COCHLEAR IMPLANT SPEECH PROCESSING, BASED ON THE COCHLEAR TRAVELLING WAVE

by

Hendrik Petrus Wolmarans

9221891

Submitted in partial fulfilment of the requirements for the degree

Master of Engineering (Bio-engineering)

In the

Faculty of Engineering, Built Environment and Information Technology

UNIVERSITY OF PRETORIA, PRETORIA

August 2005

COCHLEAR IMPLANT SPEECH PROCESSING, BASED ON THE COCHLEAR TRAVELLING WAVE

by

Hendrik Petrus Wolmarans

Supervisor: Prof JJ Hanekom

Departement Elektriese, Elektroniese en Rekenaar- Ingenieurswese

Degree: M.Eng. (Bioengineering)

Keywords: Travelling wave, basilar membrane, hydrodynamic, model, cochlea, cochlear implant, speech processing.

A cochlear implant is a prosthetic device that can provide severe-to-profoundly deaf individuals with partially restored hearing. It emulates the function of a normal cochlea through combined functioning of externally situated electronics and an electrode array surgically implanted into the cochlea. Speech coding strategies implemented in speech processors aim to stimulate the auditory nerve in a way similar to that of a normal working cochlea by modelling the way the cochlea processes sound.

Current speech processing strategies rely on the tonotopicity of the cochlea, i.e. the relation between distance from the base of the cochlea and the specific frequency that causes the highest amplitude of deflection at the specific point. The phenomenon of the travelling wave on the basilar membrane is thus reduced to its point or points of maximal deflection.

In this study, the behaviour along the full length of the basilar membrane will be investigated in the time domain, i.e. the deflection along the whole membrane for any point in time, in order to evaluate the relevance of the travelling wave in coding sound in a cochlear implant system. The additional information acquired by emulating the motion of the fluid and the basilar membrane in the cochlea, will be transmitted to the recipient in

electrical stimulus patterns, to assess whether it provides recipients of cochlear implants with better pitch perception. It will be shown that for the individuals that partook in the experiments, improvement of discrimination around 100 Hz were obtained when compared to current speech coding strategies like the advanced combination encoder (ACE) speech coding strategy in the same recipient.

KOGLEÊRE INPLANTING SPRAAKVERWERKING GEBASEER OP DIE KOGLEÊRE LOOPGOLF

deur

Hendrik Petrus Wolmarans

Studieleier: Prof JJ Hanekom

Departement Elektriese, Elektroniese en Rekenaar- Ingenieurswese

Graad: M.Ing. (Bio-ingenieurswese)

Sleutelwoorde: Loopgolf, basilaarmembraan, hidrodinamiese model, model, koglea, kogleêre inplanting, spraakprosessering.

′n Kogleêre inplanting is ′n prostese wat ′n mate van gehoor kan herstel vir individue met ′n ernstige-tot-totale gehoorverlies. ′n Kogleêre inplanting beoog om die normale funksionering van die oor te emuleer deur die werking van beide ′n eksterne spraakprosesseerder en ′n sjirurgies-geïnplanteerde elektrode en stimulator. Spraakkoderingstrategieë, wat in die spraakprosesseerder geïmplementeer word, beoog om die ouditiewe senuwee op ′n soortgelyke wyse te stimuleer as die senuwee van ′n normale oor deur so goed moontlik die werking van die koglea te modelleer.

Spraakprosesseeringstrategieë wat tans gebruik word, berus egter slegs op die tonotopisiteit van die koglea. Dit behels dat na mate opwaarts in die koglea beweeg word, die koglea toenemend meer sensitief is vir dalende frekwensies. Hierdie sensitiwiteit spruit uit die maksimale defleksie van die basilaarmembraan op elke punt vir ′n spesifieke frekwensie, wat ooreenstem met die posisie in die koglea. Die loopgolfverskynsel wat op die basilaarmembraan voorkom word dus vereenvoudig na die punt of punte van maksimale uitwyking vanuit die rusposisie.

In hierdie studie word die verplasing langs die basilaarmembraan in die tyd-domein gemodelleer, ge-enkodeer en vir kogleêre inplantinggebruikers gespeel om sodoende die relevansie van die loopgolf in die kodering van klank en veral toonhoogtes in ′n kogleêre inplanting te ondersoek. Resultate word voorgelê wat ′n verbetering in toonhoogtediskriminasie om 100 Hz aandui, wanneer vergelyk word met huidige spraakkoderingstrategieë soos ‘advance combination encoder’ (ACE) in dieselfde gebruiker.

TABLE OF CONTENTS

Chapter 1 INTRODUCTION

1.1 PROBLEM STATEMENT

A cochlear implant is a prosthetic device that can provide severe-to-profoundly deaf individuals with partially restored hearing (Loizou 1998). It emulates the function of a normal cochlea through combined functioning of externally situated electronics and an electrode array surgically implanted into the cochlea (Figure 1. 1). The electrode array of the Nucleus 24 Contour system consists of 22 half-banded electrodes on a precurved silicone carrier. Electrodes face the modiolus of the cochlea, where the spiral ganglion cell bodies are situated.

Incoming sound is picked up by a directional microphone, from where it is passed to a speech processor that digitises the analogue sound. These signals are continually streamed to the implanted receiver-stimulator by modulating a high-frequency carrier. The stimulator decodes the modulated signal, resulting in activation of appropriate electrodes on the array. Activation of electrodes in turn results in stimulation of the auditory nerve.

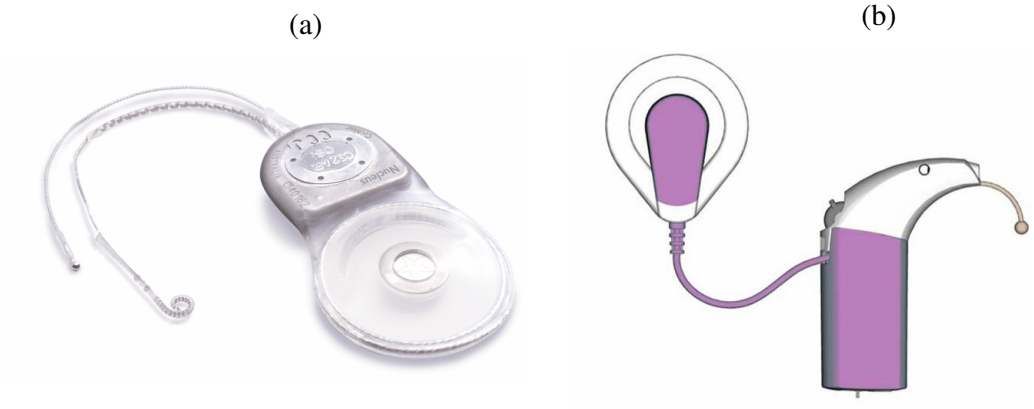

Figure 1.1 (a) The Nucleus 24 Contour cochlear implant, showing the electrode array on the silicone carrier and the subcutaneous stimulator. (b) The Nucleus Esprit 3G speech processor with directional microphone, as well as the external receiver. These components represent the internal and external parts of the implant system.

The speech coding strategies implemented in speech processors aim to stimulate the auditory nerve in a way similar to that of a normal working cochlea by modelling the way the cochlea processes sound. In a normal functioning cochlea, the sound pressure wave is transmitted via the outer ear and vibrates the eardrum. The eardrum and the three middle-ear bones in turn vibrate the oval window, matching the impedance of the cochlea in the way the movement is reduced while increasing the pressure. The vibration of the oval window sets the perilimph fluid inside the cochlea's canals in motion, causing pressure differences between the scala vestibule and scala tympani. These pressure differences cause the basilar membrane, that divides these two scala (if the scala media is ignored), to deflect. This initial deflection, combined with the changing pressure differences initiates a wave travelling from the base of the cochlea towards its apex. Due to the mechanics of the basilar membrane, the amplitude of the travelling wave increase (for a single pure tone) until it reaches a point of maximum deflection before it dies away. The growth of the travelling wave is also accompanied with a marked slowing down of the wave as it approaches the point of maximal deflection. As the basilar membrane deflects from its rest position, tiny hair-cells on the membrane cause the neurons of the auditory nerve in its vicinity to fire, causing a hearing sensation to the person.

At present, commercially available speech coding strategies for cochlear implant speech processors, for example advanced combination encoder (ACE) (Advanced Combination Encoders) and continuous interleaved sampling (CIS) strategies are based mainly on the tonotopical arrangement of the cochlea, i.e. the link between a specific frequency and the point of maximal deflection in the cochlea. For both these strategies, incoming sound is separated into a number of frequency bands and then presented to the auditory nerve via the implant.

In the continuous interleaved sampling (CIS) strategy, the outputs of the filters are sampled at a high rate and presented to the auditory nerve by means of sequential pulses, each filter or frequency band corresponding to an electrode. If the word "*choice*" is processed using the continuous interleaved sampling (CIS) speech processing strategy, the stimuli can be shown in an electrodogram (Figure 1.2), where the vertical axis represents the electrode position (corresponding to the specific filter's frequency) and on the horizontal axis, the time progression. It can be seen from Figure 1.2 how the intensity of the stimulation changes as the consonants are followed by vowels and again by the consonants.

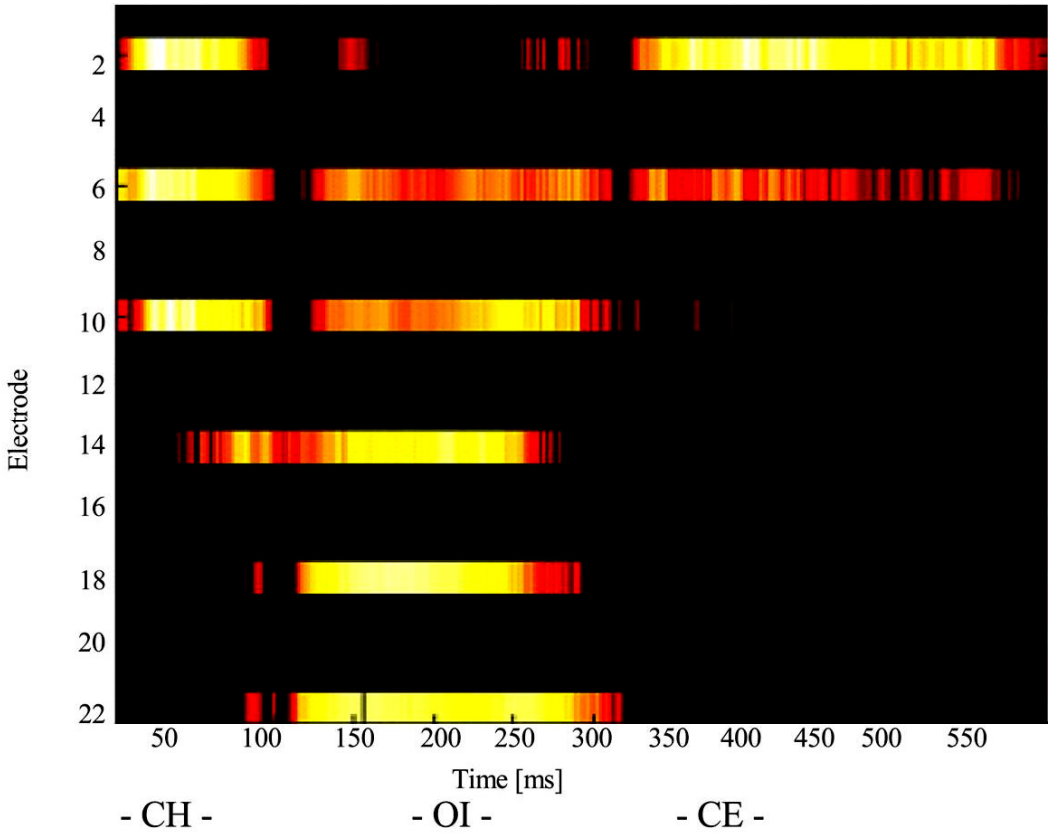

Figure 1.2 Electrodogram of the word 'Choice' processed with continuous interleaved sampling (CIS) speech processing strategy. The colours indicate the amplitude of stimulation on the specific electrode at a specific time – the brighter the colour, the higher the amplitude.

In the advanced combination encoder (ACE) speech coding strategy more channels are used for stimulation, but only the frequency bands or filter outputs with the most energy are presented to the auditory nerve for each time sample. The resultant stimuli pattern is a 'roving'-type where only a limited number of electrodes (typically more than eight) are stimulated during each time cycle, despite being selected from a large number of electrodes (see Figure 1.3). Comparing the electrodogram generated by processing the word "*choice*" with the advanced combination encoder (ACE) strategy to the one from continuous interleaved sampling (CIS) processing, it can be seen that advanced combination encoder (ACE) has an improved frequency resolution, while continuous interleaved sampling (CIS) has a finer temporal resolution due to its higher stimulation rate.

