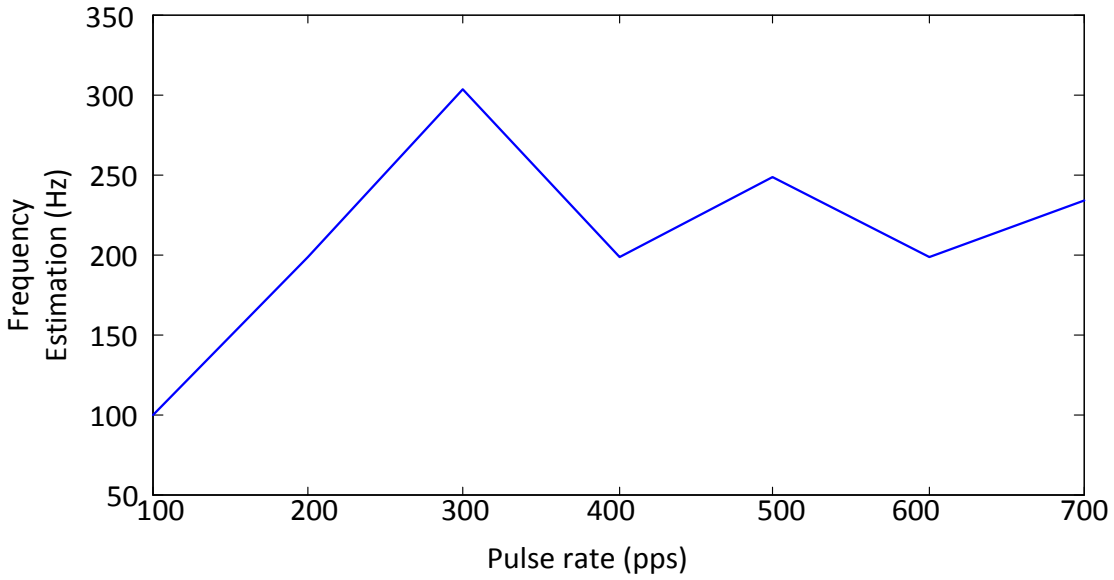

Figure 5.18. Pulse rate versus frequency estimation for the vocoder speech processor.

Venter and Hanekom (2014) showed that pulse rate difference limens track those of normal hearing (i.e. pulse rate difference limens do not suddenly increase above 300 pps) when multiple electrodes are used in a stimulus set and the pulse rate of each electrode is the same as the frequency of the pure tone stimulus that is represented. It was, however, unclear whether rate pitch also improved with multiple electrode stimuli. While Venter and Hanekom (2014) did not observe this in their data, data of Penninger, Kludt, Büchner and Nogueira (2015) does indicate improved rate pitch difference limens with multi-electrode stimuli.

When multi-electrode stimuli are simulated with the model (500 pps rate on each of 6 electrodes in this simulation, encoding a stimulus signal of 500 Hz, and using the "spread mode" stimulation pattern of Venter and Hanekom (2014)), the nerve fibre spiking pattern in Figure 5.19 shows clusters in the ISI histogram at an interval close to 4 ms (corresponding to 250 Hz), and multiples of this (i.e. clusters don't appear at the signal period (2 ms)). The predicted pitch is 253 Hz. This means that (at least in this example), there is no benefit to using multi-electrode stimuli. Multi-electrode encoding of rate pitch would not elicit spike train patterns that carry the intended rate pitch information.

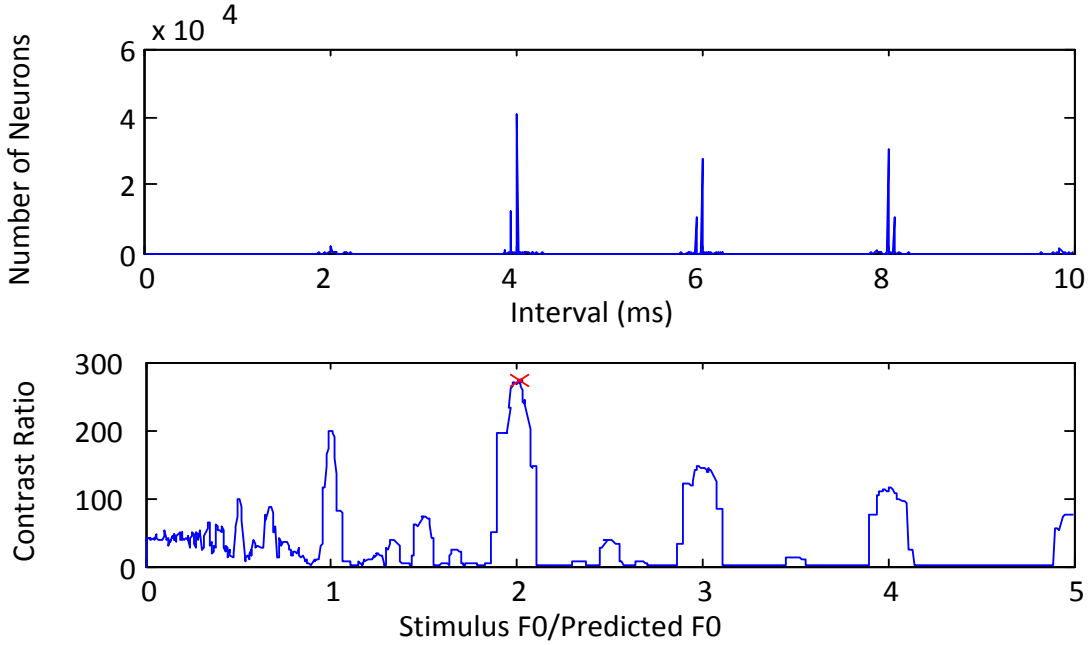

Figure 5.19. ISI histogram and pitch prediction for 6 electrodes stimulation with a pulse rate of 500 pps using the **VOC processor model**. The pitch predicted is 253 Hz.

5.3.4 Alternative CI TW processor prediction

The prediction of the pitch using the alternative CI TW processor is shown in Figure 5.20. The pitch is estimated for frequencies from 100 Hz to 1200 Hz. The pitch is accurately predicted, with a maximum error of 15.9 Hz, until 1060 Hz then drops off rapidly and is unable to predict higher frequencies.

Inability to predict pitch above the threshold of 1060 Hz is due to the first peak of the ISI histogram being substantially smaller than those of the second and third peak. This phenomenon becomes more pronounced as frequency is increased, and as seen with normal hearing listeners, the period of the stimulation rate becomes close to the refractory period of the nerve fibres.

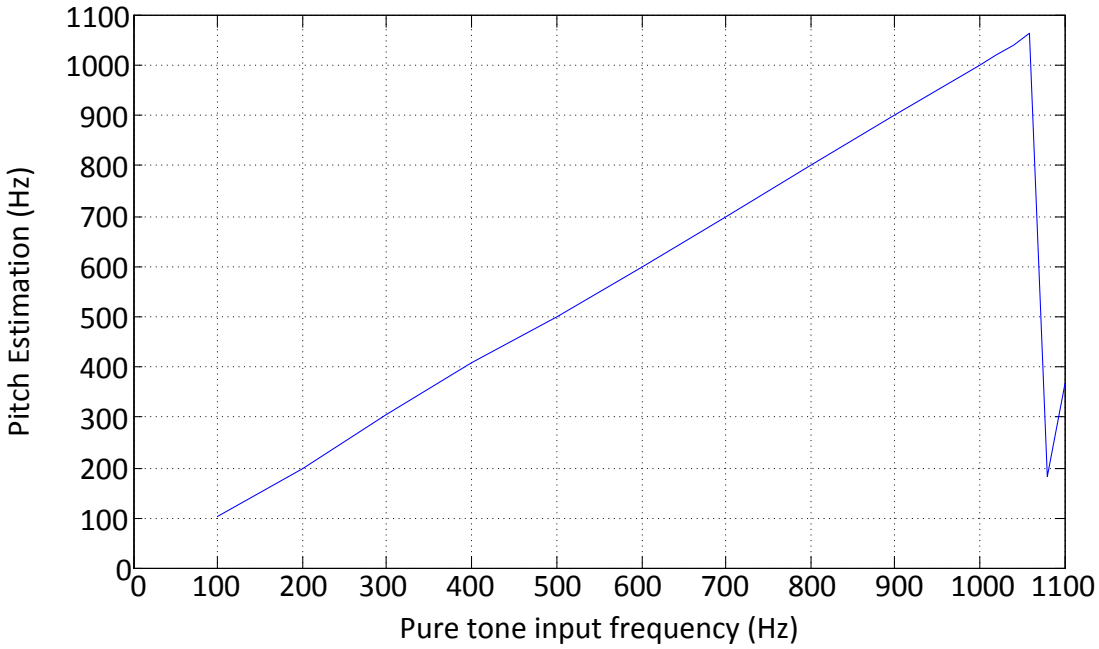

Figure 5.20. Predicted pitch for the CI TW processor. The pitch is predicted for frequencies from 100 Hz to 1200 Hz.

The pitch estimation for 110 Hz and 1000 Hz is shown in Figure 5.21 and Figure 5.22. The all-order ISI histogram is weighted and then used to predict the pitch. The pitch predicted is 115.2 Hz and 1000 Hz respectively for Figure 5.21 and Figure 5.22. The alternative travelling wave processor has a greater number of spikes clustering around the peaks in the ISI histogram. However the clustering's minimum space in between peaks is related to the maximum pulse stimulation period of 69.4 us (14400 pps sampling frequency used for the cochlear implant alternative travelling wave processor).

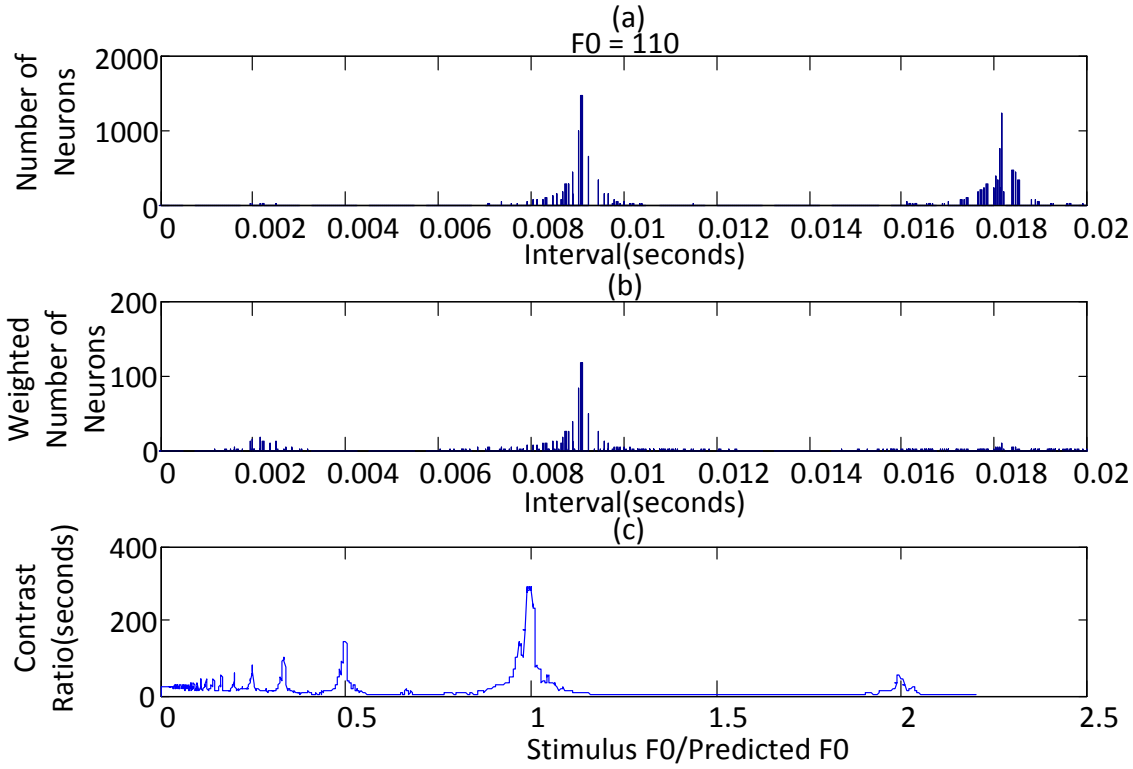

Figure 5.21. Pitch prediction results of 110 Hz pure tone input for the alternative travelling wave processor. The result is 115.2 Hz.

ISI histograms for the 1000 Hz pure tone stimulus are sharper than those of the 110 Hz ISI histograms. However pitch is still accurately detected by the algorithm, which may mean that only the main interval period and the multiples of the stimulus are required to convey pitch information using temporal encoding, not the clustering of the intervals around the peak of the ISI histograms (spikes that appear around the peak rather than at the peak). The effects of this phenomenon to hearing perception, during electrical TW stimulation, are unknown for a cochlear implant user; the sharp peaks in the interval histogram are not observed in normal hearing listeners, but the intervals are more widely clustered and distributed around the main peak of the interval.

Predicting pitch using the contrast ratio in Figure 5.22 c for the 1000 Hz pure tone may be an issue. The figure shows contrast ratio peaks of similar amplitude at 1 and 2, so that pitch may be predicted to be either 1000 Hz or 500 Hz. A similar observation can be made for normal hearing pitch predictions as the frequency is increased. This is due to temporal encoding reaching the limit of nerve fibres' ability to phase lock onto the period of the stimulus signal and the pitch detection algorithm observing more nerve fibres locking onto the second or third period of the stimulus signal.

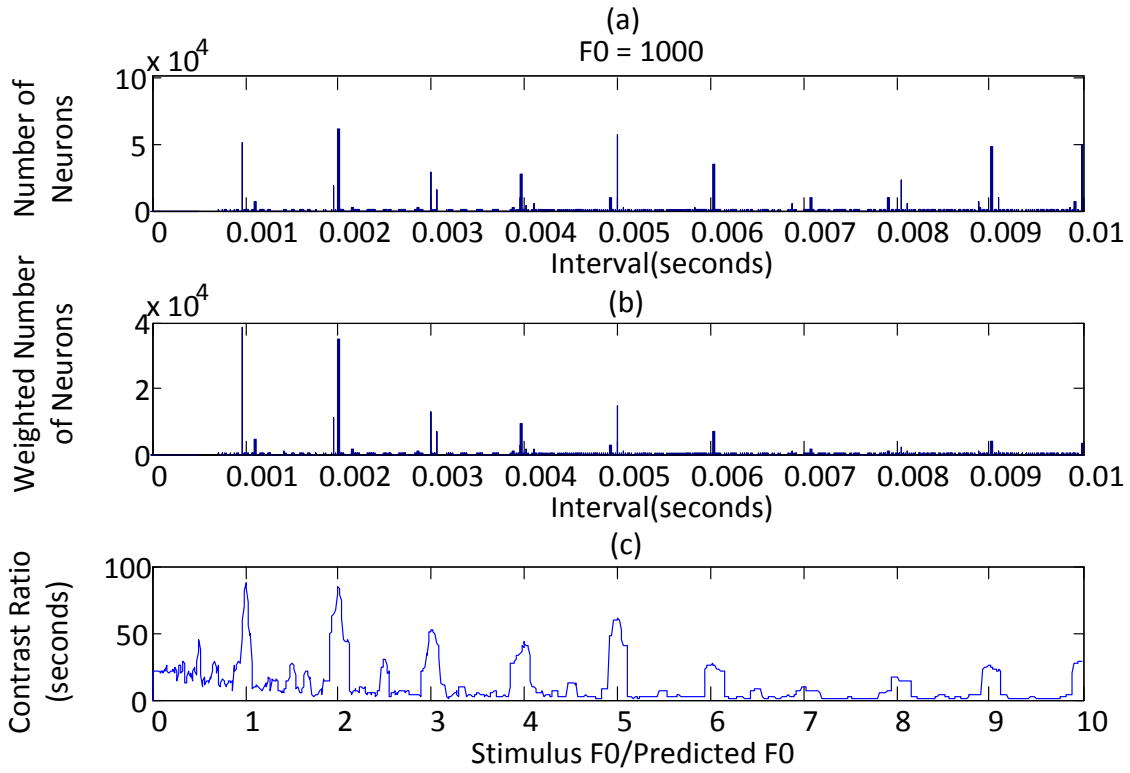

Figure 5.22. Pitch prediction results of 1000 Hz pure tone input for the alternative travelling wave processor.

The pitch predictions above show that the template match decoding technique works in both the normal hearing auditory model and the electrical stimulation models for cochlear implant users. The NH model suggests that the travelling wave is able to accurately encode the pitch of the stimulus input with great accuracy up to frequencies of 1300 Hz for normal hearing, which correlates to results of maximum upper limit of temporal pitch estimation from Cedolin and Delgutte (2005). The travelling wave processor for CI may be able to encode temporal information above 1000 Hz; 1060 Hz predicted for electrical stimulation model by the alternative wave processor. This is true at least in model predictions, but needs to be tested in an animal model.

The CI travelling wave processor, however, requires knowledge of the stimulus signal (such as peak displacement position, frequency, and stimulation duration) in order to define the electrodes activated for stimulus and pulse stimulation rate. The displacement of the travelling wave is required, as the active electrodes need to take full advantage of the spatial resolution by activating the electrodes at positions where displacement information can be transmitted to the nerve fibres. The frequency of the stimulus is required to set the

required stimulation pulse rate of the processor, as having a lower pulse rate than the stimulus signal may hinder the ability of the travelling wave processor to temporally encode information to the nerve fibres along the length of the cochlea (the entire travelling wave will be temporally down sampled). The stimulation duration is required to determine at which point in time the electrode configuration pattern and temporal pulse rate will have to update in order to encode a complex multi-tone signal to the nerve fibres. The travelling wave processor will require a complex processing algorithm to arrive at the electrode stimulation profile. This will require substantial processing power, while processing power to process complex signals which may not be available in cochlear implant processors.

Pitch prediction accuracy as shown in the figures above is unexpected. An explanation for the high accuracy of the pitch estimation observed with the alternative travelling wave electrical model could be the bin size used to develop the all-order ISI histograms. The bin size for the all-order ISI histograms was 0.1 ms, which affects the accuracy of the pitch prediction algorithm, because at high frequencies, the step size between frequency predictions is large. For example, if the frequency predicted was 1000 Hz, the closest other frequencies that the pitch predictor can predict is 909 Hz (1.1 ms) and 1111 Hz (0.9 ms).

The larger bin size was chosen to save computational simulation time, but with a bin size of 0.01 ms, the pitch predictions for the 1000 Hz simulation was re-run. The pitch estimated was 990.09 Hz, rather than the 1000 Hz, and the closet other frequencies that pitch predictor can predict is 980.32 Hz and 1000 Hz. This shows a decrease in the accuracy of the pitch prediction algorithm due to bin size choice.

5.3.5 Missing fundamental predictions

This section considers pitch predictions of tone complexes with a missing fundamental, using the normal hearing model and the alternative cochlea travelling wave model. Figure 5.23, Figure 5.24, and Figure 5.25 show the prediction of the pitch of a missing fundamental tone complex with the fundamental at 200 Hz, 500 Hz, and 1000 Hz respectively. The panels on the left show the normal hearing predictions and the panels on the right, the results for the alternative travelling wave model. The 200 Hz ISI histograms produced for the missing fundamental tone complex show tall peaks at the period corresponding to the fundamental frequency and smaller clusters between the taller peaks. This correlates to the normal hearing predictions for spike statistics.

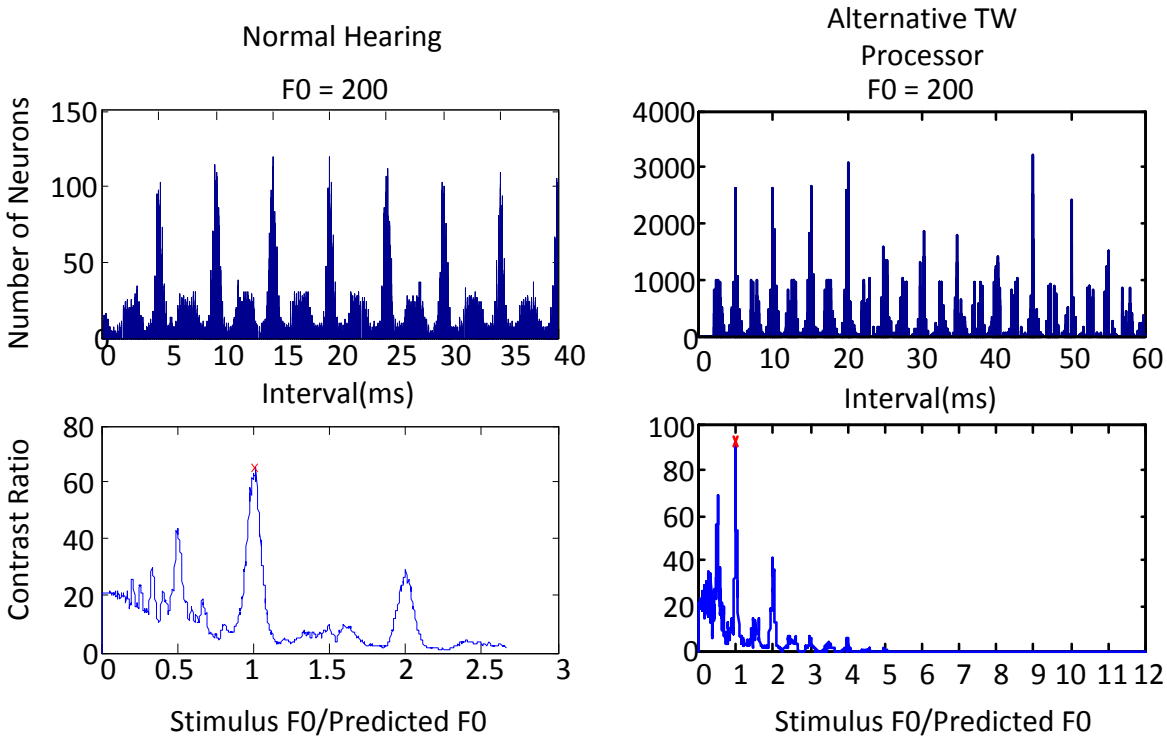

Figure 5.23. ISI histogram and pitch prediction for a missing fundamental signal with a fundamental frequency of 200 Hz. The pitch predicted is 220 Hz.

An increase in frequency of the fundamental causes the tall peaks at the period of the fundamental frequency to greatly decrease, as seen in Figure 5.24, and Figure 5.25. The results of the 1000 Hz simulation in Figure 5.25 show peaks at 1ms and the multiples of the period of the stimulus. However, there is great variance in the height of the peaks.

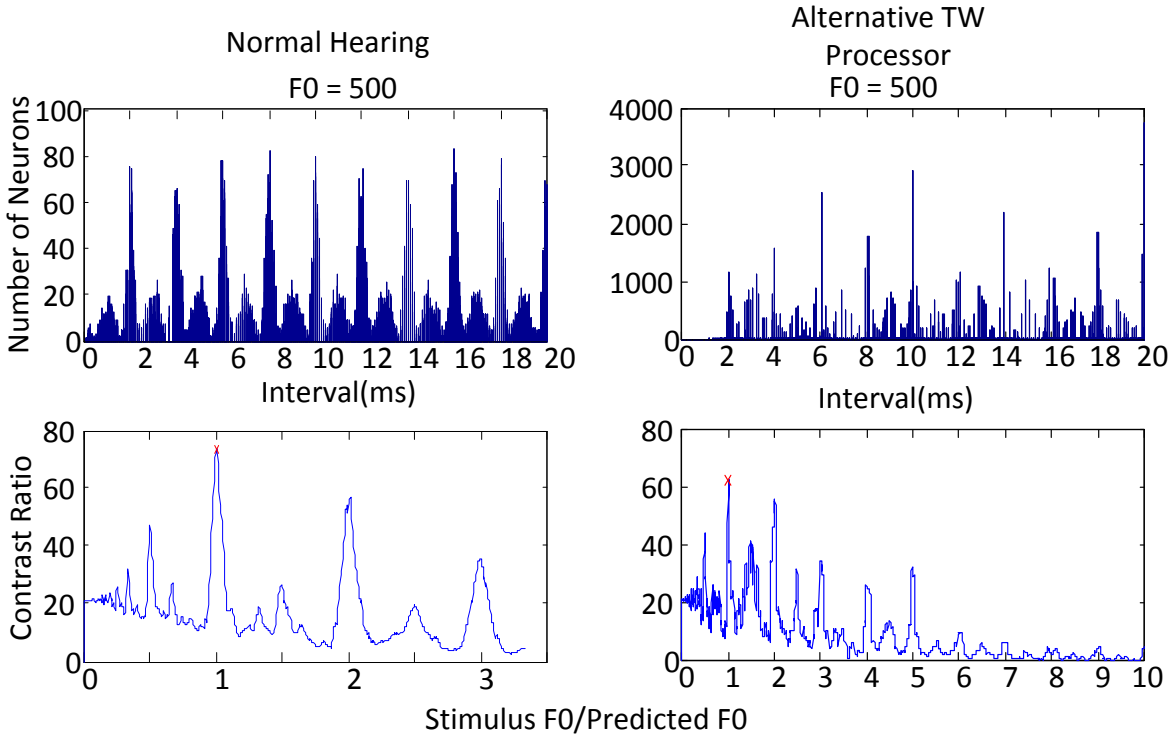

Figure 5.24. ISI histogram and pitch prediction for a missing fundamental signal with a fundamental frequency of 500 Hz. The pitch predicted is 515 Hz for the alternative travelling wave processor.

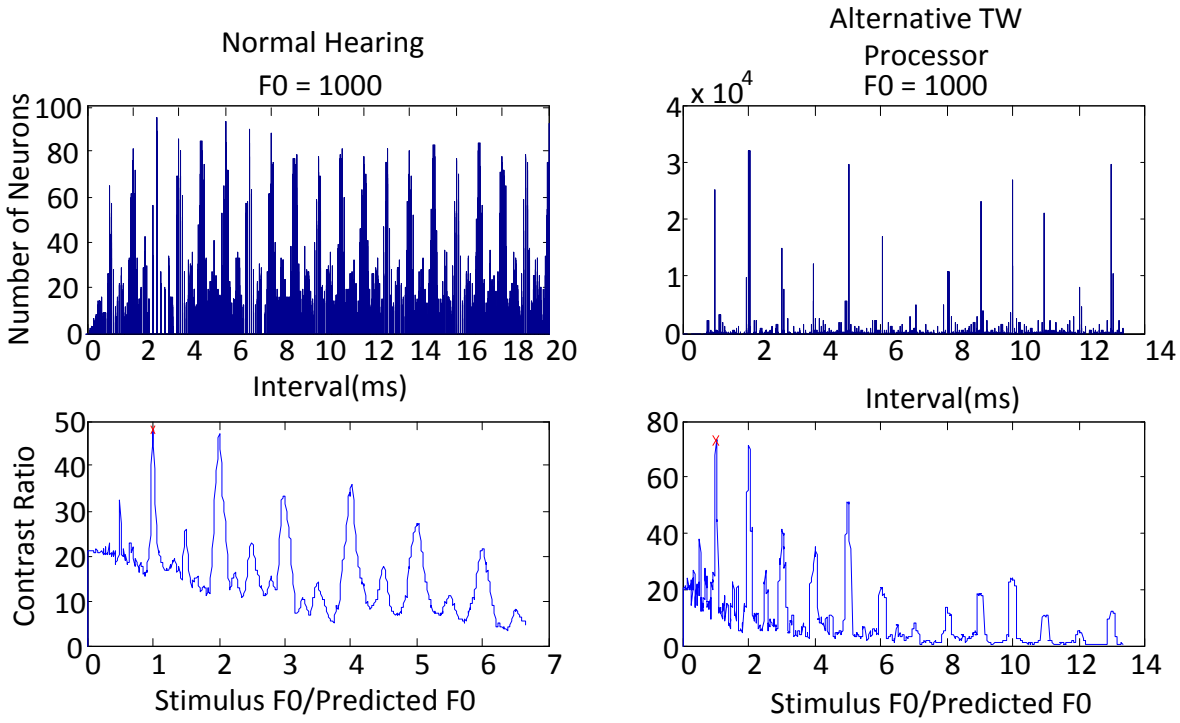

Figure 5.25. ISI histogram and pitch prediction for a missing fundamental signal with a fundamental frequency of 1000 Hz. The pitch predicted is 972 Hz for Alternative travelling wave processor.