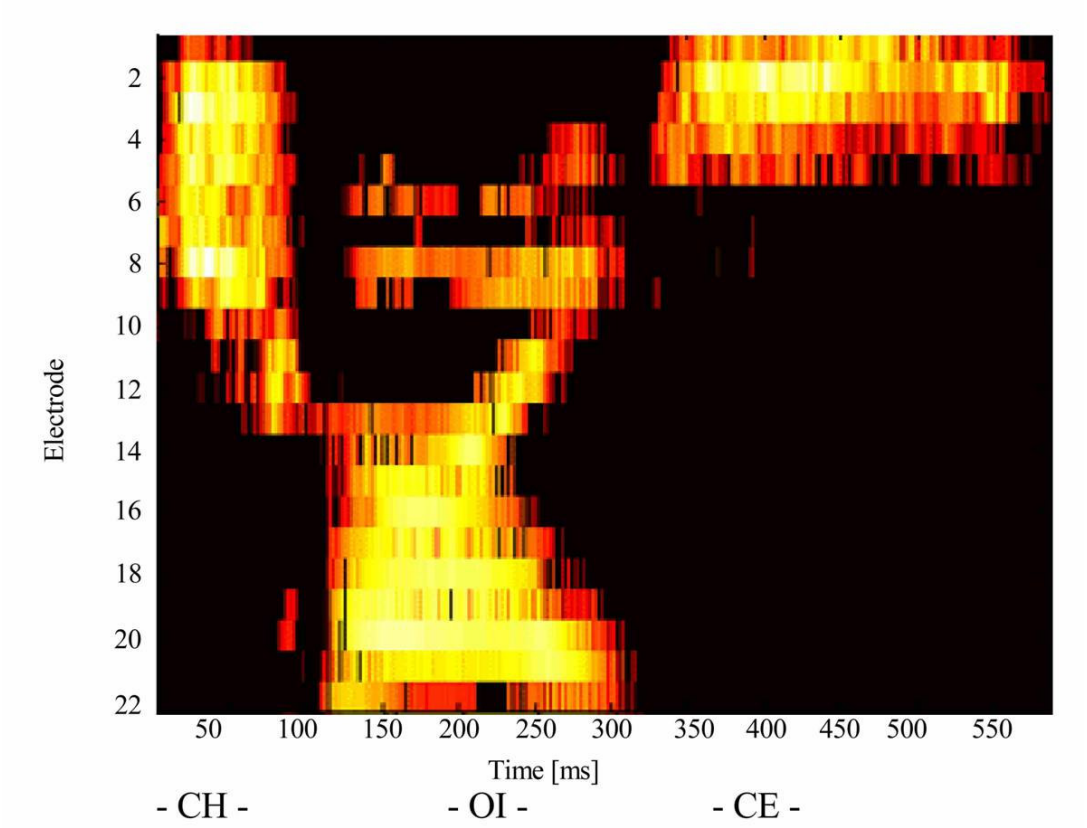

Figure 1.3 Electrodogram of the word 'Choice' processed in advanced combination encoder (ACE). The colour scheme is the same as in Figure 2

However, both these speech processing strategies are limited by the width and overlap of their filters in terms of absolute frequency discrimination, i.e. they lack fine temporal structure of the incoming audio signal. Spectral resolution in current implementations of these strategies is negatively influenced by filters that are more than 100 Hz wide (at -10 dB from the peak response), as shown in Figure 1.4. There is some overlap of the filter frequency responses, resulting in stimulation on two to six electrodes as a pure tone is sweeped from 100 Hz to 1 kHz, with frequency resolution at the low frequencies being affected most. To discriminate between two pure tones that are closely spaced in the frequency domain, the absolute loudness of the two stimuli should be compared on two adjacent electrodes.

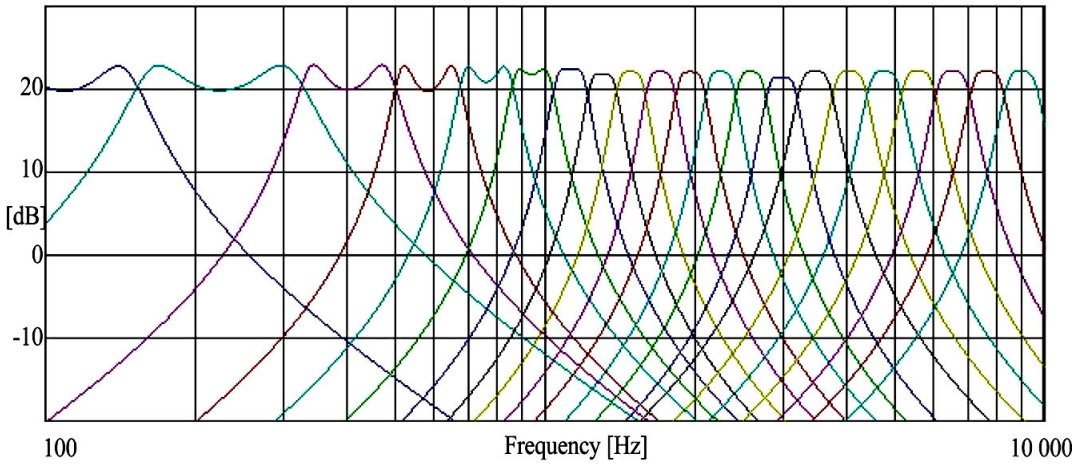

Figure 1.4 Speech processor filterbank frequency responses.

When using only the tonotopisity of the cochlea, the fine time structure of the travelling wave (Von Bekesy, 1947) on the basilar membrane is ignored. The travelling wave caused by pressure differences between the scala tympani and scala vestibuli when the oval window is displaced by the stapes, always starts at the base of the cochlea and moves toward the apex. It increases monotonically in amplitude while travelling towards the apex

until a place of maximum deflection, associated with a specific, characteristic frequency for a pure tone input (Von Bekesy, 1947; Rhode, 1971; Lighthill, 1981; Duifhuis, 1988; Ruggero, 1994), is reached and then rapidly dies away. Although the displacement amplitude of the travelling wave peaks as it reaches this point, the wave velocity slows dramatically, i.e. a rapid increase in phase delay, and its energy is dissipated easily with even the smallest amount of damping (Lighthill, 1981). It is thought possible that the temporal pattern of the basilar membrane (i.e. the synchronous firing of groups of neurones in phase with the negative deflection of the basilar membrane) and the slowing down of the travelling wave (i.e. asymptotic approach of the characteristic frequency over a relatively long time period) may assist a normal hearing individual with the detection of small pitch differences and doing so even in competing in noise. Currently implemented cochlear implant speech processing strategies assume that only the positions of the peaks of the travelling wave envelope carry information, ignoring the existence of travelling waves. The gradual growing and dying away of the travelling wave on the basilar membrane and the rapid increase in phase-shift as the wave approaches the point of maximal deflection is therefore not incorporated in current speech processing strategies such as continuous interleaved sampling (CIS) or advanced combination encoder (ACE).

Users of current speech coding strategies are limited in pitch discrimination ability by the frequency bandwidth assigned to each electrode, which in turn is a function of the number of electrodes implanted. Theoretically this limits a user's discrimination between two adjacent electrodes to a frequency resolution of roughly 100 Hz steps at 100 Hz, quite removed from the almost 1 Hz frequency resolution of a normal hearing individual around 100 Hz.

In this study, the behaviour along the full length of the basilar membrane will be investigated in the time domain, i.e. the deflection along the whole membrane for any point in time, in order to evaluate the relevance of the travelling wave in coding sound in a cochlear implant system. The additional information acquired by emulating the motion of the basilar membrane, will be transmitted to the recipient in the electrical stimulus patterns, to assess whether it provides recipients of cochlear implants with better pitch discrimination.

1.2 RESEARCH QUESTION AND HYPOTHESIS

The primary research question addressed by this study is:

"Can spectral information presented to cochlear implant recipients be improved by incorporating more information regarding the travelling wave?"

The hypothesis is that *when a hydrodynamic model of the basilar membrane is used to act as encoder for sound, spectral information should be more accurately perceived by cochlear implant users than when using current strategies.* This hypothesis was tested by using such a model as a speech coding strategy. As will be shown, sound was processed using a travelling wave emulating algorithm and presented to recipients in a range of experiments. Pitch discrimination studies, during which this travelling wave encoding strategy was compared to current, commercially available strategies, were performed. Following these, pitch ranking experiments was done to establish the tonal quality of the stimuli. Only pure tones were used as input for the speech processing in this study.

1.3 OBJECTIVES

Objectives for this dissertation were the following.

First, an appropriate model of the basilar membrane was selected from literature. Such a model should closely represent the motion of the basilar membrane. The family of hydrodynamic models, describing the motion of the basilar membrane through differential equations, were closely evaluated. These models describe what is actually happening, i.e. it limits the amount of assumptions made about the basilar membrane motion. The outputs of these models are in good agreement with measured data.

Second, a technique to solve these differential equations in the time domain has to be developed. The solution or output of these equations should be the position of the basilar membrane in discrete time steps.

Thirdly, a technique should be developed to convert the output of these equations to a format that can be used by a cochlear implant speech processor to present desired stimuli to a cochlear implant recipient. The solution will therefore be used as a signal processing strategy for the cochlear implant system, indicating where and when the implant should stimulate the cochlea.

Finally, it is necessary to experimentally determine how effective this travelling wave encoding strategy transmits frequency information to the electrically stimulated auditory system, when compared to existing strategies.

1.4 APPROACH

To achieve the objectives set out in the previous paragraph, the following approach will be taken.

1.4.1 Hydrodynamic model

A suitable model that describes the basilar membrane displacement needs to be found from literature. Such a model should closely represent the physical aspects of the cochlea, as well as the motion of both the cochlear fluid and basilar membrane. The differential equations describing the pressure distribution of the fluid on the basilar membrane, resulting in its movement, need to be solved numerically to predict the position of the basilar membrane for discrete time increments. Being able to process any input signal (not just single-frequency sinusoids) will also make such a model suitable for processing of speech and other complex sounds. A balance between complex, accurate models (e.g. three-dimensional models which include macro- and micromechanics inside the cochlea and active sharpening through the outer hair cells) (Neely & Kim, 1986; Wan & Fan, 1991; Böhnke & Arnold, 1999; De Boer & Nuttall, 1999; Steele & Lim 1999) and simplistic models (e.g. one-dimensional models using transmission line theory) (Peterson & Bogert, 1950), is likely to be two-dimensional hydrodynamic models that include only macromechanics of the cochlea without any active elements, and some third-dimension effects that can easily be translated to two dimensions. An example of such an effect is the cross-section curvature of the basilar membrane during deflection (Allen & Sondhi, 1979), which is similar to a clamped string with maximal deflection in the centre that can be approximated by merely looking at the deflection in the centre of the cross-section.

1.4.2 Implementation of model in software

The chosen model will be implemented in Matlab[1]. Two sets of differential equations were used in the model. The first describes the distribution of the pressure on the basilar membrane due to the movement of the stapes and oval window as well as the basilar membrane itself. The second set of differential equations describes the acceleration of the basilar membrane due to the distributed pressure. Solving these equations numerically, the next position of the basilar membrane can be computed using basilar membrane positions from previous time steps and current pressure distribution on the basilar membrane.

A simplified model of the hair cell transfer function is used to act as a 'rectifier' of the basilar membrane movement, increasing the likelihood of the neurons firing on negative deflections.

Application of this model to a cochlear implant imposes a number of constraints. First, the number of relevant implementation sites along the basilar membrane is much less than the amount needed for accurate simulation – 22 electrodes along the length of the cochlea for the implant used. Furthermore, the current Nucleus cochlear implant limits the total stimulation rate to 14 400 pulses per second (pps), which could be, for example 1 000 pulses per second on fourteen electrodes or 2 400 pulses per second on six electrodes. Because of these restraints, the output of the model is sub-sampled in both space and time to 22 discrete sites along the basilar membrane at a total rate of 14 400 samples per second.