5.4 DISCUSSION

5.4.1 Chapter objectives

This chapter evaluated the normal hearing model and the electrical stimulation models using a temporal pitch perception prediction model. The delivery of temporal information of an acoustic stimulus using electrical stimulation is important to understand and further develop CI implant processing strategies. Understanding which methods may be used for the delivery of temporal information may lead to more robust cochlear implant encoding of temporal pitch in all different environments.

5.4.2 Evaluation of the pitch algorithm using the normal hearing model

The normal hearing model is used to generate the nerve fibre response from a pure tone acoustic stimulus. The pitch prediction model predicted the correct pitch up to 1298 Hz using the predicted nerve fibres spiking pattern, which correlates to results of maximum upper limit of temporal pitch estimation from Cedolin and Delgutte (2005). The high accuracy of the model needs further consideration, as the model would outperform the normal listener. This may be due to the nerve fibre model, which may lack a high enough level of noise and jitter in the nerve fibres firing, even while the model produces noise levels which correlate with auditory nerve fibre noise in literature (Kelly, Johnson, Delgutte and Cariani, 1996, Zilany et al., 2009). The noise added to the normal hearing model has a Gaussian distribution with a mean of 0. This noise is added to the slow power-law adaptation path of the IHC-AN synapse model described in Chapter 3. However, take note that the noise in the auditory nerve fibre is independent of the spontaneous rate and the characteristic frequency of the nerve fibre and noise was present in all the nerve fibres.

Temporal pitch is encoded by the travelling wave using the entire length of the basilar membrane by causing nerve fibres to synchronise with the period of the stimulus signal. The population of nerve fibres stimulated by the basilar membrane displacement synchronise with the period corresponding to the pitch of the tone as well as multiples of this period, as can be seen by the ISI histogram for normal hearing. Because of the travelling wave, periodicity that reflects the pitch of the stimulus signal is not only available at the place of peak displacement, but along the length of the basilar membrane. This extent of nerve fibres synchronised to the stimulus period increases as the loudness level increases.

The limit of the range of temporal pitch perception is possibly around 1300 Hz (Cedolin and Delgutte, 2005). The model in the present work predicts the upper limit of temporal pitch prediction at around 1298 Hz. However, by inspection of the neural spiking patterns, phase locking may still be available at higher frequencies. However, the dominant change above this frequency consists of the first peak in the all-order ISI histogram, which relates to the period of the stimulus signal, being smaller than the multiples of the first order interval in the ISI histogram. This is likely due to the refractory period of the nerve fibres. As the stimulus period approaches the refractory period, the peak height of the first cluster in the ISI histogram decreases.

The temporal pitch detection algorithm developed by Cedolin and Delgutte (2005) uses an exponential weighting function with time constant of 3.6 ms, which corresponds to a minimum perceived pitch of 30 Hz. However, the weighting function strongly affects the upper limit of pitch perception and therefore, decreasing the time constant will increase the predicted upper limit of pitch perceived using temporal encoding. This is because the first cluster in the all-order ISI histogram has a greater weighting when predicting the pitch using the template matching algorithm. This does affect the minimum frequency detectable by the pitch prediction algorithm. Cedolin and Delgutte (2005) chose the 3.6 ms constant to match human data in the upper and lower limit of temporal encoding. Therefore, the normal hearing model developed here shows a good correlation to measured results in terms of the upper limit of temporal pitch prediction using the pitch prediction algorithm.

5.4.3 Vocoder CI

The vocoder CI processor relies on place coding (i.e. relies on the tonotopic arrangement of the cochlea). It implements bandpass filters that perform frequency decomposition of the acoustic signal, and then applies electrical stimuli at electrodes at specific sites in the cochlea. These stimuli reflect the energy in the corresponding filter bands. The vocoder speech processing approach implicitly assumes that place cues, accompanied by the modulation of filter output along with the envelope of the stimulus signal, contains enough information to encode acoustic stimuli for the cochlear implant user. However, the CI have limitations which limit the availability of place cues: the electrode array not being implanted along the entire length of the cochlea, the electrode array only having a resolution of 22 electrodes, nerve fibre survival that may be low in a particular area, and stimulation current levels needed to obtain enough loudness that result in current spread which deteriorates resolution. Nerve fibre survival will also affect rate pitch. The vocoder

cochlear implant speech processor relies on modulation at the filter output to convey rate information at a specific place along the length of the basilar membrane and not on a stimulation of a population of nerve fibres along the length of the basilar membrane, whereas the travelling wave relies on the population of nerve fibres along the length of the basilar membrane to encode rate pitch.

The vocoder CI implant can deliver temporal information to the auditory nerve fibres up to around 300 Hz, as shown before, by locking the neural spike trains onto the pulse rate of the electrical stimulation. However, once the pulse rate increases above the 300 Hz frequency limit, phase locking is lost rapidly and is likely to be limited by the range of temporal repetition rates that can be processed by CI users (McDermott and McKay, 1997).

The robust delivery of the pitch information is lost due to this limitation of temporal encoding and the strong reliance on place encoding (which has limited resolution) for vocoder speech processors.

5.4.4 Travelling wave processor

The TW CI processor mimics the TW along the length of the basilar membrane. This method may better reflect temporal characteristics of the acoustic input signal than the vocoder speech processor. Inspection of the nerve fibre spike train patterns suggest that nerve fibre spikes phase lock onto the emulated travelling wave stimulation, suggesting that temporal encoding of the travelling wave is delivered to the nerve fibres. However, inspection only would not validate this assumption, because the resolution of the travelling wave is lowered greatly, both temporally and spatially, and the entire length of the travelling wave is shortened due to the short electrode array (relative to the length of the cochlea).

All-order ISI histograms with a pitch detection algorithm was used to evaluate the temporal encoding of the TW CI processor. The results suggest that temporal cues may be made available to the auditory nerve fibres through TW stimulation. However, the ISI histograms differ from those of normal hearing. Histogram peaks locked onto the pitch period of the stimulus signal are sharper as the distribution of intervals around the peaks is small. This is because, as with VOC stimulation, nerve fibres phase lock to electrical stimuli with high precision (Javel, 1990). Temporal resolution of the electrically-presented

travelling wave is far lower than that present in normal hearing, which will result in less spread around the cluster peaks. The TW stimulation patterns essentially just provide more natural stimulation patterns, so that phase-locked neural spike trains reflect temporal information encoded in the travelling wave.

The sharper peaks around the pitch of the stimulus can be decreased by lowering the electrical stimulation amplitude, as the effects in the model do not cause electrical stimulation to force the nerve fibres to fire on each electrical pulse.

To increase the spread of ISI clusters, electrical stimulus amplitudes may be decreased. This will lead to less reliable phase-locking and entrainment will be smaller. A specific nerve fibre will have lower probability of firing on each electrical stimulus pulse, and will fire with more jitter, that will increase the spread around ISI peaks. The effects of lowering the electrical stimulation amplitude is unknown; it could lead to a lowering of the perceived loudness by the cochlear implant user. Smaller electrical amplitudes may not be a practical solution. The loudness thresholds in the modelled electrical stimulation, in this study, was taken from the cochlear implant users measured loudness map for vocoder speech processing. It may be found for the travelling wave processor the loudness map may be different in terms of maximum and minimum current amplitudes.

5.4.5 Alternative travelling wave processor

The alternative travelling wave processor also suffers from the sharper peaks in the ISI histogram (i.e. smaller spread around ISI histogram peaks than NH). However, when compared to the clustering of the travelling wave processor and the vocoder processor, more widely distributed clusters of nerve fibres can be seen around the acoustic signal period. This is a closer comparison to normal hearing nerve fibre predictions.

A predicted cutoff of temporal pitch is seen at 1098 Hz, meaning the alternative cochlear implant TW processor may not convey temporal pitch at higher frequencies. However, this also suggests that the alternative TW processor could convey temporal pitch above the 300 Hz limit measured in present cochlear implant users. One concern for travelling wave speech processors may be the spatial resolution that the electrode array provides, which may not clearly define the peak of the travelling wave. Spatial resolution is a specific concern in vocoder speech processors, but it seems that from the predicted nerve spiking patterns and ISI histograms that the two variants of the travelling wave processors may be less reliant on spatial high spatial resolution. However, while the alternative travelling

wave processor may be effective in encoding the rate information in pure tones and simple multi-tones, it is unclear, at this stage, whether it will adequately encode complex speech, which probably requires robust encoding of both place and temporal information.

5.4.6 Plausibility of implementation

The evaluation of the plausibility of a travelling wave cochlear implant processor from the present work, requires the evaluation of nerve fibre spike train predictions and the ability to implement the processor using present cochlear implant hardware.

The spike train patterns illustrating the nerve fibre predictions using the cochlear implant TW model allows for insight into and understanding of the delivery of spatial-temporal information to the user of a cochlear implant. The results suggest a more “natural” delivery of sound information by the alternative travelling wave speech processor for cochlear implant users; the greater delivery of temporal information to the nerve fibres. However, the loss of spatial area and resolution (short electrode and number of electrodes) may relate to spatial information loss. The stimulation along the entire length of the cochlea using the travelling wave may not show peaks in synchronisation of nerve fibres, causing the loss of spatial peak information used in current CI processing strategies.

The next steps in the development of a CI TW processor would be to conduct tests in an animal model to measure and to confirm the model predictions of nerve fibre spiking results. Subsequently, psychoacoustic tests with experimental TW processors must be conducted to test perception capabilities of CI users using CI TW processing stimulation profiles, with non-real-time processors at first (as processing is computationally expensive).

Cochlear implant TW processing may be hindered by nerve fibre survival, which would limit the capability of both place and temporal encoding of the stimulation waveform.

Real-time implementation of the TW CI processor does not appear viable for the present generation of cochlear implant speech processors. It is computationally too expensive to solve the hydrodynamic model to simulate the travelling wave for an incoming acoustic signal to allow real-time operation. However, the model of the travelling wave implemented in this work is a high resolution model of the travelling wave which may be downscaled to decrease the amount of time needed to predict basilar membrane deflection. Simpler linear cochlea models could be used and will decrease the computational

complexity of the solver. A linear model of the travelling wave, however, would cause the active properties and distortion components of basilar membrane displacement prediction to be lost. This, in turn, could cause the loss of possible benefits of encoding of loudness by a shift in peak displacement and broadening of the range basilar membrane displacement.

The basilar membrane travelling wave can be implemented using filters and hardware based platforms (Giguere and Woodland, 1993, Giguere and Woodland, 1994, Tateno, Nishikawa, Tsuchioka, Shintaku and Kawano, 2013). A hardware based model of the travelling wave could be a better implementation of a TW cochlear implant for real-time processing applications as parallel processing can be performed.

Wang et al. (2015) developed a CMOS analog cochlea which could be used with cochlear implants and had comparable responses to measured results in terms of basilar membrane gain and delay. Yang et al. (2015) developed an analog biomimetic cochlear implant processor filterbank architecture with across channels AGC (automatic gain control). The processor was able to predict the basilar membrane displacement in terms of gain but characteristics such as delay and phase are not compared to literature. These analog implementations of the cochlea processing need to be implemented with the rest of the required processing in cochlear implants, and may solve the potential issues of processing delays in a TW processor.

5.5 CONCLUSION

The main findings of this chapter may be summarised as follows.

- Temporal encoding can be detected in normal hearing at frequencies of 1000 Hz and higher (1298 Hz in the present model). An increase of the weight decay constant would enable an extension of the model's limits. However, this will affect pitch detection at the lower frequency bound of 30 Hz.
- Temporal encoding of pitch using the travelling wave processor can be detected by a model that extracts pitch from interspike intervals up to 1060 Hz. This pitch detection range is lower than that of the normal hearing model. The lower pitch predicted using the travelling alternative wave processor could be the effect of the specific nerve fibre stimulation model, but could also be the effects of the electrical stimulation as it is not a natural way of stimulation of the auditory nerves. However, this idea will have to be explored further.

- Travelling wave processing, using the alternative travelling wave processor, showed a better ability to robustly detect the pitch of the stimulus signal than both the vocoder CI processor and the normal TW processor. The latter needs to consider the space-time trade-off, while the alternative TW processor does not have this limitation. In summary, from model predictions, it is proposed that the alternative travelling wave processor could be a viable solution for the delivery of temporal information to cochlear implants. However, validation of the nerve fibres spiking patterns in an animal model is necessary.

CHAPTER 6 DISCUSSION

6.1 CHAPTER OBJECTIVES

This chapter evaluates the models that were implemented in the present study: the normal hearing model, the electrical stimulation model, and the temporal pitch detection model. The delivery of temporal information from acoustic stimuli using electrical stimulation is important to further developing CI implant processing strategies.

The model of normal hearing described in Chapter 3 will be discussed, as well as the performance of the model against measured responses in literature. Chapter 4 dealt with two travelling wave processing strategies and the possible improvements of temporal encoding in nerve fibres using these strategies in cochlear implants. Methods of encoding temporal information in cochlear implants will be discussed and improvements to the present travelling wave model will be proposed.

Chapter 5 considered the validation of encoding of temporal pitch in the auditory nerve fibres of the model and an indication toward the extent to which the travelling wave processors may be expected to improve cochlear implant pitch perception. The possible improvements seen by Chapter 5 are discussed in this chapter to identify whether the proposed travelling wave processors are viable processors.

6.2 NORMAL HEARING MODEL EVALUATION

The normal hearing models were used to generate the nerve fibre response to acoustic stimuli. Model results are then compared to literature to evaluate the worth of the normal hearing model.

6.2.1 Justification of the normal hearing model.

Normal hearing is modelled using signal processing which consists of filter banks and feedback loops, which model the characteristics of the travelling wave along the length of the basilar membrane (Carney, 1993, Zilany and Bruce, 2006). Using a hydrodynamic model of the travelling wave, which models the mechanical response, allows for the validation of the encoding performed by the travelling wave. This gives insight into the encoding present in the travelling wave, which will facilitate a greater understanding of how the travelling wave characteristics affects the nerve fibres and the encoding of sound information.

To understand the electrical stimulus, needed to be presented to auditory nerve fibres during cochlear implant stimulation, a normal hearing model along with the electrical stimulation model may be used to detangle the pre-processing of the travelling wave. Therefore, an understanding of the nerve fibre results observed in normal hearing may lead to the improvement of cochlear implant processing strategies.

6.2.2 Evaluation of the travelling wave model

The travelling wave is the first stage in the processing of the acoustic signal before it reaches the auditory nerve fibres. Detangling the pre-processing done by the cochlea mechanics before the encoded signal reaches the nerve fibres could lead to a better understanding of the spatio-temporal information. The first step in detangling the pre-processing, carried out by the cochlea mechanics, is to ensure the travelling wave model strongly resembles the measured results in literature.

The model of the basilar membrane showed good correspondence to measured results in literature. The model was able to predict the frequency-place peak excitation, the basilar membrane delay, the gain of the basilar membrane (active amplification of the basilar membrane during low loudness level stimuli), the shift in basilar membrane peak due to loudness level change, and the broadening of the travelling wave during an increase in loudness level. The importance of these characteristics needs to be detangled by observing how they would affect the perception of an acoustic signal. A good model allows for the prediction of results without invasive measurement using animal models. This allows for prediction of results (basilar membrane deflection/nerve fibres results) without conducting physically invasive measurements.

The frequency map created by the travelling wave along the length of the cochlea is one way to convey place pitch to the auditory nerve fibres by using place coding. However, a larger area of the basilar membrane is deflected (not only at a single place of excitation), the width of the area deflected is increased, and the maximum point of stimulation shifts toward the base (for pure tone stimulation) with an increase in sound pressure level (i.e. the displacement of the basilar membrane is dependent not only on the frequency components of the stimulus signal, but also the energy contained in the stimulus signal (Chatterjee and Zwislocki, 1997, Duifhuis, 2012)). The shift in place peak excitation is constantly changing due to sound pressure level. This suggests that pitch perception, during place pitch encoding, increases towards higher frequencies as loudness level is increased.

However, in this case classical place pitch theory cannot hold (Chatterjee and Zwislocki, 1997). The encoding of the loudness level in the broadening and shift of the peak of the basilar membrane is a possible explanation of the change in width and shift in peak seen during the travelling wave (Chatterjee and Zwislocki, 1998).

Single nerve fibres in cochlear implant users have a smaller dynamic range than that of normal hearing listeners in perceived loudness (Ruggero et al., 1997, Hartmann, 2005). The basilar membrane acts as an amplifier through outer hair cell electro-motility. This causes the stimulation of nerve fibres at a lower sound pressure level than when the basilar membrane acts as a purely non-active mechanical device (Kemp, 2002). The basilar membrane acting as an amplifier allows for the dynamic range to be increased. The broadening of the travelling wave and shift of peak excitation position may encode information to nerve fibres about the energy of the signal which gives cues to the sound pressure level heard (Chatterjee and Zwislocki, 1998).

The travelling wave model is able to predict both the broadening of the basilar membrane excitation region and the basilar membrane's ability to act as an amplifier, which may allow for greater understanding of the encoding of loudness and pitch. As such, the travelling wave model can be used as a tool to understand the information encoded in the nerve fibres during normal hearing.

6.2.3 Improvements to the travelling wave model

There are at least two aspects in which the present models are inaccurate and may be improved. First, the present middle ear model is a simplistic model and could be greatly improved upon over a greater frequency range to cover the frequency range of the human ear (Duifhuis, 2012). However, an improvement of the middle ear model may not be important in the context of application in cochlear implant processors in the sense that the effect on the travelling wave model as applied to cochlea electrical stimulation would be small. Encoding is severely restricted by a cochlear implant's hardware, specifically in the stimulation rate and number of electrodes available. A more complex middle ear model, e.g. O'Connor and Puria (2008), includes the entire frequency range heard by normal hearing users. However, the model used in the present work (Chapter 3), act as a bandpass filter with a centre frequency of 2 kHz which has a low cut-off frequency at 0.7 kHz and a high cut-off at 5.7 kHz.

Second, the travelling wave delay in the present travelling wave model does not correlate closely with measured travelling wave delays along the length of the basilar membrane. The delay of travelling wave displacement seems to be offset from the results measured in literature. This may be due to the fact that the model uses the Greenwood map to describe the stiffness of the basilar membrane, as seen in Chapter 3 where it is explained, which strongly influences the delay of the model along the length of the basilar membrane. A better description of the stiffness of the basilar membrane should be developed and added to the model to accurately describe the basilar membrane stiffness. However, the delay of the travelling wave shows the same trend of the measured results. In this work, the focus is on the characteristics of the travelling wave. Therefore, the effects of the travelling wave delay can still be observed in the normal hearing model.

The travelling wave model is used in the travelling wave speech processor to gain the basilar membrane displacement patterns, in order to gain the electrical stimulation pulses. Therefore, it is important that the travelling wave is in good correlation with the normal hearing measured result from literature.

6.2.4 Evaluation of the nerve fibre model

The normal hearing model showed that the nerve fibre predictions could predict the following: by inspection, it can be seen that the nerve fibres lock on to intervals of the stimulus signal, the synchronisation of nerve fibres to the pitch period along the length of the basilar membrane, and the tuning curve for nerve fibres. The latter lead to a significant finding: the modelled nerve fibres have no inherent frequency sensitivity, and are only affected by the encoding present in the travelling wave. Yet, tuning curves are observed as a result of the encoding in the travelling wave.

Nerve fibre spike train predictions generated using the normal hearing model show that nerve fibres synchronise with the travelling wave stimulus. The nerve fibres synchronise with the travelling wave from base to apex, and as the width of the deflected area increases on the BM, as a result of increased acoustic intensity, the width of the area of synchronisation increases, resulting in an increase in loudness. Low sound pressure levels have a place of peak neural synchronisation which correlates well with the average localised synchronised rate seen in literature (Young and Sachs, 1979)), and corresponds to the tonotopical frequency arrangement along the length of the basilar membrane.

The tuning curves correspond to the displacement of the basilar membrane as follows: the displacement of the basilar membrane is focussed at a particular place during low sound pressure levels, and broadens (towards the base of the cochlea) with an increase in sound pressure level (Chatterjee and Zwislocki, 1997). From the nerve fibre tuning diagrams, which correlate to the nerve fibre tuning curves from Temchin et al. (2008), the broadening characteristic indicates that lower stimulus frequencies (lower than the characteristic frequency of the nerve fibre) can stimulate the nerve fibres, irrelevant of the place pitch of the stimulus frequency, and cause synchronisation of the nerve to the stimulus frequency.

The results of the nerve fibre model were comparable to those in literature without individual nerve fibre frequency selectivity coded into the nerve fibre. This showed that using the travelling wave as input into the nerve fibre allows nerve fibres to have cochlea filtering and retained place pitch sensitivity at low loudness levels.

6.3 ELECTRICAL STIMULATION MODEL EVALUATION

Processing strategies needed to be tested and evaluated to verify the possible improvement that a TW-based processing strategy could allow for cochlear implant users. Using a modelling approach for the electrically stimulated ear could predict the possible benefits of processing strategies and create an understanding of the potential improvements.

6.3.1 Vocoder speech processor

The proposed electrical travelling wave stimulation processor differs greatly from the generally used vocoder speech processor. The vocoder speech processor relies on filter bands to elicit place perception of sound, whereas the travelling wave relies on both place and temporal aspects of perception along the length of the basilar membrane. The vocoder speech processor conveys pitch with sinusoidally amplitude-modulated pulse trains (Loizou, 2006), which has been shown in literature to be limited as nerve fibres lose synchronisation to the period of the stimulus period above a frequency of 300 Hz (McDermott and McKay, 1997). The results from this study (those in chapter 5) correspond to data from literature, showing that the period of the stimulus is lost in the ISI histograms as the pulse rate is increased past 300 pps.

The limitation of phase locking may be due to the nerve fibres' electrical refractory period, where the absolute refractory period is 0.7 ms (frequency of 1429 Hz). However, the relative refractory period is around 1.32 ms (but when added to the absolute refractory period, corresponds to approximately 455 Hz). The question may be asked: why do nerve

fibres not fire at a rate closer to the absolute refractory period if stimulation current is increased? It may be that the comfortable loudness level of stimuli, obtained from the cochlear implants user map, may not be of a high enough intensity to stimulate the nerve fibre above threshold during the relative refractory period as the nerve fibre is recovering from previous stimulus. Increasing the pulse rate to high pulse rates such as 14400 pps (the highest pulse rate in processor modelled in the present study) the nerve fibre may fire at a higher rate as seen in the travelling wave speech processor predictions due to the higher stimulation pulse rate overcoming the relative refractory period threshold. This theory will have to be tested in animal models and computation models to see if higher pulse rate will overcome the threshold caused by the relative refractory period.

Vocoder speech processors rely heavily on place-pitch encoding. This does not take advantage of the temporal aspects of encoding speech. Temporal encoding can encode information about the pitch period of the stimulus signal in multiple nerve fibres across the length of the cochlea, taking advantage of spatial area of activation without the need for place cues. Temporal encoding could use the volley principle, in which nerve fibres encode pitch using multiple nerve fibres to lock on to a high frequency stimulus (Wever and Bray, 1937). In order for pitch to be encoded temporally in a better way and to improve pitch perception, the processing strategy for cochlear implants needs to be improved and reconsidered.

6.3.2 Modelling of the electrical travelling wave speech processor

The electrical travelling wave model used two methods: the travelling wave processor model, which used a space-time trade-off to stimulate the travelling wave across the length of the basilar membrane, and the alternative travelling wave processor, which follows the maximum peak of the travelling wave during an instant of time.

The travelling wave processor model was able to encode temporal pitch into auditory nerve fibre spike train patterns and create predicted nerve fibre responses that are closer than vocoder stimulation to mimicking those of the normal hearing model. However, the application of this method requires an understanding of both the spatial resolution required, and the temporal resolution required, in order to optimally encode speech information. This varies with the frequency of the signal: higher frequency stimuli requires a higher temporal resolution, while a lower frequency signal requires a lower pulse rate, which allows one to take advantage of a higher spatial resolution in encoding of the acoustic signal. The higher

frequencies suffer from poorer spatial encoding over the length of the cochlea, due to a smaller number of active electrodes because of increasing the maximum pulse rate. The latter is required to cause the phase locking of nerve fibres to the period of the stimulus.

The selection of active electrodes is important, as an electrode that is active but outside the range of the wavefront of the travelling wave displacement, in a sense, wastes a sample of the travelling wave (which could be used as sample at a different electrode to convey information to the auditory nerve fibres). A sample is wasted, in that it does not add information about the (mimicked) travelling wave along the length of the basilar membrane. Deciding which electrode is active for a simple stimulus, such as a pure tone, could be decided by pre-determined parameters. However, a complex signal may not be able to be determined with such ease, as the signal is decomposed along the entire length of the cochlea during multi- tone stimulation. It may require sophisticated processing to optimally select electrode number and placement for a particular input signal.

The travelling wave processor is able to create nerve fibre response which mimics encoding seen during normal hearing (Chapter 3). However, the ISI histograms are sharper (i.e. the peaks cluster closely around the period of the stimulus signal). This clustering is sharper than that observed in normal hearing and the resulting effect on perception, with a cochlear implant, is unknown.

The travelling wave speech processor is far more complex to implement than the vocoder speech processor and is limited by the space-time trade-off, which may not allow for implementation in real-time CI processing.

6.3.3 Modelling of the electrical alternative travelling wave speech processor

The alternative travelling wave speech processor (Chapter 4) was developed to see the effects of following the maximum energy of the signal during a time sample. This method of travelling wave simulation in electrical stimulation patterns affects the length of the travelling wave during multiple stimulation periods (Chapter 4, Figure 4.7). The alternative travelling wave processor follows the peak displacement of the travelling wave, and while the travelling wave caused by the first period of the stimulus signal is still being sampled on to the electrode array, the second period begins. The travelling wave caused by the second period has a smaller displacement than the travelling wave caused by the first period of the stimulus signal, and the second travelling wave only gets sampled once the displacement exceeds that of the first travelling wave. The effects include a shortening of

the travelling wave, low spatial resolution, and fewer nerve fibres along the length of the cochlea phase lock to the travelling wave signal as that seen during normal hearing. Fewer activated nerve fibres could affect the robustness of the encoding and may hinder processing at further stages of the auditory cortex, such as the encoding provided by the octopus cells of the CN (McGinley et al., 2012).