1. Matlab is a mathematical interpreting-type package that excels in manipulating matrices in calculations and has good graphic capabilities to display results. www.mathworks.com

The process employed by advanced combination encoder (ACE[2]) is diagrammatically shown in Figure 1.5. Digitised sound in wave-file format (usually with *.wav extension) is used as input. The sound is processed in time blocks, adjusted by a Hanning window to limit boundary-effects during the Fourier transform before going through a fast Fourier transform filter with 64 linear output bins. The output bins are combined using frequency allocation tables to equal the number of channels available in the recipient's MAP, a set of recipient specific parameters that controls the presentation of sound to the recipient to optimise hearing performance. Of the total number of channels present in the recipient's MAP, only those with the most information or highest levels of filter outputs are chosen. The numbers of channels that are presented to the recipient at each time interval are called the number of 'maxima'. The current level necessary for stimulation is determined by converting the filter output from each of the selected channels to a logarithmic current scale by means of a non-linear transfer function.. This information is then transmitted to the cochlear implant by modulating the radio frequency carrier that also provides power.

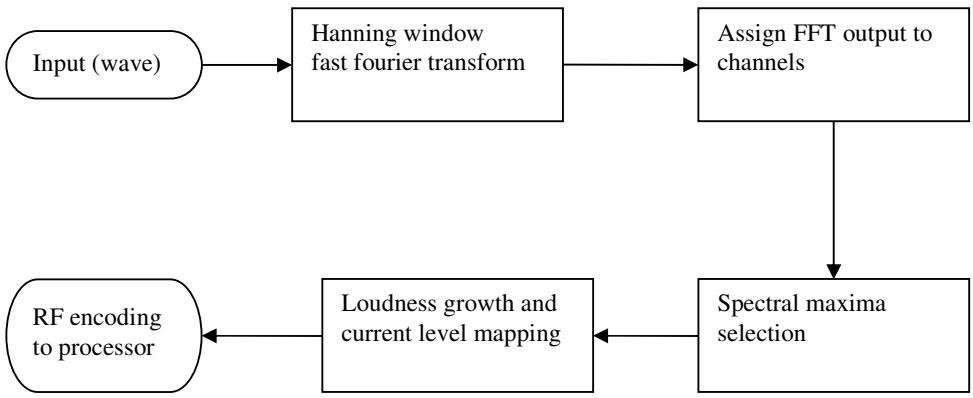

Figure 1.5 Processing steps in advanced combination encoder (ACE) encoding strategy. The input is in wave-sound file format.

2. Advanced combination encoder (ACE) is the default speech coding strategy used with the Nucleus 24 cochlear implant system of Cochlear Ltd., a leading, Australian based cochlear implant manufacturer. www.cochlear.com.au

The travelling wave encoding strategy retains most of the processes used for the advanced combination encoder (ACE) encoding strategy in order to ease implementation and comparison of their respective results. Digitised pure tones in wave-sound file format are also used as input, which the hydrodynamic model block uses to determine the displacement of the basilar membrane for each sampled time increment (Figure 1.6). To comply with the restrictions of the current Nucleus cochlear implant system, the basilar membrane displacement is sub-sampled in both time and space, with sub-sampling in space done on 22, linearly spaced positions along the basilar membrane. Electrodes to be stimulated were selected similarly to the maxima selection in the advanced combination encoder (ACE) strategy, i.e. the highest values of the basilar membrane displacement was selected for stimulation of the corresponding electrode. The loudness growth transfer function that computes the current levels corresponding to basilar membrane displacement was changed to a linear function, instead of the non-linear one used in the advanced combination encoder (ACE) strategy. This was done to ensure that the basilar membrane displacement is analysed without distortion of non-linear transfer functions. The RF coding and transmission of the individual stimuli were kept the same as in the advanced combination encoder (ACE) strategy.

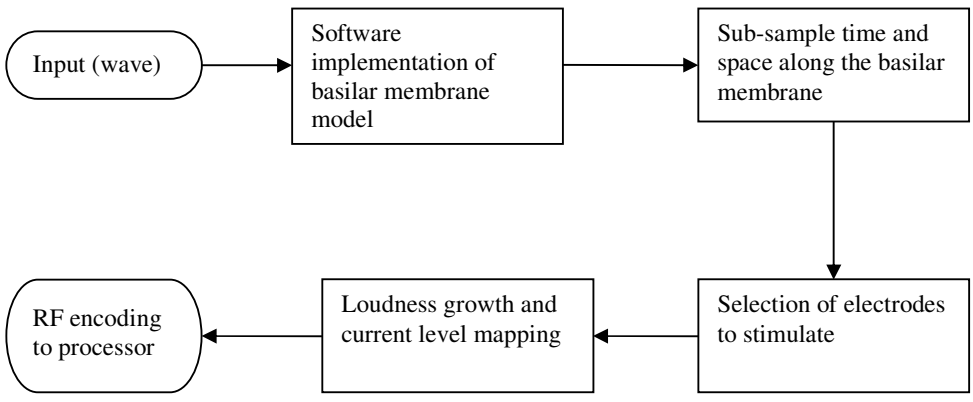

Figure 1.6 Processing steps of the proposed encoding strategy.

1.4.3 Evaluating the model (experimental work)

To evaluate the implementation of the model as a possible speech processing strategy, a first step is to compare listeners' pitch discrimination abilities for pure tones processed by the advanced combination encoder (ACE) strategy (Figure 1.5) and the travelling wave encoding strategy (Figure 1.6) respectively. There is a suggested link between increased pitch discrimination ability and improved speech perception (Moore, 1986; Hanekom & Shannon, 1996; Henry, McKay *et. al.*, 2000). By using simple stimuli in an experimental setup where sound can be pre-processed, the potential usefulness of a new strategy can be evaluated without the need to implement the strategy in real time within the constraints of the speech processor.

To ensure that discrimination is not linked to loudness, an adaptive loudness balancing procedure will be performed to eliminate loudness cues from the different stimuli. It will be followed by a discrimination experiment using 21 pure tones in 1 Hz increments, around 100 Hz, to establish the ability of the recipient to distinguish between each of the 21 tones. A similar experiment will be done with 21 pure tones in 10 Hz increments, around 1 kHz. The output of this experiment will be displayed in a confusion matrix that is expected to have a triangular form, with good discrimination for tones far apart and poorer discrimination for tones close together.

If sufficient discrimination is achieved with the travelling wave encoding strategy, a pitch-ranking experiment will be done to establish whether the new sounds only differ from each other or whether they can also be assigned a rising tonal quality linked to the rising frequency of the inset stimuli.

1.5 CONTRIBUTION

This work contributes to current knowledge by implementing a more complex basilar membrane model as speech processing algorithm in cochlear implant speech processors. The model incorporates the travelling wave on the basilar membrane and adds both temporal pitch, by way of the periodic deflection of the basilar membrane, and the deceleration effect of the wave as it approaches its characteristic frequency, to the information presented to the cochlear implant recipient.

It also provides a basic model from which more complex models of the outer, middle and inner ear (and hair cell function) can be implemented during future research. The modular approach of the Nucleus Matlab toolbox not only makes the additional processing steps created for current experiments ideal building blocks for future research, but will also be easily compatible with additional building which may be added to the proposed speech coding strategy.

1.6 OVERVIEW

In the second chapter, a detailed literature study will be presented, explaining the background of the problems experienced with current speech coding strategies implemented in cochlear implant speech processors. Models describing the functioning of the basilar membrane will be discussed, as well as their possible applications for and effects on frequency and speech detection. Possible solutions will also be discussed.

In the third chapter, the different hydrodynamic models will be discussed, considering their respective strengths and weaknesses. One-, two- and three-dimensional models with solutions in both time- and frequency-domain will be explored and discussed. One such two-dimensional model with solutions in the time-domain will be chosen from these models and implemented in Matlab software. Problems encountered during this process, as well as solutions or alternatives, will also be discussed in this chapter.

The fourth chapter will focus on the integration of the above mentioned model into the Nucleus Implant Communicator (NIC) toolbox in Matlab, to provide a fast-prototyping of this model and comparison with current, commercial strategies. Cochlear Ltd.'s NIC software will be used as platform to do a rapid prototype evaluation of sound processing using this model. NIC acts as a software interface between the programming system and speech processor. It consists of a dynamic linked library (dll) and a toolbox written for use with Matlab (Nucleus Matlab Toolbox). Using the toolbox functions in Matlab, almost any possible stimulation pattern can be presented to a recipient. Detailed discussion of how the toolbox was used will be used and changes made to the model to integrate it with the toolbox will be discussed in Chapter 3.

The fifth chapter will discuss the experimental procedure followed. It will mention the case histories of the recipients that participated in the study and will focus specifically on electrodograms and captured radio-frequency data that show the stimulus pattern and levels of stimulation. The experimental setup of each experiment will be discussed and results will be shown.

The sixth chapter will discuss the results obtained during the experimental sessions. Results obtained from current speech coding strategies will be compared to those from the evaluated strategy and discussed in terms of the hydrodynamic model. The effectiveness of using the discussed model as a basis for speech coding strategies will be evaluated, followed by some recommendations for improvements and further investigations of using the same building blocks.

In the final chapter, the dissertation will be analysed and discussed in terms of the hypothesis, objectives and problem stated in the first chapter.

Chapter 2 LITERATURE STUDY

2.1 INTRODUCTION

In this chapter the background of current, commercial speech coding strategies and some of the problems still experienced with these are discussed. Various, more complex possibilities of how the cochlea processes sound is also discussed and compared with one another.

2.2 BACKGROUND LITERATURE STUDY

The concept of the travelling wave on the basilar membrane has long been known (Von Bekesy, 1947) and is commonly accepted (Allen, 1977; Shamma, 1985a; Duifhuis, 1988; Prosen & Moody, 1991; Ruggero, 1994; Uppenkamp, Fobel & Patterson., 2001; Moore, Alcatrana & Glasberg, 2002), yet the models of the cochlear mechanics implemented in current, commercial cochlear implant speech coding strategies only include the points of maximal deflection on the basilar membrane (Loizou, 1998). Considering only the tonotopicity of the cochlea simplifies speech processing to only filtering of incoming sound, followed by e.g. maxima selection and current level mapping when using advanced combination encoder (ACE). This basic model of how the cochlea processes sound led to the development of signal processors that used symmetrical filter-banks to decompose the input signal into its spectral components (Loizou, Dorman and Fitzke, 2000).

Loizou (1998) gave a short overview of the history of cochlear implants and existing speech processing strategies. Speech perception performance of cochlear implant recipients improved over the years, with many recipients being able to understand speech without lip-reading and some even able to use telephones (Loizou,1998). Speech

production skills are closely related to speech perception and Loizou reported gradual improvements over time up to four years after the start of using the cochlear implant.

One of the improvements of continuous interleaved sampling (CIS) and advanced combination encoder (ACE) over previous speech processing strategies is the ability to encode more temporal information into the stimuli presented to the cochlea, i.e. representing rapid envelope variations (Loizou, 1998). The travelling wave encoding strategy continues along this line by not only considering the envelope variations, but also looking the travelling wave on the basilar membrane, in an attempt to incorporate it in the stimuli presented to the cochlear implant recipient.

Other parameters of how the cochlea processes sound have also been reported. Von Bekesy (1947) suggested that phase measurements may be better in determining the pattern of vibration (i.e. position and pitch) than considering the amplitude envelope alone. Shamma (1985a) has shown that the distribution of the relative phases of synchronised activity in the auditory nerve is an important indication of spectral parameters. He has shown in particular that, for regions basal to the characteristic frequency point on the basilar membrane, the relative phase shifts are small (fast propagating travelling wave) in comparison to those at the characteristic frequency point as the travelling wave decelerates. These rapid phase shifts can be seen in the auditory nerve as steep and localised spatial discontinuities along the tonotopical axis of fibres that otherwise fire in a relatively uniform instantaneous pattern if the frequency is lower than the characteristic frequency point. This is visible for the characteristic frequency locations that correspond to low frequencies (less than or equal to 1.5 – 2 k Hz for cats). Shamma (1985b) further argues that lateral inhibitory networks contribute to processing of these localised spatial discontinuities. In his defence of the travelling wave phenomenon in the cochlea, Ruggero

(1994) also uses the experimentally measured phase delays associated with points close to the characteristic frequency as proof of the existence of the travelling wave. It can be concluded from the above that the change in phase shifts and deceleration of the travelling wave close to the characteristic frequency carry some useful information regarding the pitch of incoming sound. Using a model of the basilar membrane that can show such deceleration of the travelling wave should therefore be able to convey this information to the cochlear implant user.

An implementation that used phase information of the incoming sound to time the stimuli of the cochlear implant was developed by Hartmann, Tönder and Klinke (1998), but it failed to reflect the changing phase delays seen in the cochlea as the travelling wave travelled on the basilar membrane. In order to simplify calculation, it ignored the relative phase shifts and in particular the sharp discontinuity in phase shift/delay, seen close to the characteristic frequency. This implementation was not clinically evaluated.

In the normal cochlea, for low stimulation rates, the neural firing pattern locks on to the phase of the travelling wave and is most likely to stimulate when the basilar membrane is moving towards the scala vestibuli, with almost no stimuli when moving towards the scala tympani (Rose, Hind, *et. al.,* 1971). A high degree of phase locking exists up to 1 kHz, after which it declines to approach zero at 5 kHz. This dependence on stimulus frequency can be attributed to the capacitance of the basal walls of the inner hair cells (Russel & Sellick, 1978). Maximum firing rate of auditory fibres appear to be affected by the saturation of the input/output function of the inner hair cells (for frequencies lower than the characteristic frequency), as well as the basilar membrane non-linearity at frequencies higher than the characteristic frequency, resulting in the travelling wave not propagating beyond the characteristic frequency position (Pattuzzi & Sellick, 1983). Moore (1986)

suggested, in general agreement with the above, that information in the phase locking of neural pulses up to 5 kHz is essential for effective frequency discrimination. He further suggests that both phase locking information and place information are used for complex stimuli. The relevance of these outcomes to the proposed study will be discussed extensively in Chapter 6, but it implies that, in the low frequency range, the information contained within the periodic displacement of the basilar membrane can convey additional pitch information to the cochlear implant recipient.

In current speech coding strategies, a non-linear transfer function is used when processing sound (e.g. loudness growth function in Figure 1.6). Such an additional transfer function should not be necessary for the travelling wave encoding strategy if the non-linearity of the basilar membrane displacement is incorporated in the hydrodynamic model. The non-linear response of the basilar membrane has been attributed to the non-linear transfer function of the stereocillia deflection to the receptor potential of the outer hair cells (Preyer & Gummer,1996). Others have suggested that the mechanical non-linearity of the Dieters cells is a more relevant cause for the non-linear basilar membrane motion (Böhnke & Arnold, 1998). Sellick, Pattuzzi and Johnstone (1982) stated that the cochlear mechanics are actually sharply tuned in comparison to neural tuning, which can result in the sharp tuning seen in inner hair cells and auditory nerve fibres. In general, it can be stated that the stimulation of the auditory nerve follows the low frequency movement of the basilar membrane. The relatively broad sensitive region around 1 kHz that is independent of the position in the cochlea might be linked to the change in viscous coupling from velocity-related to displacement-related coupling (Pickles, 1986).

2.3 SUMMARY

In conclusion, current cochlear implant speech processing strategies provide users with good speech discrimination, but are as yet unable to convey all the auditory information needed for good music appreciation and discrimination of closely spaced pitch differences at low frequencies. It is argued that this inability is intrinsic to the current model of the cochlea and basilar membrane underlying current commercial speech coding strategies. This model reduces the cochlea and basilar membrane to a simple tonotopic function that stimulates auditory nerves only at the point corresponding to the characteristic frequency and all across the cochlea at roughly the same time. However, development of speech processing strategies that convey fine temporal structure of the envelope (e.g. advanced combination encoder (ACE) and continuous interleaved sampling (CIS)) has significantly improved recipients' speech perception performance. This study aims to further improve performance by not only providing additional information regarding fine temporal structure, but by also re-evaluating the basilar membrane model that is used. The concept of a travelling wave on the basilar membrane is well accepted when considering the functioning of a normal hearing individual's cochlea. The implementation of hydrodynamic models in cochlear implant speech processing strategies has until now, however, been prevented by the sheer computational power needed to solve such models – even with the travelling wave encoding strategy sound currently needs to be processed prior to stimulation because of the amount of time required for information processing. Implementation of this model provided additional information about the phase and progressive change of phase delay of stimuli, period of low frequency sounds that enables the recipient to discriminate pitch because of phase locking of the auditory nerve and the intrinsic non-linear amplitude response of the basilar membrane.

Chapter 3 HYDRODYNAMIC MODEL AND IMPLEMENTATION

3.1 INTRODUCTION

Following the discussion in the previous chapter, the different hydro-dynamic models will be discussed in this chapter, considering their various strengths and weaknesses. One-, two- and three-dimensional models with solutions in both the time- and frequency-domains will be explored and discussed. From these models one will be chosen and implemented in Matlab software. Problems encountered during this process (with their solutions or alternatives) will also be discussed.

3.2 BACKGROUND

In order to overcome the constraints of current models, mathematically modelling the fluid mechanics of the cochlea was investigated. As computing power and modelling methodology improved, increasingly complex models were developed. The earliest of these were one-dimensional (transmission line) models, which showed only qualitative correlation with experimental values (Peterson & Bogert, 1950) - specifically a broad area of maximal deflection around the characteristic frequency that was not observed in later measurements (Rhode, 1971). Two-dimensional models, incorporating two-dimensional flow of fluid in the scala and output in the frequency domain were subsequently developed (Allen, 1977; Sondhi, 1978). These models were refined to consider longitudinal stiffness but more pertinent to the aim of the current study, to also provide output in the time domain (Allen & Sondhi, 1979). Refinements of such two-dimensional models included the addition of the effect of viscous fluid on the whole model (Wan & Fan, 1991). Subsequent three-dimensional models analysed the effect of a viscous fluid (Steele & Taber, 1979b).

Johnstone, Pattuzzi and Yates (1986) discussed some of the experimental evidence pointing to possible 'feed forward' elements that contribute to the non-linear amplitude response of the basilar membrane. Such response behaviour is characterised by high sensitivity and sharp spectral tuning at low amplitudes which decreases with increasing amplitude levels. Although one of the earlier models that included active elements (Neely & Kim, 1986) was linear and only in one dimension, it included micromechanics, considering longitudinal stiffness and using eardrum pressure as input. As computing power increased, some older two-dimensional models were greatly refined to include the micromechanics of the Organ of Corti (Allen & Neely, 1992) as well as active elements. More recently, active elements were also included in three-dimensional models (Steele & Lim, 1999; De Boer & Nuttall, 1999), all implemented in the frequency domain.

The use of finite difference and finite element methods further assisted the solution of this complex hydrodynamic system. Finite difference methods were first used in two dimensions, since they required less computing power and were easier to implement directly from the differential equations (Allen & Sondhi, 1979; Steele & Taber, 1979a; and Neely, 1981). Implementation in three dimensions followed later (Steele & Taber, 1979b). Similarly, finite elements were also used first in two-dimensional solutions (Janssen, Segal & Viergever, 1978), followed by later three-dimensional solutions which included the micromechanics (Organ of Corti) (Böhnke & Arnold,l 1998) as well as fluid – structure couplings (Böhnke & Arnold, 1999).

For purposes of this study, the model of Allen and Sondhi (1979) was most useful, since it is the most advanced existing model that describes the position of the basilar membrane in the time domain. Since it is aimed to link the basilar membrane position for each sampled

point in time as a reference for stimulation, the output has to be in the time domain. This solution also includes some non-linear basilar membrane properties that are best modelled in the time domain, since working in the frequency domain has various assumptions that rely on linear superposition.

3.3 IMPLEMENTATION

Allen and Sondhi (1979) describe the fluid motion and that of the basilar membrane motion separately by using Green's theorem to relate the distributed pressure across the basilar membrane to its acceleration. Together with the plate equation that describes the basilar membrane motion resulting from this acceleration, these two equations completely specify the basilar membrane motion resulting from movement of the oval window. This model assumes the cochlear fluid to be incompressible, does not account for coiling and also excludes lateral fluid motion (movement in the third dimension). Some non-linear basilar membrane properties, for example damping, are included.

Allen (1977) extended these assumptions so that basilar membrane velocity and scala pressure are related through by the integral equation

$$p(x,t) = \rho \int_0^L G(x,x') \ddot{\xi}(x',t) dx' + (L-x)\rho \dot{u}(t) \qquad \textbf{(3.1)}$$

where $p(x,t)$ represents the scala pressure, ρ the cochlear fluid density, L the length of the cochlea, $\xi(x,t)$ the basilar membrane displacement, x the position along the basilar membrane and $u(t)$ the stapes velocity. G represents Green's function. Green's function has been derived previously (Allen, 1977; Sondhi, 1978) but is in principle the solution to Laplace's equation for the pressure distribution on the basilar membrane, given the acceleration at any point along the membrane. Allen and Sondhi (1979) unfolded the

cochlea and transformed it into a periodic extension with period 4L, resulting in a 'real-odd harmonic' symmetry that can be expanded as a Fourier series, the following equations were mostly taken from Allen and Sondhi (1979). This symmetry is so called because the Fourier series expansion of such a periodic function consists only of odd-harmonic cosine terms.

$$2p((x)) = \rho\int_{-2L}^{2L} F((x-x'))\ddot{\xi}((x'))dx' + 2((L-|x|))\rho\dot{u} \qquad \textbf{(3.2)}$$

This then leads to where the replacement $G(x,x') = F(|x-x'|) + F(|x+x'|)$ has been made from Equation 3.1 and F has real-odd harmonic symmetry. The double brackets denote the function as being periodic with period 4L.

Equation 3.2 summarises the fluid coupling inside the cochlea and the resultant pressure on the basilar membrane. The second part of Allen and Sondhi's (1979) hydrodynamic model describes the basilar membrane displacement, $\xi(x,t)$, in terms of all the forces acting on it. If the basilar membrane is modelled as a homogeneous, anisotropic plate, the equation for the plate's displacement is

$$\mathrm{q(x,z,t)} = D_x\frac{\partial^4\xi}{\partial x^4} - 2(D_x D_z)^{1/2}\left(\frac{\pi}{\mathrm{W}(x)}\right)^2\frac{\partial^2\xi}{\partial x^2} + D_z\left(\frac{\pi}{\mathrm{W}(x)}\right)^4\xi \qquad \textbf{(3.3)}$$

where q is the load on the membrane and the parameters D_x and D_y are functions of the internal structure, thickness and modulus of elasticity of the basilar membrane. These parameters are assumed to be constants but not necessary equal (homogeneous and anisotropic).

The load on the basilar membrane consists of the inertia of the basilar membrane $m\ddot{\xi}$, the viscous damping $R(x)\dot{\xi}$ and the trans-basilar membrane pressure, $2p$. When combining

Equations 3.2, 3.3 and these loads on the basilar membrane, the basilar membrane motion is given by

$$D\xi + R(x)\dot{\xi} + m\ddot{\xi} = -\rho F \otimes \ddot{\xi} - 2(L - |x|)\rho\dot{u} \qquad \textbf{(3.4)}$$

where

$$D\xi = D_x \frac{\partial^4 \xi}{\partial x^4} - 2(D_x D_z)^{1/2}\left(\frac{\pi}{W(x)}\right)^2 \frac{\partial^2 \xi}{\partial x^2} + D_z\left(\frac{\pi}{W(x)}\right)^4 \xi \qquad \textbf{(3.5)}$$

The symbol $\otimes$ denotes the circular convolution $\rho\int_{-2L}^{2L} F(x-x')\dot{v}(x')dx'$ with each term having a period of 4L, allowing it to be expanded into a Fourier series. R*(x)* denotes the damping, ξ the basilar membrane displacement (cm), $\dot{\xi}$ the basilar membrane velocity (mm/s) and $\ddot{\xi}$ the basilar membrane acceleration in (cm/s^2). ρ is the cochlear fluid density, L the length of the unrolled cochlea (mm), x the position along the basilar membrane (mm) and $\dot{u}$ the velocity of the stapes movement (mm/s). *Dx* and *Dz* are the longitudinal and transversal stiffness of the basilar membrane respectively, while W*(x)* describes the basilar membrane width as a function of *x*.

This equation can be solved as an initial value problem in time, with boundary values in space. The boundary problem is solved using Fourier transforms, while the initial value problem is solved recursively, starting with the basilar membrane's initial displacement and velocity. Equations 3.4 and 3.5 can be solved numerically if the space and time variables are discretisised as follows:

$$\ddot{\xi}(t) = F_s^2(\xi_n - 2\xi_{n-1} + \xi_{n-2}) \qquad \textbf{(3.6)}$$

and

$$\dot{\xi}(t) = F_s\left(\xi_{n-1} - \xi_{n-2}\right) \tag{3.7}$$

where n is the discrete time interval and F_s is the sampling frequency. The circular convolution in Equation 3.4 can also be solved in the discrete time domain when written as $\frac{\rho L F_s^2}{M}\sum_{m=1}^{4M} F_{k-m}\left(\xi_{m,n} - 2\xi_{m,n-1} + \xi_{m,n-2}\right)$. M represents the number of spatial samples along the basilar membrane, k the discrete spatial intervals and both F_{k-m} and ξ_m is periodic with a period of 4M. A sampled version of F_{k-m} is given by Allen and Sondhi (1979),

$$F_k = \begin{cases} \dfrac{L}{2H} - \dfrac{1}{\pi}\left\{\ln\left(\dfrac{\pi L}{2MH}\right) - 1\right\}, k = 0 \\ \dfrac{L}{2H}\left(1 - \dfrac{k}{M}\right) - \dfrac{1}{\pi}\ln\left(1 - e^{\frac{-\pi k L}{HM}}\right), k \neq 0 \end{cases} \tag{3.8}$$

where H is the height of the scala.