The octopus cell possibly compensates for the travelling wave delay by having dendrites which extend across the length of the basilar membrane, with longer dendrites at placement closer to the base of the cochlea and shorter dendrites closer to the octopus cell (closer to the apex) of the cochlea. The octopus cell could possibly compensate for the travelling wave delay by optimally firing with travelling wave delays across the length of the basilar membrane. The octopus cells' dendrites only extend 1/5 to 2/3 of the length of the basilar membrane. This means that the alternative travelling wave speech processor, with shortened travelling waves, may still be able to activate the octopus cell as some of the octopus cells extend in the range of the travelling wave cochlear implant stimulation.

The travelling wave is able to cause nerve fibres to phase lock on to the stimulus period of the signal, allowing for the encoding of the temporal information for processing by further auditory processing stages. The alternative travelling wave processor may allow for place encoding (the point of peak excitation during the travelling wave stimulation is followed) and temporal encoding (phase locking of nerve fibres along the length of the electrode array), which may possibly account for the multiple encoding mechanics delivered to the auditory cortex (Greenberg, 1997).

The place encoding of the proposed alternative travelling wave processor may be a concern as the electrode spatial resolution is low and there is a potential mismatch of the place cues due to the electrode configuration, size, and number of electrodes. This may cause a mismatch in the temporally encoded information and the place encoded information presented to central auditory processing mechanism for decoding (Oxenham et al., 2004). Oxenham et al. (2004) showed that transposed tones, expected to result in nerve fibre spiking patterns, do not correlate to the place pitch of the signal. When transposed tones were used in pitch discrimination tests, the listeners performed poorer than listeners conducting the test used pure tone signals. This could mean that expected performance of the alternative travelling wave processor may not be as good as expected due to the place-pitch mismatch caused by the electrode array. However, pitch was still perceived by listeners in the experiments from Oxenham et al. (2004) although performance was

degraded, which means that the alternative travelling wave speech processor will still deliver temporal pitch above 300 Hz for cochlear implant users where this is not the case for vocoder speech processors.

The predicted benefits of the alternative TW processor seen in this work is the phase locking of nerve fibres to the stimulus signal for frequencies over a large spatial range of the basilar membrane. The ISI histograms show intervals at the period of the stimulation signal. These predictions could indicate that the travelling wave speech processor may allow for strong temporal encoding across the length of the basilar membrane.

The encoding of temporal pitch is accurately decoded from 100 Hz up to 1060 Hz for the alternative TW model using the pitch detection algorithm; also, stimulation signals with a missing fundamental pitch could be detected for signals with a (missing) fundamental frequency of 200 Hz, 500 Hz, and 1000 Hz. These predictions suggest that the alternative travelling wave processor may have the ability to temporally encode complex signals in spike train patterns (i.e. presented in a format that should enable decoding by the central auditory nervous system).

The nerve fibre ISI histogram predictions are sharper (higher peaks and closer clustering around the period of the fundamental frequency) than those of the normal hearing predictions. The sharpness of the clustering on the ISI histograms increases with frequency increases, which show similarities with the ISI histogram observed for the vocoder speech processors. The effects of sharper ISI histograms could lead to poorer temporal encoding and poorer temporal pitch perception by the cochlear implant user, this will still have to be investigated, as the effects are unknown. However, the pitch prediction algorithm still showed that the correct temporal pitch encoding was present in the sharp ISI histograms of the high frequency stimulation signals. The pitch prediction algorithm showed that the alternative travelling wave processor has the ability to present the required information to the auditory nerve fibre for pitch perception above the 300 Hz limit seen during vocoder speech processing.

6.3.4 Relevance

To evaluate the plausibility of a travelling wave cochlear implant processor from the present work; the nerve fibre spike train patterns prediction and the possibility of implementing the processor using present cochlear implant hardware needs to be evaluated.

The spike train patterns illustrating the nerve fibre predictions using the cochlear implant travelling wave (CI TW) model, allows for insight and understanding of delivery of spatial-temporal information to the user of a cochlear implant. The results for the travelling wave speech processors suggest potentially more natural spike train patterns for cochlear implant users by more accurate delivery of temporal information to the auditory nerve fibres (i.e. spike train patterns that are closer to natural spike train patterns for the same acoustic stimulus). However, the CI travelling wave processor does have limitations such as the loss of spatial area (due to short electrode array which cannot be implanted along the length of the cochlea) and resolution (limited number of electrodes on a cochlear implant array) which may relate to spatial information loss.

The next steps in the development of a cochlear implant travelling wave processor would be: 1) to conduct tests in an animal model to measure and confirm the model predictions of nerve fibre spiking diagrams and 2) to use psychoacoustic tests with experimental processors to test perception capabilities of CI users using the CI TW processing stimulation profiles, without the need for real-time stimulation as the stimulation profiles may be pre-processed before the time of the psychoacoustic experiment.

The limitation of the CI TW processor is its ability to be hindered by nerve fibre survival, which would limit the capability of both place and temporal encoding of the stimulation waveform. The cells in the cochlea nucleus, acting as incident detectors of temporal information (Greenberg et al., 1998, McGinley et al., 2012), could also be affected by cell survival and damage. This could greatly affect the results during psychoacoustic experiments.

6.3.5 Implementation

The real-time implementation of the cochlear implant travelling wave processor is not viable for implementation in present cochlear implants. The model of the travelling wave introduces delays in creating the electrode stimulation patterns. This is due to the complex nature of the travelling wave model which has large computational load. The travelling wave speech processor would not be viable in a normal acoustic environment. However, the model of the travelling wave implemented in this work is a high resolution model of the travelling wave which can be downscaled to decrease processing times of basilar membrane deflection predictions. Other travelling wave model methods, such as linear

travelling wave models, could be used and will decrease the computational complexity of the solver.

The basilar membrane filtering can be predicted using filters and hardware base platforms (Giguere and Woodland, 1993, Giguere and Woodland, 1994, Tateno et al., 2013). The hardware based model of the travelling wave could be a better implementation of a TW cochlear implant where real-time operation is required.

Hardware implementations of the travelling wave are possible, and will not have long processing times, as the travelling wave displacement predictions will be done using analog hardware predictions, as was done in Yang et al. (2015) and Wang et al. (2015). However, understanding and validating the results are important before implementation can take place. This will require validation of the model results by psychoacoustic experiments (non-real time i.e. with pre-processed stimulus patterns testing pitch perception) and animal models, in which nerve fibre firing patterns in the latter can be compared to the predictions found in this study.

Validation of the normal hearing model is done by showing that the model can accurately predict the nerve fibres spiking patterns. This allows for the use of the model to predict nerve fibre spiking patterns which may allow for a simpler processing algorithm development for cochlear implant speech processors. The model will be able to easily validate speech processing algorithms which may deliver pitch and speech information to the nerve fibres. The presented travelling wave processor model provides a starting point to develop a practical TW speech processing algorithm for cochlear implant users which convey pitch information to cochlear implant users.

6.4 TEMPORAL PITCH PERCEPTION

This section considers the delivery of temporal pitch perception to the auditory nerve fibres during the normal hearing listener (vocoder CI stimulation) and CI TW stimulation. The value of using a decoder to extract pitch is that it enables validation of the predicted nerve fibre spike train patterns, which may contain useful temporal pitch information and allow for a repeatable measure of the performance of the models and processing strategies.

6.4.1 Normal hearing

Some of the cues which can be used to detect pitch by central auditory processing mechanisms are the temporal periodicities, average firing rates, and their distribution

across the auditory-nerve fibre array (Shamma, 1985a). Average firing rates saturate under higher stimulus intensities (Cedolin and Delgutte, 2010). Therefore, different encoding mechanisms need to be explored to describe the encoding of pitch in the auditory system. Cedolin and Delgutte (2010) found that the spatiotemporal representation (time difference between spikes, with components along the length of the cochlea) of pitch is more consistent with pitch psychoacoustic trends and is more robust at higher stimulus intensities than rate-place representation (i.e., a place specific rate of firing of a nerve fibre during a time interval).

The synchronisation of individual nerve fibres is limited to a firing rate of 150-300 spikes/second (Javel, 1990). A population of nerve fibres, however, make use of the volley principle: nerve fibres lock on to sections of the stimulus signal, and by using the cumulative sum of the multiple nerve fibre responses, the pitch period of the acoustic signal is encoded (Wever and Bray, 1937, Rattay, 1990). Therefore, higher stimulus frequency is encoded using the volley principle in conjunction with a population of nerve fibres. The travelling wave distribution of stimulation along the length of the basilar membrane activating many fibres allows for the volley principle to take place, and is not localised to an individual site.

The limitation of temporal pitch perception is possibly around 1300Hz (Cariani and Delgutte, 1996b, Cedolin and Delgutte, 2005). In this work, a limitation of 1298 Hz for the normal hearing model predictions is evident (Chapter 6). However, by inspection, the phase locking to the stimulus signal may still be present, but the first peak in the all-order ISI histogram is far smaller than that which was observed during lower frequency stimulation. This is most likely due to the refractory period of the nerve fibres. The temporal pitch detection algorithm developed by Cedolin and Delgutte (2005) uses an exponential weighting function of 3.6 ms which correlates to minimum pitch perception of 30 Hz. However, the weight function strongly affects the upper limit of pitch perception and decreasing the weighting will increase the upper limit frequency of pitch perceived using temporal encoding.

Predicted temporal pitch perception with a TW processor is robust. The phase locking of nerve fibres is not localised to a point along the length of the basilar membrane (limiting the deleterious effect of damaged nerve fibres during hearing loss) because the travelling wave stimulates along the length of the cochlea.

This work only focuses on the temporal cues which could be delivered. However, a combination of both temporal and place cues is most likely the method in which speech information is encoded by the travelling wave (Oxenham et al., 2004, Cedolin and Delgutte, 2010, Oxenham, 2013). The combination of the temporal and place cues is evident in the auditory system and lateral inhibitory networks (LIN) serve to highlight both temporal and spatial cues in the nerve fibres (Shamma, 1985b). A LIN may be able to highlight the phase locking of nerve fibres further, therefore increasing the upper limit of the frequency detectability of temporal encoding. Using a LIN to decode pitch in the normal hearing model may help to decode the pitch encoding done by the travelling wave at higher frequencies where both place and temporal encoding play a role in the perception of pitch.

6.4.2 Vocoder CI

The vocoder CI relies on filter bands along the length of the basilar membrane, used for the frequency decomposition of the signal. The vocoder processor relies on place encoding. Vocoder speech processing essentially assumes that place cues give enough information for the cochlear implant user to perceive pitch. However, the CI has limitations that hamper the value of place cues: the electrode array is typically not implanted along the entire length of the cochlea, the electrode array only has a physical resolution of 22 electrodes, nerve fibre survival patterns probably vary greatly among CI users, and relatively high stimulation currents are often required to gain the necessary loudness level. The latter may lead to increased current spread which will again deteriorate spatial resolution.

The vocoder CI implant can deliver temporal information to the auditory nerve fibres up to 300 Hz – 400 Hz, by locking the nerve fibres onto the pulse of the electrode stimulation (McDermott and McKay, 1997, Shepherd and Javel, 1997). However, once the pulse rate increases above the limit at around 300 pps, the phase locking is progressively lost. This may partially explain the inability of many cochlear implant users to perceive pitch for higher rates of stimulation.

This suggests that vocoder-type (or ACE-like) processing strategies do not capture periodicity information of the acoustic signal or do not elicit the neural phase locking that is related to the characteristics of the acoustic signal for complex signal.

6.4.3 TW CI

The cochlear implant travelling wave processor mimics the TW along the length of the basilar membrane; this method may produce spike train patterns that are far closer to the temporal encoding found in normal hearing. Inspection of the nerve fibre spike train patterns suggest that nerve fibre spikes may phase lock on to the travelling wave stimulation, which would indicate that temporal encoding of the travelling wave is delivered to the nerve fibres. However, inspection alone would not justify this assumption, because the resolution of the travelling wave is lowered greatly and the entire length of the travelling wave shortened.

An all-order ISI histogram with a temporal detection algorithm was used to evaluate the temporal encoding of the TW CI processor. The results of the temporal pitch detection suggest that temporal cues may be made available to auditory nerve fibres. The ISI histograms show differences from normal hearing predictions; the ISI histograms peaks, clustered around the pitch (period) of the stimulus signal, are sharper, not as widely distributed around the period of the stimulus signal, and no random firing occurs during the travelling wave cochlear implant speech processor.

The sharper peaks around the pitch period of the stimulus can be decreased by lowering the electrical stimulation amplitude. This would, however, result in lower perceived loudness. The delivery of temporal pitch is limited to 1060 Hz in the model, meaning the travelling wave cochlear implant processor may not improve temporal pitch perception at higher frequencies. However, the upper limit for temporal information is around 1300 Hz and other encoding mechanics such as rate-place encoding may be dominant at frequencies higher than 1300 Hz (Cedolin and Delgutte, 2005, Cedolin and Delgutte, 2010). Therefore, this is not of concern.

During the travelling wave speech processor, phase locking greater than 300 Hz is observed. The pitch detection algorithm does not account for the volley principle, so the phase locking of nerve fibres to pitch greater than the 300 Hz limit is questionable, as it is not seen during the vocoder speech processor which uses the same electrical stimulation model. However, this is possible as the absolute refractory rate of the electrical model nerve fibre is 0.7 ms (which relates to a frequency of 1428.57 Hz), so the electrical nerve fibre model has the ability to fire and synchronise with a pulse rate far above the 300 Hz limit but this is not observed in the vocoder speech processor simulations. The possible

answer to this stimulation is that stimulation current is not able to overcome the threshold due to relative refractory period.