If the terms containing $\ddot{\xi}$ in the discretisised Equations 3.4 and 3.5, are moved to the left, it results in

$$\frac{\rho L F_s^2}{M}\sum_{m=1}^{4M} F_{k-m}\left(\xi_{m,n} - 2\xi_{m,n-1} + \xi_{m,n-2}\right) + mF_s^2\left(\xi_n - 2\xi_{n-1} + \xi_{n-2}\right) = b(x, t) \tag{3.9}$$

where

$$b(x,t) = -D\xi - R(x)F_s\left(\xi_{n-1} - \xi_{n-2}\right) - 2\left(1 - \frac{|k|}{M}\right)L\rho F_s\left(u_{n-1} - u_{n-2}\right) \tag{3.10}$$

To solve for ξ, the left of Equation 3.9 needs to be changed in such a way that $\ddot{\xi}$ can be isolated. Allen and Sondhi (1979) solved this by using the Fourier transform, thereby reducing the convolution to a multiplication in the frequency domain.

In order to do so, they defined

$$Q(x) = \rho F(x) + \left(\frac{m}{2}\right)\left(\delta(x) - \delta(2L - x)\right) \tag{3.11}$$

where $\delta(x)$ is the Dirac δ function, with the factor of 0.5 and the second $\delta(x)$ -term necessary to give Q(x) real-odd harmonic symmetry. Equation 3.9 and Equation 3.10 can now be solved, using the Fourier transform.

$$F_s^2\left(\xi_{\mathrm{n}} - 2\xi_{\mathrm{n-1}} + \xi_{\mathrm{n-2}}\right) = \mathbf{F}^{-1}\left(\frac{\mathbf{F}\{b(x,t)\}}{\mathbf{F}\{Q(x)\}}\right) \tag{3.12}$$

Computing $\mathbf{F}\{Q(x)\}$ once therefore computes the model. The following steps are subsequently iterated until all the input samples are processed.

- Compute $b(x,t)$ in Equation (3.10) using previously computed values of ξ.
- Compute the fast Fourier transform of this result.
- Divide by the previously computed $F(Q_m)$.
- The inverse fast Fourier transform will yield acceleration vector $\ddot{\xi}$.
- Using recursive methods, the next displacement vector ξ, is computed.

The resulting estimate of the basilar membrane position is proportional to the probability of the nerves firing at each point, with a higher probability in the case of deflection in the negative direction.

During implementation of the model of Allen and Sondhi (1979) in Matlab, some problems were encountered.

- The spatial fast Fourier transform, when first implemented, caused zeros in the frequency domain. These zero points prevented the division in the third step of the above algorithm. It was later realised that they were created because of the periodic extension of function F prior to computing the fast Fourier transform, knowing that the fast Fourier transform inherently expand any such function. The initial inability to prevent these zeros resulted in seeking alternative methods for computing the basilar membrane movement.

  An alternative method of solving Equations 3.9 and 3.10 required many repetitions to minimise the error function, since the expression could not be solved explicitly. The second-order terms (unknowns) were moved to the left of the equation with the standard and first-order terms (known values of previous calculations) on the right. The method attempting to solve Equations 3.9 and 3.10 without using the fast Fourier transform follows.

  If the terms containing ξ_n are isolated on the left, the following is obtained:

$$\mathrm{m}F_s^2\xi_\mathrm{n} + \frac{\rho \mathrm{L}F_s^2}{\mathrm{M}} \sum_{m=0}^{4\mathrm{M}-1} \mathrm{F}_{\mathrm{k}-\mathrm{m}} \left(\xi_{\mathrm{m,n}}\right) = \mathrm{A}(\xi) \tag{3.13}$$

where

$$\mathrm{A}(\xi) = -\mathrm{D}\xi - \mathrm{R}(x)F_\mathrm{s}\left(\xi_{\mathrm{n-1}} - \xi_{\mathrm{n-2}}\right) - \mathrm{m}F_s^2\left(-2\xi_{\mathrm{n-1}} + \xi_{\mathrm{n-2}}\right)\ldots$$
$$-\frac{\rho \mathrm{L}F_s^2}{\mathrm{M}} \sum_{m=0}^{4\mathrm{M}-1} \mathrm{F}_{\mathrm{k}-\mathrm{m}} \left(-2\xi_{\mathrm{m,n-1}} + \xi_{\mathrm{m,n-2}}\right) - 2\left(1 - \frac{|k|}{\mathrm{M}}\right) \mathrm{L}\rho F_\mathrm{s}\left(u_{n-1} - u_{n-2}\right) \tag{3.14}$$

$D\xi$ is the same as in Equation 3.5.

ξ_n (the new displacement vector of the basilar membrane) can now be solved using numerical methods. $A(\xi)$ can be computed from already known values of ξ (ξ_{n-1} and ξ_{n-2}) and the input variables u_{n-1}, u_{n-2}. If the summation in Equation 3.13 is expanded,

$$mF_s^2\xi_n + \frac{\rho L F_s^2}{M}\begin{bmatrix} F_{1-1}\xi_{n,1} & F_{2-1}\xi_{n,1} & \cdots & F_{M-1}\xi_{n,1} \\ F_{1-2}\xi_{n,2} & F_{2-2}\xi_{n,2} & \cdots & F_{M-2}\xi_{n,2} \\ \cdots & \cdots & \cdots & \cdots \\ F_{1-4M}\xi_{n,M} & F_{2-4M}\xi_{n,M} & \cdots & F_{M-4M}\xi_{n,M} \end{bmatrix} = A(\xi) \qquad \textbf{(3.15)}$$

where i in $\xi_{n,i}$ denotes the position along the basilar membrane. If each position along the basilar membrane considered individually (i.e. the columns in Equation 3.15),

$$mF_s^2\xi_{n,i} + \frac{\rho L F_s^2}{M}\left(F_{i-1}\xi_{n,1} + F_{i-2}\xi_{n,2} + F_{i-3}\xi_{n,3} + \ldots + F_{i-4M}\xi_{n,4M}\right) = A(\xi_i) \qquad \textbf{(3.16)}$$

for every i, where

$$A(\xi_i) = -D\xi_i - R(x)F_s\left(\xi_{n-1,i} - \xi_{n-2,i}\right) - mF_s^2\left(-2\xi_{n-1,i} + \xi_{n-2,i}\right) - K_i \ldots \\ -2\left(1 - \frac{|i|}{M}\right)L\rho F_s\left(u_{n-1} - u_{n-2}\right) \qquad \textbf{(3.17)}$$

and

$$K_i = \frac{\rho L F_s^2}{M}\left(F_{i-1}\left(-2\xi_{n-1,1} + \xi_{n-2,1}\right) + \ldots + F_{i-4M}\left(-2\xi_{n-1,4M} + \xi_{n-2,4M}\right)\right) \qquad \textbf{(3.18)}$$

Equation 3.16 can be written in matrix form, so that it can easier be solved in Matlab:

$$m F_s \begin{bmatrix} \xi_{n,1} & \dots & \xi_{n,M} \end{bmatrix} - \frac{L\rho F_s^2}{M} \begin{bmatrix} \xi_{n,1} & \dots & \xi_{n,4M} \end{bmatrix} \begin{bmatrix} F_0 & \dots & F_{M-1} \\ F_{-1} & \dots & F_{M-2} \\ \dots & \dots & \dots \\ F_{1-4M} & \dots & F_{-3M-1} \end{bmatrix} = \begin{bmatrix} a_1 & \dots & a_M \end{bmatrix} \quad \textbf{(3.19)}$$

Since Equation 3.17 can be solved using previous displacement values of the basilar membrane and displacement of the oval window, the right side of Equation 3.19 is known. ξ_n can now be solved by choosing a starting value for ξ_n and computing the left side of Equation 3.19. The vector values on both sides of Equation 3.19 will differ. To solve for the correct value of ξ_n, the difference between the left and right side of Equation 3.19 should be minimised. Repetitions, with appropriate minimising or zero-seek methods resulted in a reasonably accurate guess with an acceptably small error.

- The method described above required excessive computing times, e.g.168 hours to compute the basilar membrane movement during 25 ms of sound. Attempts to shorten the computing time by accepting larger errors or by increasing step sizes in approaching the zero error caused the models to become unstable.
- Zero-seek methods, using the gradient of changes in the error function to reduce the number of iterations without causing instability, did not reduce the time required to calculate the basilar membrane response to an acceptable level. To minimise the error typically required between a few and tens of millions of iterations.
- In order to minimise the above mentioned calculation time, multiresolution analysis was investigated (Vetterli, 1991). The spatial resolution in the basal portion of the cochlea needs to be higher than in the apical portion, in order to have an accurate and stable calculation of basilar membrane movement. Analysing the basilar

membrane in several portions, with varying spatial sampling, helped to minimise the size of the matrices and therefore computational time, without the subsequent restrictions on highest analysed input frequency.

- Functions for finding zero points built into the MATLAB toolboxes also improved calculation time, but often led to instability.

As mentioned above, when using the equation from Allen and Sondhi (1979) without expanding it in the frequency domain, the above problems could be solved. With current MATLAB code, one second of basilar membrane movement due to a sound, takes approximately 90 seconds to compute and store when simulated on a Pentium IV, 1.4 GHz processor, using a Windows 2000 operating system.

3.4 RESULTS AND DISCUSSION

When using similar cochlear parameters, the results obtained are in good agreement with the graphical output of the model used by Allen and Sondhi (1979). The steep roll-off above the characteristic frequency can clearly be seen as the travelling wave slows down to approach characteristic frequency (Figure 3.1).

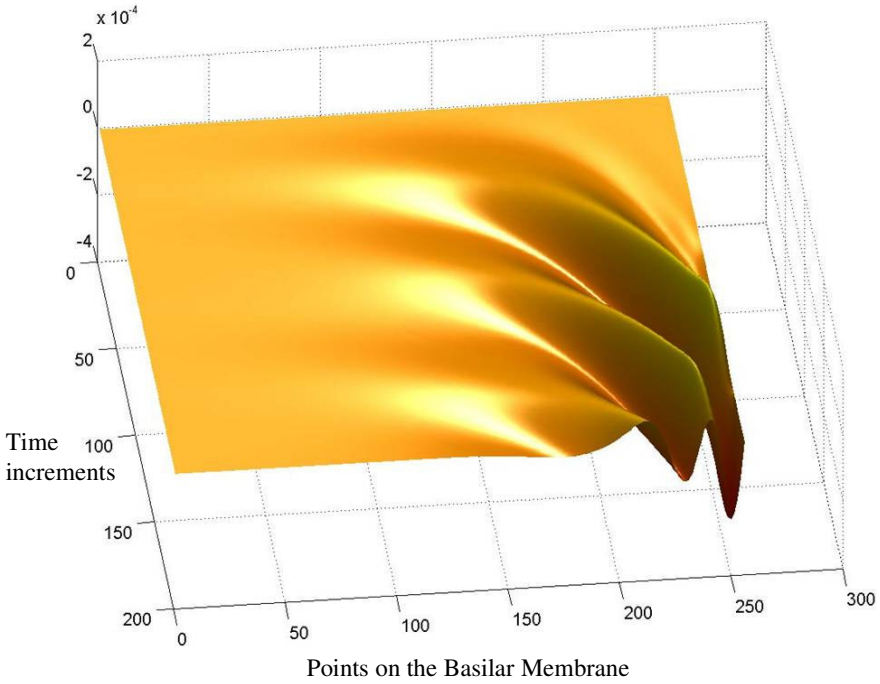

Figure 3.1 Basilar Membrane response to a low-frequency pure tone input

The travelling wave motion can also be seen in Figure 3.2, with the blue, green and red curves indicating the displacement of the basilar membrane on three neighbouring points along the basilar membrane. The red curve represents the displacement of the most apical point. The figure not only shows how the amplitude of the travelling wave increases as it approaches the point of maximal deflection, but also how the wave travels towards the apex, i.e. there is a time delay in the peaks of the deflection waveform.