One way in which the travelling wave cochlear implant speech processor could overcome the relative refractory period is that the pulse rate is higher for the travelling wave speech processor (14400 pps can occur on one electrode in the processor modelled). The higher pulse rate could overcome the higher threshold during the relative refractory period causing nerve fibres to lock on to a higher pitch (shorter period) of the stimulus. This higher pulse rate does not affect the nerve fibres' synchronisation, as the period of the acoustic stimulus signal is not encoded in the pulse rate but in the travelling wave, thereby causing nerve fibres to phase lock to the period of the stimulus and not to the higher possible period of the pulse rate.

6.4.4 Possible issues in temporal pitch prediction

The temporal pitch prediction algorithm does not take into account the volley principle (Rattay, 1990), which can account for predicted pitch perception for frequencies higher than the limit of around 1300 Hz observed here, and by Cariani and Delgutte (1996a). The population of nerve fibres can lock on to the period of the stimulus signal using multiples firings and later be decoded in the central auditory nervous system.

Cedolin and Delgutte (2010) argue that the temporal pitch prediction algorithm has trouble explaining the poor pitch perception in cochlear implants, as phase locking with nerve fibres is excellent with electrical stimulation and much higher than the 400 pps tested in psychoacoustic experiments (Dynes and Delgutte, 1992, Shepherd and Javel, 1997). However, the phase locking is only for newly deafened cats. After a period of time, deafened cats only show phase locking up to 400 Hz (Shepherd and Javel, 1997). Shepherd and Javel (1997) hypothesize that lower phase locking upper limit in deafened cats could be due to the increase in threshold, which may be due to the demyelination of the nerve fibres.

The higher threshold required could lend support to the higher stimulation pulse rate required to overcome the relative refractory period threshold of the auditory nerve fibre due to the demyelination of the nerve fibre, thereby causing phase locking of the nerve fibre higher pulse rates than the 300 pps seen in literature.

6.4.5 Discovering information contained in the travelling wave using decoding algorithms

This study only considered the potential of extracting information using the temporal properties of nerve fibres, which may contain pitch information. However, place encoding, rate encoding, phase encoding, and temporal encoding may all play important roles in the encoding of sound using the travelling wave as the mechanism for encoding.

Using the decoding algorithms, we are able to see what potential decoding is done by the central auditory nervous system. The decoding algorithms need to be extended to include both place encoding and temporal encoding algorithms in order to fully obtain the results seen during psychoacoustic experiments. It may be necessary to understand in far more detail how decoding works more internally in the auditory pathway. If understood, this will allow for the development of algorithms which mimic the decoding of the brain and allow for cochlear implant processors to more correctly present information to the auditory nerve fibres for decoding.

6.5 SUMMARY

This chapter expanded on the discussion sections from the chapters of 3 to 5. Here the implications of the findings and how they relate to literature is discussed.

The normal hearing model is discussed as a valid tool to be used in prediction of nerve fibres' firing patterns for research. The normal hearing model can accurately predict many of the travelling and nerve fibres data measured in literature.

The second part of the study, which includes the travelling wave processors, is discussed and the interpretation of the results is examined. The travelling wave processors may have the ability to allow for temporal encoding using cochlear implants, as they are not reliant on rate encoding. However, animal model and psychoacoustic experiments still need to be conducted. The effects of the findings of the research questions are considered for the next chapter, which is to be discussed.

CHAPTER 7 CONCLUSION

7.1 CHAPTER OBJECTIVES

This chapter summarises the findings of the research study, the influence on the approach to future work, and the discussion of the research questions posed in Chapter 1. The study results are analysed in order to discuss the influence of this work on the research field of cochlea modelling and cochlear implant speech processing.

7.2 FINDINGS

The findings from this research, namely in regards to the travelling wave, nerve fibre model and the electrical model effects, will be discussed briefly below.

7.2.1 Finding 1: Intrinsic filtering in the travelling wave

The travelling wave is the first stage in processing of the acoustic signal before it reaches the auditory nerve fibres. Detangling the pre-processing done by the cochlea mechanics before an encoded signal reaches the nerve fibres could lead to a better understanding of the spatiotemporal information. The travelling wave has important characteristics including: the travelling wave delay, the increase in basilar membrane displacement from the base of the cochlea to the point of maximum excitation, the shift of maximum excitation point towards the base of the cochlea with an increase in sound pressure level and the broadening of the excitation area along the length of basilar membrane with the increase in sound pressure level.

There is a strong correlation between the tuning curves of the nerve fibres and the displacement of the basilar membrane. For example, the displacement of the basilar membrane is extremely sharp during low sound pressure levels and broadens (towards the base of the cochlea) during an increase in sound pressure level. We can see from the nerve fibre tuning diagrams that the broadening of the travelling wave characteristic indicates that lower stimulus frequencies (the characteristic frequency of the nerve fibre) can stimulate the nerve fibres at the base of the cochlea and cause synchronisation of the nerve fibres in that region. This indicates that nerve pitch may not only be place specific.

The threshold tuning curve of the nerve fibres is determined by the travelling wave response. This study showed that with nerve fibres having no frequency selectivity, tuning curve data could be predicted from literature.

The travelling wave model can be used to predict the filtering characteristics seen during basilar membrane displacement measurements. This means that the travelling wave models can be used in hearing models.

The travelling wave displacement characteristics have a strong influence on the response of the nerve fibres due to stimulation. Mimicking this process would allow for the delivery of robust speech cues to the auditory nerve fibres during electrical stimulation.

7.2.2 Finding 2: Nerve fibres can synchronise to the acoustic signal

The nerve fibre results generated, using the normal hearing model, show that nerve fibres synchronise with the travelling wave stimulus. The nerve fibres synchronise with the travelling wave from base to apex, increasing the area of synchronisation with an increase in stimulation loudness level (low sound pressure level has a “place” of synchronisation) and tonotopical frequency arrangement along the length of the basilar membrane.

The prediction of pitch temporal encoding of nerve fibres in this study correlated to results from literature. The predicted result of 1298 Hz correlated with the result from Cedolin and Delgutte (2005) of 1300 Hz.

7.2.3 Finding 3: Normal hearing model

Auditory modelling showed that a model of the travelling wave can be used to model the processing of the basilar membrane. The auditory model was able to accurately detect the nerve fibre responses during pure tone stimulation.

The results predicted by the auditory model included: nerve fibre spiking patterns, synchronisation rate, ISI histograms (which showed phase locking of nerve fibres on to the pure tone stimulus period), and the nerve fibre tuning curves. The nerve fibre tuning curve is found without the nerve fibres having frequency selectivity, therefore the tuning curves are only affected by the travelling wave displacement.

7.2.4 Finding 4: The alternative travelling wave speech processor performs better than the travelling speech processor

The electrical travelling wave stimulation model differs greatly from the vocoder speech processor. The vocoder speech processor relies on filter bands to stimulate place perception of sound, whereas the travelling wave relies on both place and temporal aspects of perception along the length of the basilar membrane.

The results of the electrical model for the alternative travelling wave speech processor show nerve fibres firing patterns, which are closer to mimicking the normal hearing firing patterns, the ISI histograms, which closely mimic the ISI histograms for low frequency stimulation, and the synchronisation of nerve fibres to higher frequencies than those of the vocoder speech processor.

The results here showed the vocoder speech processor has accurate pitch prediction up to 300 Hz; this correlated to the upper limit in literature of 300 – 400 Hz phase locking to the pulse rate of the cochlear implant (McDermott and McKay, 1997). The alternative travelling wave processor was able to increase that prediction to 1060 Hz using temporal encoding of the travelling wave, and accurately encoded missing fundamental pitch in cochlear implant using temporal encoding for multiple frequencies.

The travelling wave cochlear implant processor may be a viable solution for encoding and the delivery of temporal information to cochlear implant users. However, the present research implementation has limitations, which includes non-real-time application of the travelling wave cochlear implant processor. It also should be noted that results have not been validated using psychoacoustics and animal models.

7.3 DISCUSSION OF RESEARCH QUESTIONS

Basilar membrane modelling for the use in auditory nerve fibre firing prediction uses measured basilar membrane displacement to predict the basilar membrane filters. These filters are then used as the input for nerve fibre models. Does a more realistic model of the cochlea filtering, by using the travelling wave as a signal processing technique, allow for the accurate prediction of nerve fibre spiking patterns and cochlea filters?

The auditory model developed in this work used a travelling wave model to describe the displacement of the basilar membrane (Chapter 3). The use of a realistic travelling wave model, which implements the travelling wave characteristics, was used to predict nerve fibre spiking diagrams. These results correlated to measured data in literature such as, ISI histograms, nerve fibre synchronisation, nerve fibre spiking diagrams, and nerve fibre tuning curve results (Chapter 3.3).

The prediction of the nerve fibre spiking patterns was important as it shows that realistic implementation of the travelling wave has the ability to reflect place encoding, which is important in the cochlear implant pitch recognition. However, the travelling wave still has

a strong ability to encode temporal information along the length of the basilar membrane, while containing the place encoding mechanics.

Temporal synchronisation of nerve fibres across the entire stimulation of the travelling wave can be observed. The travelling wave does not only contain pitch information at the peak of excitation, but along the entire length of the basilar membrane, allowing for pitch extraction from multiple nerve fibres which are not place specific, making the encoding of pitch more robust to nerve fibre damage. The travelling wave, as a processing mechanism for an auditory model, can accurately encode temporal pitch information in the auditory nerve fibres.

The use of the travelling wave hydrodynamic model as the pre-processing of the nerve fibres allows for accurate prediction of the nerve fibres spiking patterns (as seen in results found in Chapter 3) by prediction the tuning curves, spiking diagrams, ISI histograms of both pure tones and multi-tone stimulation.

What stimulation pulse rate and number of electrodes (Du Preez, 2011) are required to find an optimal solution for implementing the cochlea travelling wave model in a cochlear implant?

The work of Du Preez, 2011 showed that the travelling wave process could achieve synchronisation of nerve fibres to a pure tone stimulus. However, it required space and time sampling which reduced the spatial resolution to gain temporal sampling resolution. The correct encoding of pitch is dependent on the correct choice of the space-time trade-off for the electrodes and sampling rate. The method of space-time sampling is too difficult to map out to gain an optimal solution to encode the frequency information in the auditory nerve fibres. The answer to solving this research question is in the development of a new stimulation method, resulting in the alternative travelling processor algorithm.

The newly developed alternative travelling wave processing algorithm, that follows the peak displacement of the basilar membrane at an instance of time (correlating to the maximum sampling rate of the cochlear implant), is able to encode the temporal information of the travelling wave up for frequencies above 1060 Hz (Chapter 5.3.4), eliminating the need for a space-time trade-off (Chapter 4 and 5).

The alternative travelling wave speech processor is able to predict the pitch in missing fundamental stimulus signals correctly (Chapter 5.3.5). The alternative may allow for pitch

discrimination of complex tones and speech. This shows that the entire space and time of the travelling wave may not need to be sampled to deliver quality pitch information to the auditory nerve fibres.

Can the implementation of a travelling wave processor for cochlear implant users potentially be used to improve pitch perception?

The alternative travelling wave processor pitch prediction simulations show that the travelling wave processor may have the ability to improve both pure tone temporal pitch and missing fundamental pitch encoding for cochlear implants (Chapter 5.3.4 and 5.3.5). However, due to unknowns such as the effect of place pitch mismatch caused by the cochlear implant electrode array and if spike train patterns produced by the model are valid. The alternative travelling wave processor will have to be tested in animal models. However, the electrical stimulation model used in this work (Smith, 2011) is a high quality model and the results produced in this work are promising.

7.4 CONTRIBUTION

The contribution of this work is the implementation of the auditory model, using the travelling wave model as the prediction of the basilar membrane displacement. This work expands on the normal hearing models used in literature, such as model from Zhang et al. (2001) and (Moore et al., 1990), which use filters to explain the frequency turning of the basilar membrane to stimulate the nerve fibres. The normal hearing model in this work uses the travelling wave model to describe the frequency tuning of the basilar membrane and links it to a nerve fibre model in literature which has not been done before. The auditory model is able to predict the nerve fibre characteristics that are observed and measured in normal hearing users. The normal hearing auditory model shows that without individual nerve fibre frequency selectivity, the tuning curves, temporal synchronisation, and phase locking are able to be predicted.

The auditory model was used to further develop the travelling wave processor model to be used in cochlear implants where vocoder speech processors are mainly used. This work extends on the results seen in Harczos et al. (2013) by modelling the benefits of using a travelling wave based speech processor. Travelling wave processors cause the electrical nerve fibre model to synchronise to pure tone stimuli for frequencies greater than those observed in vocoder speech processors. This work shows that cochlear implants can be used to synchronise nerve fibres to a pure tone.

The auditory model and the electrical model results were used in a pitch prediction model to estimate pitch for both the auditory model and the electrical model. The results of the auditory model showed a correlation of the upper limit maximum pitch prediction from temporal encoding.

This work allowed for the development of a normal hearing model and electrical model in order to understand and validate the encoding of temporal information in auditory nerve fibres. The validation of encoding of temporal information is important for development of processing strategies for cochlear implant users. This study utilized work from Cariani and Delgutte (1996a) to validate the development of the speech processor for cochlear implants. The alternative travelling wave processor developed in this work showed that including temporal encoding in the travelling wave speech processor could improve pitch delivery in cochlear implant users with present cochlear implant hardware.

7.5 FUTURE WORK

Future work spurred on by this study should include psychoacoustic testing, animal model testing, further advancement in the auditory models for normal hearing and electrical models, and the improvement of the cochlea travelling wave processing algorithm.

The proposed future work includes:

7.5.1 Validation of the results from the electrical model using psychoacoustics

Psychoacoustic testing will allow for the validation of the pitch prediction algorithm for pure tone stimulus in cochlear implant users. A research processor with predetermined stimuli can be used to do pitch discrimination tests on cochlear implant users.

Use of pitch discrimination tests will validate whether or not the travelling wave speech processor is improving cochlear implant pitch perception. Pitch ranking tests can be conducted using pre-processed electrode stimulation patterns.

7.5.2 Validation of the results from the electrical model using animal models

An animal model, measuring the nerve fibre spike train patterns, will validate the nerve fibre spike train patterns used in the electrical stimulation model using the alternative travelling wave processor. The results of the animal model nerve fibre spiking diagram can then be used as an input to the pitch prediction algorithm to detect the possible pitch the

animal model is hearing. This would help as we would know if the animal's nerve fibres has the correct pitch information in the nerve fibre spiking diagrams, therefore providing insight into the perceived pitch, as the animal would not be able to give feedback.

7.5.3 Improvement and further development of the normal hearing model

The normal hearing model can be further improved upon by including a more detailed middle ear model and by increasing the frequency range over the entire audible hearing range (O'Connor and Puria, 2008), this is discussed in Chapter 3, where it's evident that the middle ear model is a simple bandpass filter. A second improvement would be to accurately describe the change in basilar membrane stiffness along the length of the cochlea. The basilar membrane is like a series of tuned strings (shown in Von Békésy (1960)), changing the basilar membrane stiffness model would improve the basilar membrane delay response, allowing delay response to tuned match literature.

7.6 Testing the electrical model with superior decoding algorithms

More information may be present in the travelling wave than has been decoded in the models implemented in this work by the pitch prediction algorithm. Improving the pitch detection algorithm to include multiple pitch encoding mechanics (Cedolin and Delgutte, 2010) could improve an understanding of what information is presented by the alternative travelling wave speech processor and what information is presented by the vocoder speech processor.

7.7 SUMMARY

This work showed the implementation of an auditory model, an electrical auditory model, and a pitch estimation algorithm. It illustrated the use of a travelling wave in both an auditory model and an electrical stimulation model. The auditory model can be used as a research tool to predict nerve fibre spiking diagrams.

The electrical model shows that the travelling wave speech processors may potentially benefit cochlear implant users.

REFERENCES

Baskent, D. and Shannon, R. V. (2005). Interactions between cochlear implant electrode insertion depth and frequency-place mapping, *Journal of the Acoustical Society of America* 117(3 I): 1405-1416.

Battmer, R. D., Reid, J. M. and Lenarz, T. (1997). Performance in quiet and in noise with the Nucleus Spectra 22 and the clarion CIS/CA cochlear implant devices, *Scandinavian Audiology* 26(4): 240-246.

Berenstein, C. K., Vanpoucke, F. J., Mulder, J. J. S. and Mens, L. H. M. (2010). Electrical field imaging as a means to predict the loudness of monopolar and tripolar stimuli in cochlear implant patients, *Hearing Research* 270(1-2): 28-38.

Bialek, W., Rieke, F., De Ruyter Van Steveninck, R. R. and Warland, D. (1998). *Spikes : exploring the neural code*, edn, The MIT Press, United States of America.

Bingabr, M., Espinoza-Varas, B. and Loizou, P. C. (2008). Simulating the effect of spread of excitation in cochlear implants, *Hearing Research* 241(1-2): 73-79.

Blamey, P. J., Dowell, R. C., Clark, G. M. and Seligman, P. M. (1987). Acoustic parameters measured by a formant-estimating speech processor for a multiple-channel cochlear implant, *Journal of the Acoustical Society of America* 82(1): 38-47.

Bohne, B. A. and Carr, C. D. (1985). Morphometric analysis of hair cells in the chinchilla cochlea, *Journal of the Acoustical Society of America* 77(1): 153-158.

Bruce, I. C., White, M. W., O'leary, S. J., Dynes, S., Javel, E. and Clark, G. M. (1999a). A stochastic model of the electrically stimulated auditory nerve: Pulse- train response, *IEEE Transactions on Biomedical Engineering* 46(6): 630-637.

Bruce, L. C., White, M. W., Irlicht, L. S., O'leary, S. J., Dynes, S., Javel, E. and Clark, G. M. (1999b). A stochastic model of the electrically stimulated auditory nerve: Single-pulse response, *IEEE Transactions on Biomedical Engineering* 46(6): 617-629.

Cariani, P. A. and Delgutte, B. (1996a). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience, *Journal of Neurophysiology* 76(3): 1698-1716.

REFERENCES

Cariani, P. A. and Delgutte, B. (1996b). Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch, *Journal of Neurophysiology* 76(3): 1717-1734.

Carlyon, R. P., Deeks, J. M. and Mckay, C. M. (2010). The upper limit of temporal pitch for cochlear-implant listeners: Stimulus duration, conditioner pulses, and the number of electrodes stimulated, *Journal of the Acoustical Society of America* 127(3): 1469-1478.

Carney, L. H. (1993). A model for the responses of low-frequency auditory-nerve fibers in cat, *Journal of the Acoustical Society of America* 93(1): 401-417.

Cedolin, L. and Delgutte, B. (2005). Pitch of complex tones: Rate-place and interspike interval representations in the auditory nerve, *Journal of Neurophysiology* 94(1): 347-362.

Cedolin, L. and Delgutte, B. (2010). Spatiotemporal representation of the pitch of harmonic complex tones in the auditory nerve, *Journal of Neuroscience* 30(38): 12712-12724.

Chatterjee, M. and Zwislocki, J. J. (1997). Cochlear mechanisms of frequency and intensity coding. I. The place code for pitch, *Hearing Research* 111(1-2): 65-75.

Chatterjee, M. and Zwislocki, J. J. (1998). Cochlear mechanisms of frequency and intensity coding. II. Dynamic range and the code for loudness, *Hearing Research* 124(1-2): 170-181.

Cheatham, M. A. (2008). Comment on "Mutual suppression in the 6kHz region of sensitive chinchilla cochleae" [J. Acoust. Soc. Am.121, 2805–2818 (2007)], *The journal of the acoustical society of america* 123(2): 602-605.

Clark, G. (2003). *Cochlear implants: fundamentals and applications*, edn, Springer, New York.

Diependaal, R. J., Duifhuis, H., Hoogstraten, H. W. and Viergever, M. A. (1987). Numerical methods for solving one-dimensional cochlear models in the time domain, *Journal of the Acoustical Society of America* 82(5): 1655-1666.

REFERENCES

Diependaal, R. J. and Viergever, M. A. (1983). *Nonlinear and active modelling of cochlear mechanics: A precarious affair*. In*:* De Boer, E. and Viergever, M. A. (eds.) *Mechanics of hearing.* Nijhoff, The Netherlands, pp.

Donaldson, G. S. and Ruth, R. A. (1993). Derived band auditory brain-stem response estimates of traveling wave velocity in humans. I: Normal-hearing subjects, *Journal of the Acoustical Society of America* 93(2): 940-951.

Dreyer, A. and Delgutte, B. (2006). Phase locking of auditory-nerve fibers to the envelopes of high-frequency sounds: Implications for sound localization, *Journal of Neurophysiology* 96(5): 2327-2341.

Du Preez, C. C. (2011). *Considerations in the practical implementation of a travelling wave cochlear implant processor*, Master of Engineering (Bioengineering), University of Pretoria.

Duifhuis, H. (2012). *Cochlear Mechanics : Introduction to a Time Domain Analysis of the Nonlinear Cochlea*, edn, Springer, New York.

Duifhuis, H., Hoogstraten, H. W., Van Netten, S. M., Diependaal, R. J. and Bialek, W. (1987). Modelling the cochlear partition with coupled van der pol oscillators,

Duke, T. and Jülicher, F. (2003). Active traveling wave in the cochlea, *Physical Review Letters* 90(15): 158101/1-158101/4.

Dynes, S. B. C. and Delgutte, B. (1992). Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea, *Hearing Research* 58(1): 79-90.

Eguia, M. C., Garcia, G. C. and Romano, S. A. (2010). A biophysical model for modulation frequency encoding in the cochlear nucleus, *Journal of Physiology Paris* 104(3-4): 118-127.

Epp, B., Verhey, J. L. and Mauermann, M. (2010). Modeling cochlear dynamics: interrelation between cochlea mechanics and psychoacoustics, *The journal of the acoustical society of america* 128(4): 1870-1883.

Eskridge, E. N., Galvin Iii, J. J., Aronoff, J. M., Li, T. and Fu, Q. J. (2012). Speech perception with music maskers by cochlear implant users and normal-hearing listeners, *Journal of Speech, Language, and Hearing Research* 55(3): 800-810.

REFERENCES

Fastl, H. and Stoll, G. (1979). Scaling of pitch strength, *Hearing Research* 1(4): 293-301.

Friesen, L. M., Shannon, R. V., Baskent, D. and Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants, *Journal of the Acoustical Society of America* 110(2): 1150-1163.

Giguere, C. and Woodland, P. C. (1993). Wave digital filter model of the entire auditory periphery, *Proceedings of the 1993 IEEE International Conference on Acoustics, Speech and Signal Processing*, 1993, Piscataway, NJ, United States, Minneapolis, MN, USA, Publ by IEEE, pp. II-708-II-711.

Giguere, C. and Woodland, P. C. (1994). A computational model of the auditory periphery for speech and hearing research. I. Ascending path, *Journal of the Acoustical Society of America* 95(1): 331-342.

Goldsworthy, R. L. and Shannon, R. V. (2014). Training improves cochlear implant rate discrimination on a psychophysical task, *Journal of the Acoustical Society of America* 135(1): 334-341.

Greenberg, S. 1997. The significance of the cochlear traveling wave for theories of frequency analysis and pitch. *In:* Lewis, E. R., Steele, C. and Lyon, R. F. (eds.) *Diversity in auditory mechanics.* 1 ed. Berkeley: World Scientific Publishing.

Greenberg, S., Poeppel, D. and Roberts, T. (1998). *A space-time theory of pitch and timbre based on cortical expansion of the cochlear traveling wave delay*. In*:* Rees, A. and Meddis, R. (eds.) *Psychophysical and Physiological Advances in Hearing.* Whurr Publishers, London, pp. 293-300

Greenwood, D. D. (1990). A cochlear frequency-position function for several species - 29 years later, *Journal of the Acoustical Society of America* 87(6): 2592-2605.

Gundersen, T., Skarstein, O. and Sikkeland, T. (1978). A study of the vibration of the basilar membrane in human temporal bone preparations by the use of the mossbauer effect, *Acta Oto-Laryngologica* 86(1-6): 225-232.

REFERENCES

Harczos, T., Chilian, A. and Husar, P. (2013). Making use of auditory models for better mimicking of normal hearing processes with cochlear implants: The SAM coding strategy, *IEEE Transactions on Biomedical Circuits and Systems* 7(4): 414-425.

Hartmann, R. and Klinke, R. 1990. Response characteristics of nerve fibers to patterned electrical stimulation. *In:* Miller, J. M. and Spelman, F. A. (eds.) *Cochlear Implants.* New York: Springer-Verlag.

Hartmann, W. M. (2005). *Signals, Sound, and Sensation*, edn, Springer Science, New York.

Hodgkin, A. L. and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve, *The Journal of physiology* 117(4): 500-544.

Javel, E. 1990. Acoustic and Electrical Encoding of Temporal Information. *In:* Miller, J. M. and Spelman, F. A. (eds.) *Cochlear Implants.* New York: Springer-Verlag.

Javel, E. and Mott, J. B. (1988). Physiological and psychophysical correlates of temporal processes in hearing, *Hearing Research* 34(3): 275-294.

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(4): 1115-1122.

Kelly, O. E., Johnson, D. H., Delgutte, B. and Cariani, P. (1996). Fractal noise strength in auditory-nerve fiber recordings, *Journal of the Acoustical Society of America* 99(4 I): 2210-2220.

Kemp, D. T. (1979). Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea, *Archives of Oto-Rhino-Laryngology* 224(1-2): 37-45.

Kemp, D. T. (2002). Otoacoustic emissions, their origin in cochlear function, and use, *British Medical Bulletin* 63(223-241.

Kim, D. O., Parham, K., Sirianni, J. G. and Chang, S. O. (1991). Spatial response profiles of posteroventral cochlear nucleus neurons and auditory-nerve fibers in unanesthetized decerebrate cats: Response to pure tones, *Journal of the Acoustical Society of America* 89(6): 2804-2817.

REFERENCES

Knill, D. C. and Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation, *Trends in Neurosciences* 27(12): 712-719.

Lachs, G., Al-Shaikh, R. and Bi, Q. (1984). A neural-counting model based on physiological characteristics of the peripheral auditory system. V. Application to loudness estimation and intensity discrimination, *IEEE Transactions on Systems, Man and Cybernetics* 14(6): 819-836.

Laneau, J. and Wouters, J. (2004). Multichannel place pitch sensitivity in cochlear implant recipients, *JARO - Journal of the Association for Research in Otolaryngology* 5(3): 285-294.

Larsen, E., Cedolin, L. and Delgutte, B. (2008). Pitch representations in the auditory nerve: Two concurrent complex tones, *Journal of Neurophysiology* 100(3): 1301-1319.

Loizou, P. C. (1999). Introduction to cochlear implants, *IEEE Engineering in Medicine and Biology Magazine* 18(1): 32-42.

Loizou, P. C. 2006. Speech processing in vocoder-centric cochlear implants. *In:* Møller, A. R. (ed.).