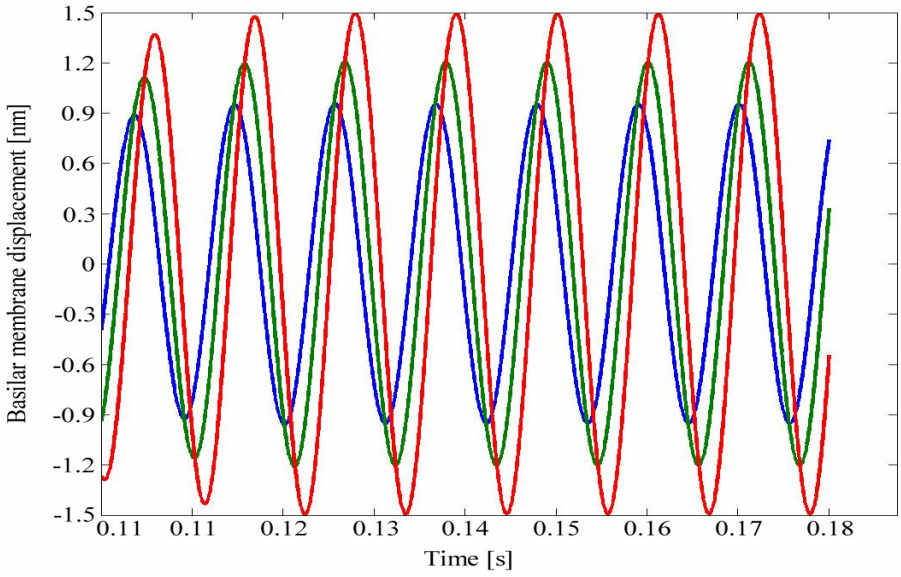

Figure 3.2 Traveling wave on the basilar membrane.

It was therefore possible to implement the model, described by Allen and Sondhi (1979), in Matlab. The output from this model is in the time domain, which allows the output to be implemented as a cochlear implant speech coding strategy by simply translating the basilar membrane displacement to stimulus intensity.

3.5 SUMMARY

This chapter discussed different hydrodynamic models along with their respective strengths and weaknesses. One-, two- and three-dimensional models with solutions in both the time- and frequency-domains has been discussed. A two-dimensional model developed by Allen and Sondhi (1979) was chosen for its time domain solutions which can be implemented as a possible speech coding strategy. Processing incoming sound provides basilar membrane displacement in the time domain. The results obtained were as expected, showing the growth and steep decrease in travelling wave amplitude along the basilar membrane length as well as the increase in phase delay, or deceleration of the

wave, as it approaches the point of maximal deflection. This complex, computational intensive solution, currently limits the implementation of the travelling wave encoding strategy in applications that can process sound prior to its presentation to cochlear implant recipients.

The current restraint on using the travelling wave encoding strategy is its computational intensive solution, currently limiting its implementation to applications that can pre-process the sound before presenting it to the recipient.

Chapter 4 INTEGRATION OF THE HYDRODYNAMIC MODEL WITH THE NUCLEUS SPEECH PROCESSING ALGORITHM

4.1 INTRODUCTION

This chapter will focus on the integration of the hydrodynamic model into the Nucleus Implant Communicator (NIC) toolbox in Matlab. This toolbox enables a fast prototyping of the suggested model, thereby circumventing development of a new code that processes sound in an entirely different way compared to current strategies. The interaction between the speech processor and cochlear implant therefore stays the same, while only the speech processing modules that needs to change, are altered. The stimuli produced using this travelling wave encoding strategy can then be compared to current, commercial strategies. Specific changes to the implementation of the model to allow integration with the NIC toolbox will be discussed. De-bugging and controls to eliminate the possibility of over-stimulation will also be discussed. The advanced combination encoder (ACE) strategy served as baseline, with only changing the modules that impacted the travelling wave encoding strategy, so as to compare the two strategies as closely as possible.

4.2 SUB-SAMPLING THE OUTPUT OF THE MODEL

The output of the model, as described in Chapter 3 has 128 discrete points along the basilar membrane, with a sampling or output frequency of 16 kHz. The Nucleus system used here has 22 electrodes in total, with a total stimulation rate limited to 14 400 pulses per second. For this specific implementation of the suggested model, discrete, linear samples along the basilar membrane were used. Since pure tones have been used for the initial experiments, it was decided to stimulate either on fourteen electrodes (those with the largest negative basilar membrane deflection during each time sample) at a rate of 1 000 pulses per second or on six electrodes at a rate of 2 400 pulses per second (see Paragraph 4.7 for results of

sub-sampling issues). The first option will only allow frequencies up to 500 Hz to be analysed before aliasing effects, due to the under sampling in time, occur whereas the second option will be used to process pure tones in the area around 1 kHz.

An example of how the basilar membrane displacement is sampled is shown in Figure 4.1, with the red solid curve representing the displacement of a more apical part of the basilar membrane and the blue curve the displacement of a more basal part. This example shows a 90 Hz pure tone sampled at 1 kHz (dashed curves).

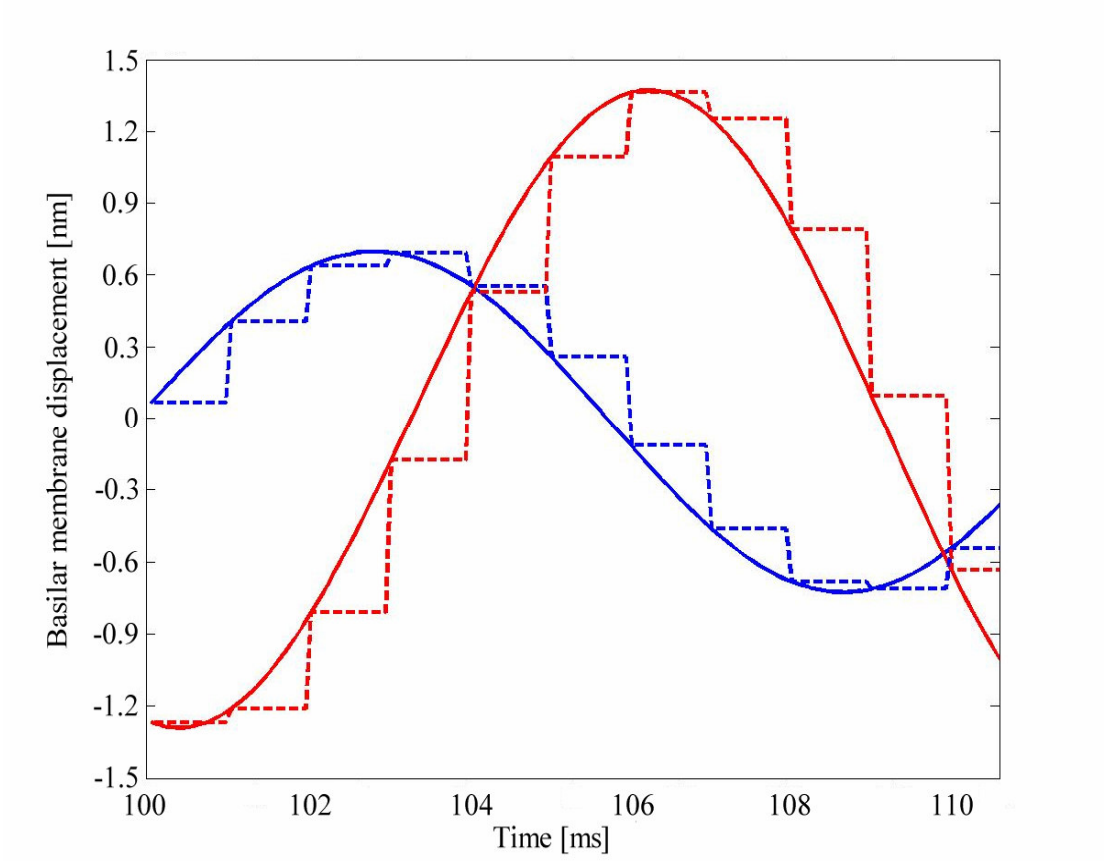

Figure 4.1 Sub-sampling of basilar membrane displacement – 90 Hz pure tone.

4.3 OVERVIEW OF PROCESSING BLOCKS IN THE NUCLEUS MATLAB TOOLBOX WHEN USING ADVANCED COMBINATION ENCODER (ACE)

The speech processing blocks available in the advanced combination encoder (ACE) speech coding strategy (Figure 4.2) were used as basis for the development of the travelling wave encoding strategy. The function of each processing block in the advanced combination encoder (ACE) speech processing strategy, as implemented in the Nucleus Matlab toolbox, is described below.

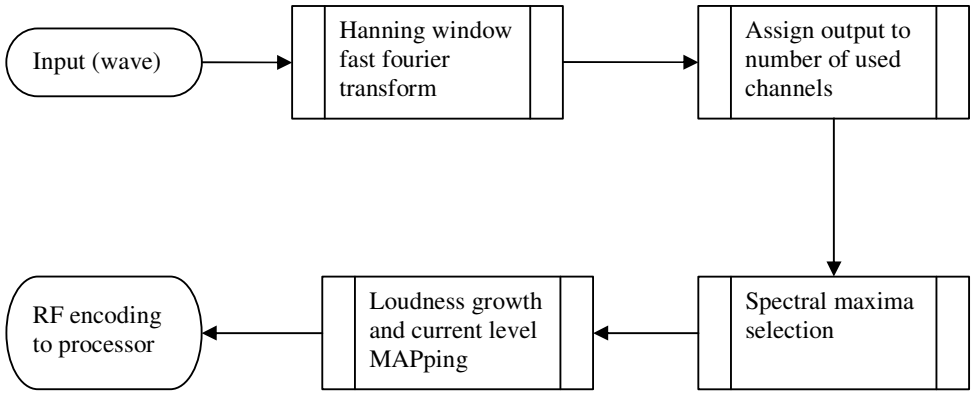

Figure 4.2 Processing steps in advanced combination encoder (ACE) encoding strategy.

- The input for the advanced combination encoder (ACE) speech coding strategy is usually sampled audio files that also contain the frequency that was used during sampling. To ensure that frequency information up to almost 8 kHz is retained, the default sampling frequency is 16 kHz.
- The Hanning window ensures that boundary errors, caused by assuming the processed window to be periodic, are minimised when processing a window of sampled sound by means of fast Fourier transform. This processing block spectrally represents an 8 ms sound window. The processing windows are overlapped to ensure that the processed spectral information that is linked to the per-channel stimulation rate is updated at an adequate rate, up to a maximum of

about 800 Hz. This window size (8 ms) therefore limits the strategy to a maximum low-frequency resolution of 125 Hz. Although increasing the window size will improve the spectral resolution, more computing time or higher power consumption will be required. Implementation constraints have therefore limited the low frequency resolution. The fact that output of fast Fourier transform-based filters is linearly distributed along the frequency axis further limits their use in processing strategies. For example, at low frequencies, 125 Hz resolution is worse than for normal hearing at low frequencies, while probably being unnecessarily good at 7 kHz.

- The 64 output values of the fast Fourier transform processing block are distributed among the number of channels that will be used. This distribution is dictated by a set of Frequency Allocation Tables (FATs), but can be modified in Matlab to include any choice of frequency boundaries. The energy of the fast Fourier transform output values, combined in a specific channel, is added and averaged to obtain the mean energy in the specified frequency range.
- In the 'spectral maxima selection' processing block a number of user-specified number of channels with the highest amount of spectral energy is selected for stimulation during a particular time window (Figure 4.3). The number of maxima is chosen by the clinician who programs the speech processor. As the incoming sound changes, the stimulated channels rove across the entire electrode array. This processing block not only filters background noise (low spectral energy), but also conserves power or increases the per-channel stimulation rate by stimulating only on selected electrodes.

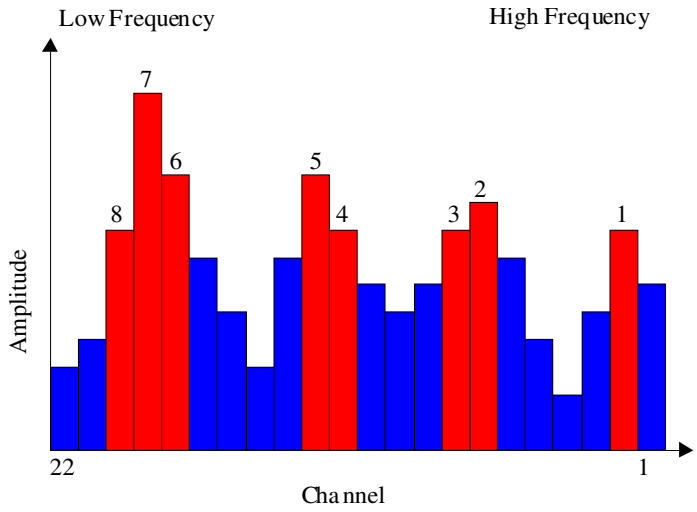

Figure 4.3 Maxima selection

- The loudness growth and current level mapping block translates the energy of the channel to specific current levels, which specify the amount of current delivered to the electrodes. In the Nucleus processor and implant used, current levels range from 1 to 255, corresponding to 10 μA and 1.75 mA respectively. The actual current changes logarithmically for linear increases in current level, i.e. an increase of 30 current levels corresponds to an increase of 3 dB in current. In the advanced combination encoder (ACE) strategy, the loudness growth is a non-linear function that compensates for the increase in loudness that recipients experience during electrical stimulation. The full scale of the amount of energy possible in each channel is mapped to fit within the dynamic range of the recipient. An example of how a specific amount of energy in a channel is mapped to a value between the threshold- (T) and comfort (C)- levels is shown in Figure 4.4. The threshold-level is the current level corresponding to a current flow in a single channel that is just detectable to the recipient. Conversely, the comfort-level is the current level that corresponds to the recipient perceiving a sound of comfortable loudness.

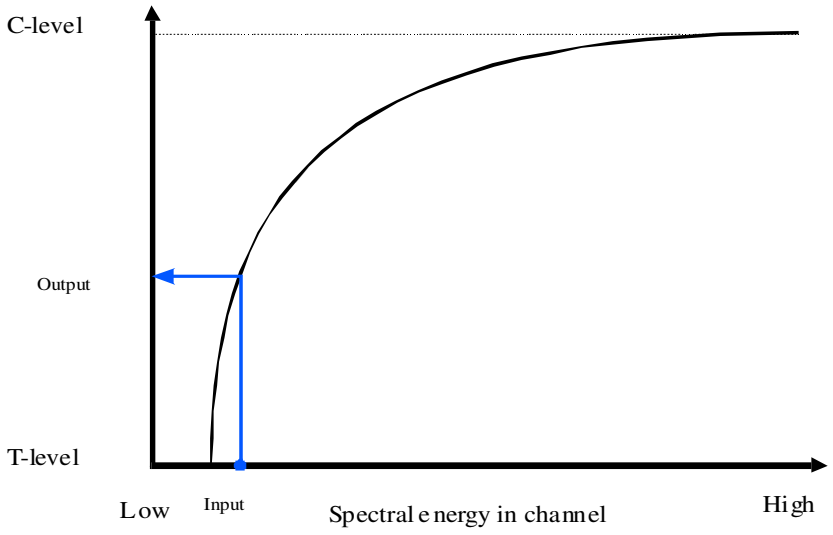

Figure 4.4 Loudness growth and current level mapping

- The final processing block completes the coding of stimulation instructions (e.g. active electrode, reference electrode, current level, and pulse width) by compiling the information received from the previous processing blocks and sending it to the speech processor. The speech processor uses this information to modulate the radio frequency transmission to the implant.

4.4 OVERVIEW OF CONVERSION FROM ADVANCED COMBINATION ENCODER (ACE) TO TRAVELLING WAVE ENCODING STRATEGY

To implement the hydrodynamic model into a speech coding strategy, some of the processing blocks of the advanced combination encoder (ACE) speech coding strategy were removed or modified to fit into the travelling wave encoding strategy (see Figure 4.5).

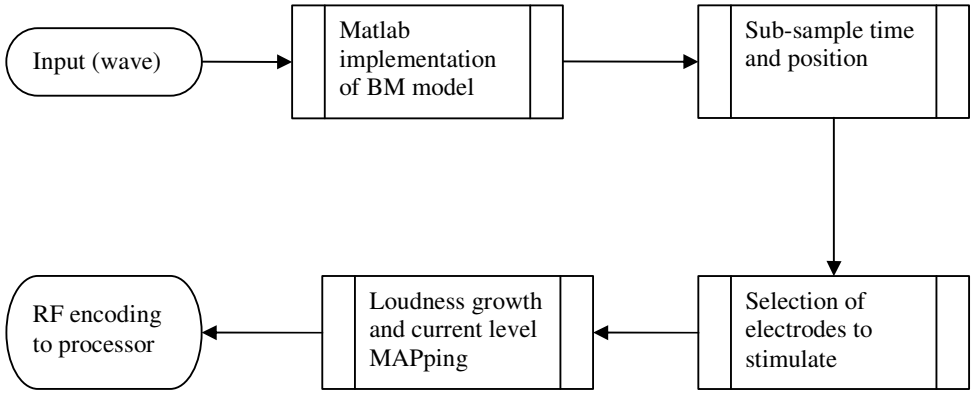

Figure 4.5 Processing steps in the traveling wave encoding strategy

- The Hanning window and fast Fourier transform processing block was replaced by the Matlab implementation of the hydrodynamic model described in Chapter 3 .
- The processing block that assigns fast Fourier transform outputs to specific channels was replaced by the sub-sampling of the model output in both time and space. It assigns each sampled displacement of the basilar membrane, to a specific channel.
- The existing maxima selection processing block was retained, specifying either fourteen or six maxima for per-channel stimulation rates of 1 000 pulses per second or 2 400 pulses per second respectively. This ensures stimulation on as many electrodes as possible within the stimulus rate restrictions of the system used. (Recall that the maximum total stimulation rate is 14 400 pulses per second.)

- The loudness growth and current level mapping processing block was modified to include an option for a linear loudness growth function. A linear loudness growth function ensures that results can directly be linked to the implemented model. Changing back to a non-linear function could possibly be investigated in future research.

4.5 IMPLEMENTATION OF THE TRAVELLING WAVE ENCODING STRATEGY IN SOFTWARE

The Matlab implementation of the above processing blocks in was performed stepwise (Figure 4.6) similar to the organization of processing blocks.

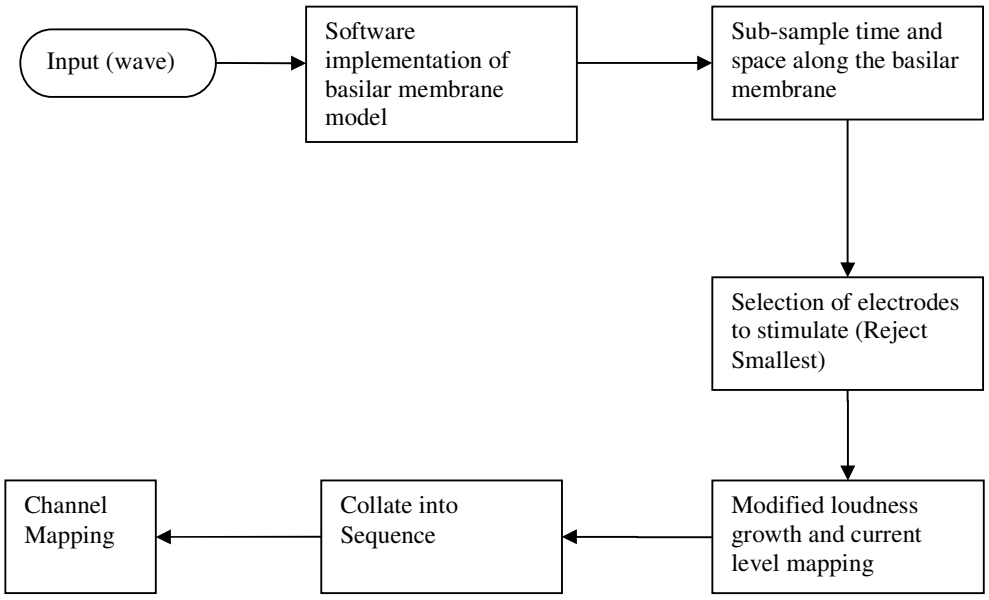

Figure 4.6 Processing of sound in Matlab.

The processing of sound progresses as follows.

4.5.1 Input

Normal wave-files, treated as a stereo input and sampled at 16 kHz, serve as input. Sampled analogue signals have sampling errors in both the time domain and in the amplitude of the signal. The wave-format in Matlab uses 32 bits per sample, with an amplitude sampling error of 2^{-32}, considered to be small enough for this application. The sampling error in the time domain restricts frequency analysis to at least half the sampling frequency, i.e. less than 8 kHz, which is also deemed more than sufficient for this application. The wave-audio standard is stereo sound represented in two column vectors. To enable the model to process generic wave-files, one column vector is discarded (if present) and only one of the mono-channels is processed.

4.5.2 Software implementation and sub-sampling

The travelling wave encoding strategy is the implementation of the hydrodynamic model of the basilar membrane displacement as discussed in Chapter 3 . The wave-type input is processed, followed by computation of basilar membrane displacement for each consecutive time step. The displacement of the basilar membrane is sampled linearly (in space) at 22 sites along the cochlea, since the implant system comprises 22 electrodes. The displacement is re-sampled to a per-channel stimulation rate of 1 000 pps, i.e. only one in sixteen calculations is used as output to the next computational block and only 22 of the 128 positions along the basilar membrane are used.

4.5.3 Selection of electrodes to stimulate

This procedure retains the values of the fourteen sites associated with the largest positive displacement along the basilar membrane. This restriction is used because of the Nucleus 24 cochlear implant system's maximum stimulation rate (14 400 pulses per second). An electrodogram, that displays the output intended for specific electrodes

(representing specific points on the basilar membrane) against time, is generated. (Figure 4.7). The vertical axis displays the position along the basilar membrane that corresponds to electrode position and time is displayed on the horizontal axis. The colour indicates increasing amplitude at the specific electrode, with blue and red indicating small and large amplitudes respectively.

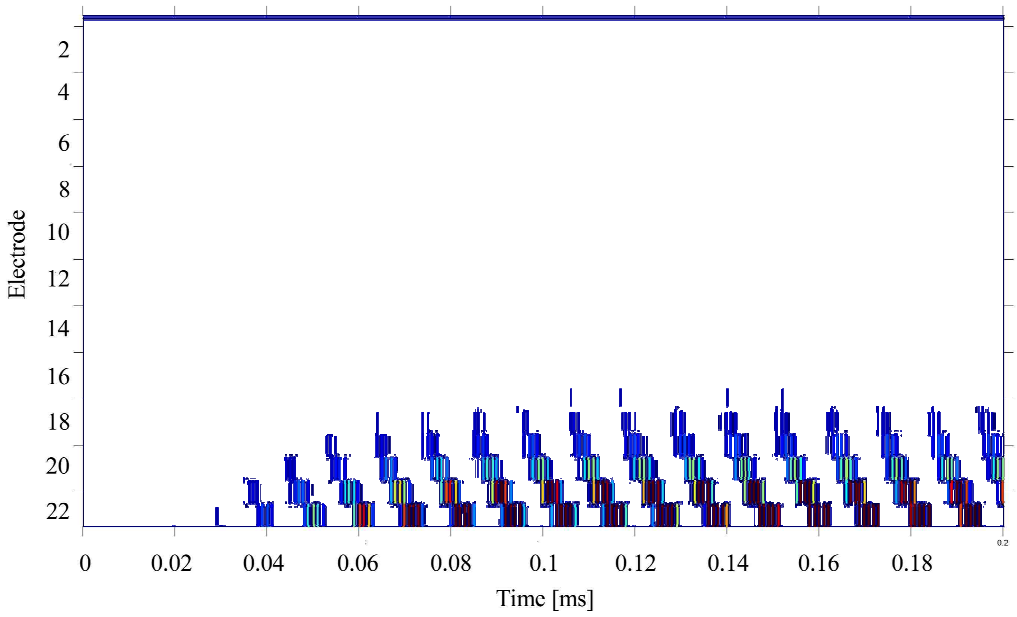

Figure 4.7 Example of electrodogram

4.5.4 Modified loudness growth function and current level mapping

This procedure scales the output of the travelling wave encoding strategy to a value between 0 and 1, corresponding to the range between the recipient's threshold- and comfort-levels (see Figure 4.8). When implementing this for the travelling wave encoding strategy, the curve becomes a straight line, assuming linear loudness growth (compare to Figure 4.4). This allows the model to be evaluated on its own, i.e. the loudness growth measured in the normal population when using the advanced combination encoder (ACE) speech coding strategy does not necessarily apply to the travelling wave encoding strategy. The output of the previous processing block was

normalised using a pure tone signal at maximum value as input and measuring the output. This output was mapped to the comfort-level, while all other inputs were converted to an appropriate level between the threshold- and comfort-levels using the linear transfer function as shown in Figure 4.8.