Loizou, P. C., Dorman, M. and Tu, Z. (1999). On the number of channels needed to understand speech, *Journal of the Acoustical Society of America* 106(4 I): 2097-2103.

Mcdermott, H. J. and Mckay, C. M. (1997). Musical pitch perception with electrical stimulation of the cochlea, *Journal of the Acoustical Society of America* 101(3): 1622-1631.

Mcginley, M. J., Charles Liberman, M., Bal, R. M. and Oertel, D. (2012). Generating synchrony from the Asynchronous: Compensation for cochlear traveling wave delays by the dendrites of individual brainstem neurons, *Journal of Neuroscience* 32(27): 9301-9311.

Mckay, C. M., Mcdermott, H. J. and Carlyon, R. P. (2000). Place and temporal cues in pitch perception: Are they truly independent?, *Acoustic Research Letters Online* 1(25-30.

REFERENCES

Meddis, R. and Hewitt, M. J. (1991a). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification, *Journal of the Acoustical Society of America* 89(6): 2866-2882.

Meddis, R. and Hewitt, M. J. (1991b). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. II: Phase sensitivity, *Journal of the Acoustical Society of America* 89(6): 2883-2894.

Moore, B. C. J. and Ernst, S. M. A. (2012). Frequency difference limens at high frequencies: Evidence for a transition from a temporal to a place code, *Journal of the Acoustical Society of America* 132(3): 1542-1547.

Moore, B. C. J. and Glasberg, B. R. (2004). A revised model of loudness perception applied to cochlear hearing loss, *Hearing Research* 188(1-2): 70-88.

Moore, B. C. J., Peters, R. W. and Glasberg, B. R. (1990). Auditory filter shapes at low center frequencies, *Journal of the Acoustical Society of America* 88(1): 132-140.

Mountain, D. C. and Hubbard, A. E. (1995). *Computational Analysis of Hair Cell and Auditory Nerve Processes*. In*:* Hawkins, H. L., Mcmullen, T. A., Popper, A. N. and Fay, R. R. (eds.) *Auditory Compuatation.* Springer-Verlay, New Your, pp.

Neely, S. T. (1981). Finite difference solution of a two-dimensional mathematical model of the cochlea, *Journal of the Acoustical Society of America* 69(5): 1386-1393.

Nogueira, W., Büchner, A., Lenarz, T. and Edler, B. (2005). A psychoacoustic "NofM"-type speech coding strategy for cochlear implants, *Eurasip Journal on Applied Signal Processing* 2005(18): 3044-3059.

O'connor, K. N. and Puria, S. (2008). Middle-ear circuit model parameters based on a population of human ears, *Journal of the Acoustical Society of America* 123(1): 197-211.

Oxenham, A. J. (2013). Revisiting place and temporal theories of pitch, *Acoustical Science and Technology* 34(6): 388-396.

Oxenham, A. J., Bernstein, J. G. W. and Penagos, H. (2004). Correct tonotopic representation is necessary for complex pitch perception, *Proceedings of the National Academy of Sciences of the United States of America* 101(5): 1421-1425.

REFERENCES

Oxenham, A. J., Micheyl, C., Keebler, M. V., Loper, A. and Santurette, S. (2011). Pitch perception beyond the traditional existence region of pitch, *Proceedings of the National Academy of Sciences of the United States of America* 108(18): 7629-7634.

Penninger, R. T., Kludt, E., Büchner, A. and Nogueira, W. (2015). Stimulating on multiple electrodes can improve temporal pitch perception, *International Journal of Audiology* 54(6): 376-383.

Peterson, L. C. and Bogert, B. P. (1950). A Dynamical theory of the cochlea, *The Journal of the Acoustical Society of America* 22(3):

Plonsey, R. and Barr, R. C. (2007). *Bioelectricity A quantitive approach*, Third edn, Springer science, New York.

Pressnitzer, D., Patterson, R. D. and Krumbholz, K. (2001). The lower limit of melodic pitch, *Journal of the Acoustical Society of America* 109(5 I): 2074-2084.

Rattay, F. (1990). *Electrical nerve stimulation*, First edn, Springer-verlag Wien, New York.

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(4):

Rhode, W. S. (2007). Mutual suppression in the 6 kHz region of sensitive chinchilla cochleae, *Journal of the Acoustical Society of America* 121(5): 2805-2818.

Rhode, W. S., Roth, G. L. and Recio-Spinoso, A. (2010). Response properties of cochlear nucleus neurons in monkeys, *Hearing Research* 259(1-2): 1-15.

Rodieck, R. W., Kiang, N. Y. and Gerstein, G. L. (1962). Some quantitative methods for the study of spontaneous activity of single neurons, *Biophysical journal* 2(351-368.

Rose, J. E., Brugge, J. F., Anderson, D. J. and Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey, *Journal of Neurophysiology* 30(4): 769-793.

Rose, J. E., Brugge, J. F., Anderson, D. J. and Hind, J. E. (1968). *Patterns of activity in single auditory nerve fibres of the squirrel monky*. In*:* De Reuck, A. V. S. and

Knight, J. (eds.) *Hearing Mechanisms in vertebrates.* J. & A. Churchill LTD., London, pp. 144 - 168

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(4): 2151-2163.

Ruggero, M. A. and Temchin, A. N. (2007). Similarity of traveling-wave delays in the hearing organs of humans and other tetrapods, *JARO - Journal of the Association for Research in Otolaryngology* 8(2): 153-166.

Sachs, M. B. and Abbas, P. J. (1974). Rate versus level functions for auditory nerve fibers in cats: tone burst stimuli, *Journal of the Acoustical Society of America* 56(6): 1835-1847.

Shamma, S. A. (1985a). 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(5): 1612-1621.

Shamma, S. A. (1985b). 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(5): 1622-1632.

Shamma, S. A., Chadwick, R. S., Wilbur, W. J., Morrish, K. A. and Rinzel, J. (1986). A biophysical model of cochlear processing: intensity dependence of pure tone responses, *The journal of the acoustical society of america* 80(1): 133-145.

Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics, *Hearing Research* 11(2): 157-189.

Shannon, R. V., Cruz, R. J. and Galvin, J. J. (2011). Effect of stimulation rate on cochlear implant users' phoneme, word and sentence recognition in quiet and in noise, *Audiology and Neurotology* 16(2): 113-123.

Shapiro, W. H. and Bradham, T. S. (2012). Cochlear implant programming, *Otolaryngologic Clinics of North America* 45(1): 111-127.

REFERENCES

Shaw, E. a. G. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane, *Journal of the Acoustical Society of America* 56(6): 1848-1861.

Shepherd, R. K. and Javel, E. (1997). Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status, *Hearing Research* 108(1-2): 112-144.

Smith, P. H., Massie, A. and Joris, P. X. (2005). Acoustic stria: Anatomy of physiologically characterized cells and their axonal projection patterns, *Journal of Comparative Neurology* 482(4): 349-371.

Smith, S. (2011). *Modelling psychoacoustic experiment outsomes from space-time neural patterns for acoustic and electrical hearing*, Bioengineering Master's thesis, University of Pretroria.

Spahr, A. J. and Dorman, M. F. (2004). Performance of Subjects Fit with the Advanced Bionics CII and Nucleus 3G Cochlear Implant Devices, *Archives of Otolaryngology - Head and Neck Surgery* 130(5): 624-628.

Srinivasan, A. G., Padilla, M., Shannon, R. V. and Landsberger, D. M. (2013). Improving speech perception in noise with current focusing in cochlear implant users, *Hearing Research* 299(29-36.

Stenfelt, S., Puria, S., Hato, N. and Goode, R. L. (2003). Basilar membrane and osseous spiral lamina motion in human cadavers with air and bone conduction stimuli, *Hearing Research* 181(1-2): 131-143.

Stickney, G. S., Zeng, F. G., Litovsky, R. and Assmann, P. (2004). Cochlear implant speech recognition with speech maskers, *Journal of the Acoustical Society of America* 116(2): 1081-1091.

Stöver, T., Issing, P., Graurock, G., Erfurt, P., Elbeltagy, Y., Paasche, G. and Lenarz, T. (2005). Evaluation of the Advance Off-Stylet insertion technique and the cochlear insertion tool in temporal bones, *Otology and Neurotology* 26(6): 1161-1170.

Strydom, T. and Hanekom, J. J. (2011). An analysis of the effects of electrical field interaction with an acoustic model of cochlear implants, *Journal of the Acoustical Society of America* 129(4): 2213-2226.

REFERENCES

Sumner, C. J., O'mard, L. P., Lopez-Poveda, E. A. and Meddis, R. (2003). A nonlinear filter-bank model of the guinea-pig cochlear nerve: Rate responses, *Journal of the Acoustical Society of America* 113(6): 3264-3274.

Swanson, B. 2005. Nucleus MATLAB Toolbox. 4.03 ed.: Cochlear Ltd.

Taft, D. A., Grayden, D. B. and Burkitt, A. N. (2009). Speech coding with traveling wave delays: Desynchronizing cochlear implant frequency bands with cochlea-like group delays, *Speech Communication* 51(11): 1114-1123.

Taft, D. A., Grayden, D. B. and Burkitt, A. N. (2010). Across-frequency delays based on the cochlear traveling wave: Enhanced speech presentation for cochlear implants, *IEEE Transactions on Biomedical Engineering* 57(3): 596-606.

Tateno, T., Nishikawa, J., Tsuchioka, N., Shintaku, H. and Kawano, S. (2013). A hardware model of the auditory periphery to transduce acoustic signals into neural activity, *Frontiers in Neuroengineering* NOV):

Temchin, A. N., Recio-Spinoso, A. and Ruggero, M. A. (2011). Timing of cochlear responses inferred from frequency-threshold tuning curves of auditory-nerve fibers, *Hearing Research* 272(1-2): 178-186.

Temchin, A. N., Rich, N. C. and Ruggero, M. A. (2008). Threshold tuning curves of chinchilla auditory-nerve fibers. I. Dependence on characteristic frequency and relation to the magnitudes of cochlear vibrations, *Journal of Neurophysiology* 100(5): 2889-2898.

Van Der Marel, K. S., Briaire, J. J., Wolterbeek, R., Snel-Bongers, J., Verbist, B. M. and Frijns, J. H. M. (2014). Diversity in cochlear morphology and its influence on cochlear implant electrode position, *Ear and Hearing* 35(1): e9-e20.

Varsavsky, A. (2010a). Cochlear implant design for better representation of basilar membrane mechanics, *Proceedings of the*, Sept 2010a, Buenos Aires, pp. 5831-5834.

Varsavsky, A. (2010b). Cochlear implant design for better representation of basilar membrane mechanics, *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference* 2010(5831-5834.

REFERENCES

Venter, P. J. and Hanekom, J. J. (2014). Is there a fundamental 300 HZ limit to pulse rate discrimination in cochlear implants?, *JARO - Journal of the Association for Research in Otolaryngology* 15(5): 849-866.

Viergever, M. A. and Diependaal, R. J. 1983. Simulataneous amplitude and phase match of cochlear model calculations and basilar membrane vibration data. *In:* De Boer, E. and Veirgever, M. (eds.) *Mechanics of hearing 1983.* The Netherlands: Delft University Press.

Von Békésy, G. (1960). *Experiments in hearing*, edn, McGraw-Hill, New York.

Wang, S., Koickal, T. J., Hamilton, A., Cheung, R. and Smith, L. S. (2015). A bio-realistic analog CMOS cochlea filter with high tunability and ultra-steep roll-off, *IEEE Transactions on Biomedical Circuits and Systems* 9(3): 297-311.

Westerman, L. A. and Smith, R. L. (1988). A diffusion model of the transient response of the cochlear inner hair cell synapse, *Journal of the Acoustical Society of America* 83(6): 2266-2276.

Wever, E. G. and Bray, C. W. (1937). The Perception of Low Tones and the Resonance-Volley Theory, *The Journal of Psychology* 3(1): 101-114.

Wojtczak, M., Beim, J. A., Micheyl, C. and Oxenham, A. J. (2012). Perception of across-frequency asynchrony and the role of cochlear delays, *Journal of the Acoustical Society of America* 131(1): 363-377.

Yang, G., Lyon, R. F. and Drakakis, E. M. (2015). A6μ W per channel analog biomimetic cochlear implant processor filterbank architecture with across channels agc, *IEEE Transactions on Biomedical Circuits and Systems* 9(1): 72-86.

Yates, G. K., Manley, G. A. and Köppl, C. (2000). Rate-intensity functions in the emu auditory nerve, *Journal of the Acoustical Society of America* 107(4): 2143-2154.

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(5): 1381-1403.

Zeng, F. G. (2002). Temporal pitch in electric hearing, *Hearing Research* 174(1-2): 101-106.

REFERENCES

Zeng, F. G. and Galvin Iii, J. J. (1999). Amplitude mapping and phoneme recognition in cochlear implant listeners, *Ear and Hearing* 20(1): 60-74.

Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A., Wei, C. and Cao, K. (2005). Speech recognition with amplitude and frequency modulations, *Proceedings of the National Academy of Sciences of the United States of America* 102(7): 2293-2298.

Zhang, X. and Carney, L. H. (2005). Analysis of models for the synapse between the inner hair cell and the auditory nerve, *Journal of the Acoustical Society of America* 118(3 I): 1540-1553.

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(2): 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(3): 1446-1466.

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(5): 2390-2412.

Zilany, M. S. A. and Carney, L. H. (2010). Power-law dynamics in an auditory-nerve model can account for neural adaptation to sound-level statistics, *Journal of Neuroscience* 30(31): 10380-10390.

Zwislocki, J. (1950). Theory of the Acoustical Action of the Cochlea, *The Journal of the Acoustical Society of America* 22(6): 778 - 784.