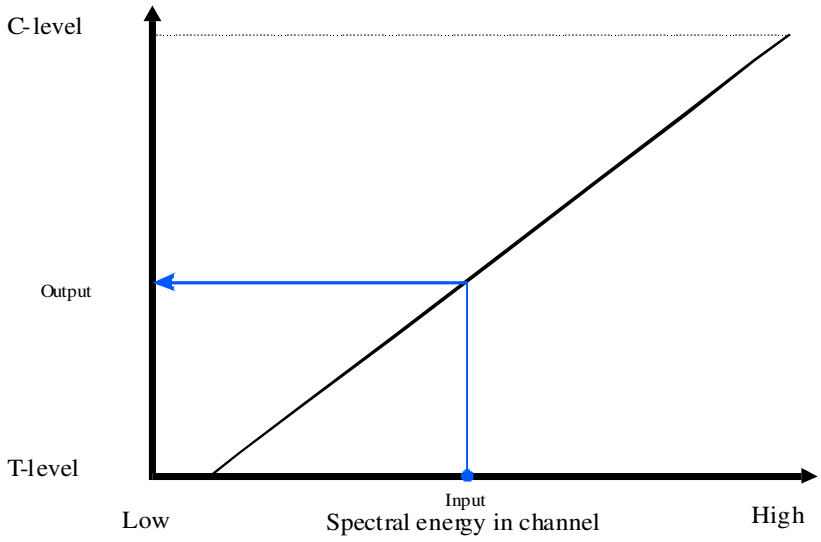

Figure 4.8 Loudness growth and current level mapping.

4.5.5 Collate into sequence

This procedure uses the output from the travelling wave encoding loudness growth function block and organizes it according to a sequential stimulation order for each sampling period, since the Nucleus system employs sequential stimulation rather than simultaneous stimulation. The stimulation sequence is by default from base to apex. It was decided to keep the defaults, to allow comparison with the advanced combination encoder (ACE) speech coding strategy. Channels not selected for stimulation are ignored so that only relevant information is coded to be sent through to the implant.

This is the last processing block that contains recipient independent data. Provided threshold- and comfort-levels are known, this output can be used, with some additional

parameters, to stimulate any recipient given the recipient's threshold- and comfort-levels. In order to speed up the experiments, all the sounds that were presented to the recipients were pre-processed up to this point. Output data of this processing block was stored for easy and quick use with any recipient.

4.5.6 Channel mapping

This procedure inserts patient-specific details, such as threshold- and comfort-levels, into the sequence. The output of the previous procedure created a sequence of stimuli mapped between a threshold-level of 0 and a comfort-level of 1 (see Figure 4.8). The threshold-level and range between threshold- and comfort-levels are used to scale the magnitudes of the sequence of the previous procedure to current levels. Apart from threshold- and comfort-levels, all recipients used the same parameters (see Table 4.1).

Table 4.1 Parameter values used for recipient stimulation

Parameter	Values
Mode of stimulation	MP1+2 (Both mono-polar electrodes are used together)
Per-channel stimulation rate	1 000 pps for the 100 Hz pure tones 2400 pps for the 1 kHz pure tones
Inter phase gap	8.0 μs
Number of electrodes used	22
Maximum number of electrodes stimulated per time sample	100 Hz pure tones: 14 1 kHz pure tones: 6
Pulse width	25 μs

4.6 DEBUGGING AND RX-FRAMES

Debugging in this application is of utmost importance, since it is possible to cause over stimulation to the recipient if stimulated at levels above the comfort-level of a specific channel. In the normal hearing ear, the tensor tympani that contracts on perception of loud sounds, thereby limiting perception of high sound levels. This contraction causes the ossicular chain to miss-align and so attenuate incoming sound. If incoming sound increases even further the tympanic membrane may rupture, causing an additional loss of about 40 dB. Since all of these systems are bypassed with a cochlear implant, there are no such safety mechanisms to prevent high stimuli to reach the central auditory system and it is possible, therefore to stimulate a cochlear implant recipient far beyond their pain threshold.

Since the cochlear implant manufacturer does not have full control over the final product created using the NIC software or Nucleus Matlab toolbox, the responsibility falls upon the software user to ensure that the recipient is not over-stimulated when presenting any stimulus by means of software using the above packages. In order to prevent over-stimulation a precautionary program (Rx-Frames, Cochlear Ltd.) allows one to capture and store the output to the speech processor connected to a PC, prior to actual stimulation of the recipient. This output data can then be compared to the intended stimuli to ensure all stimulation is below the comfort-levels of each channel.

In Table 4.2, an example of the output from Rx-Frames is shown. Each line represents a single stimulus, together with all its parameters. The first column shows the active electrode, with electrode 24 indicating the ball electrode used during non-stimulus periods to keep the implant powered while waiting for new information to be received. “Mode”

refers to the electrode(s) used as reference. A value of 30 indicates the monopolar 1+2 mode, where the two external reference electrodes are connected together to act as a common reference, while a value of 25 indicates the plate electrode as reference during the non-stimulus periods. The "current level" column shows the current level used for the stimulation, with zero indicating the lowest possible current used during the non-stimulus periods. This column's data are compared to the comfort-levels set for each active electrode, to prevent any stimulation being louder than this. Each pulse width is given separately and inter-phase gap is also shown. Constant rate stimulation is a constraint within the current software version, resulting in a constant period between pulses and requiring many non-stimulus frames to be used when no stimuli are needed.

Table 4.2 Rx-Frames stimulus data

Active Elec.	Mode	Current level	pulse 1 width	Inter-phase gap	Pulse 2 width	Period
14	30	2	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
5	30	11	24.8	8	24.8	70
6	30	78	24.8	8	24.8	70
9	30	100	24.8	8	24.8	70
12	30	66	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	8	24.8	70
24	25	0	24.8	7	24.8	70
5	30	30	24.8	8	24.8	70
6	30	34	24.8	8	24.8	70
8	30	100	24.8	8	24.8	70
10	30	94	24.8	8	24.8	70
12	30	100	24.8	8	25.2	70

RESULTS

The results obtained after software implementation can be analysed visually using electrodograms, as mentioned in Paragraph 4.5.3 or by looking at specific time samples to see how the stimulus amplitudes change with time. Such amplitude changes show the travelling wave moving towards the apex. The results from the travelling wave encoding strategy can also be compared visually to the results obtained with advanced combination encoder (ACE).

In Figure 4.9, the first 200 μs of an electrodogram obtained with a 90 Hz pure tone are shown. The lower numbered electrodes represent the basal part of the cochlea from where the travelling wave originates to travel towards electrode 22 in the apex of the cochlea. The bursts of stimulation coincide with the positive deflections of the basilar membrane and serve as an indication of how the travelling wave progresses along the basilar membrane. The arrow shows the direction of travel. If the travelling wave moved at a constant speed, the stimulation bursts would have continued to start at the arrow line, as is the case with electrodes 17 and 18. The fact that the bursts lag in time as it approaches the point of maximal deflection agrees with the increasing phase delay as the travelling wave approaches the point of maximal deflection. The bursts are also spaced apart in time with a period of 11 ms. This perceived modulation of the stimulus train on the electrodes should in itself also convey pitch information to the recipient (McDermott & McKay, 1997). The lack of stimulation during the first 40 ms is due to the smooth ramp-up of the pure tone to prevent possible transient effects.

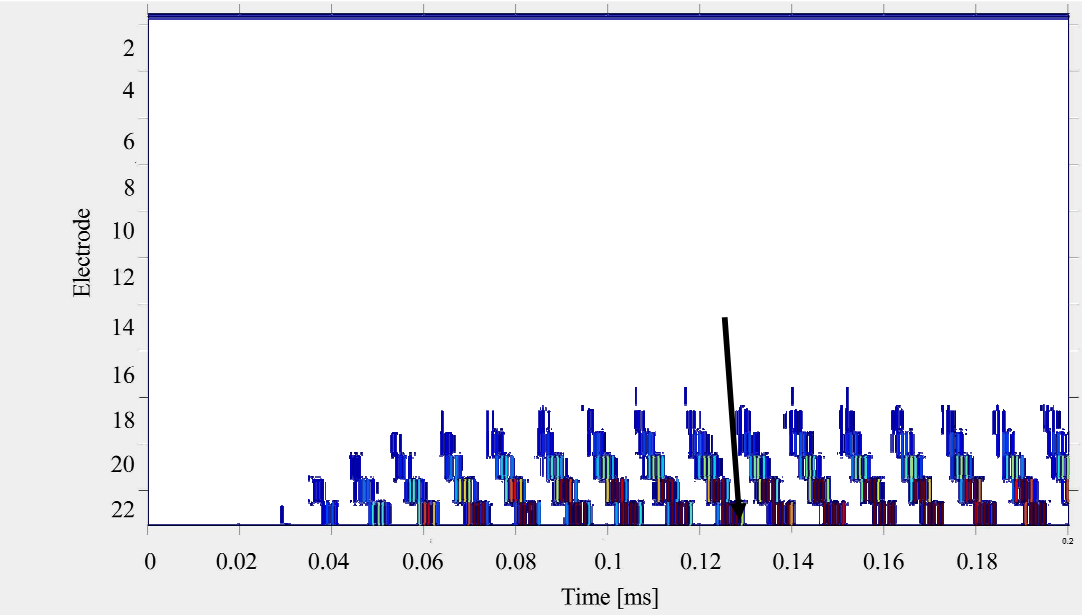

Figure 4.9 Electrodogram for 90 Hz pure tone processed with the travelling wave encoding strategy

The electrodogram of a 90 Hz pure tone processed with the advanced combination encoder (ACE) strategy is shown in Figure 4.10. Comparing Figures 4.9 and 4.10 highlights the difference between advanced combination encoder (ACE) and the travelling wave encoding strategy. Only two electrodes are stimulated in advanced combination encoder (ACE), dictated by the band-pass filters used (through the Fast Fourier Transform calculations as described in Paragraph 4.3) and is stimulated on two channels due to the overlap in the filter bands (see Figure 1.). The stimulation amplitude of 125 Hz on electrode 22 is higher than on electrode 21, as shown in Figure 4.10. Following the smooth ramp-up of the stimuli, the stimulation is essentially constant both in amplitude and in electrodes, relying fully on the position of stimulation to convey the pitch of the pure tone.

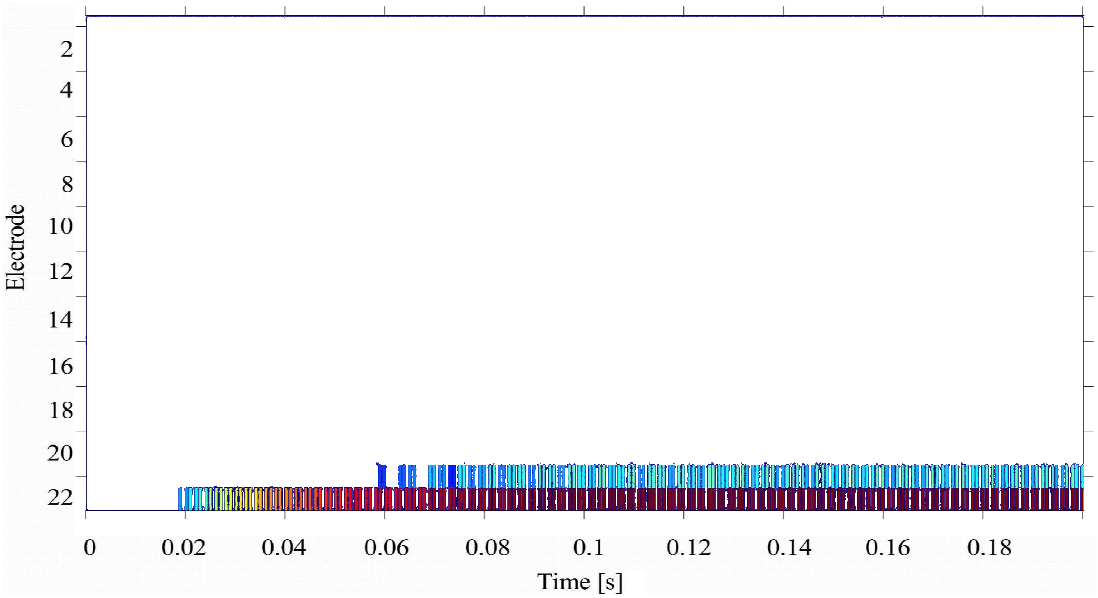

Figure 4.10 Electrodogram of a 90 Hz tone through advanced combination encoder (ACE)

In order to visually evaluate the between two adjacent frequencies, the difference between two such frequencies was displayed as an electrodogram. The difference between a pure tone of 90 Hz and 91 Hz respectively, after it was processed using the advanced combination encoder (ACE) speech processing strategy, is displayed in Figure 4.11. The only visible difference is a slight relative increase (light blue is for low amplitudes) of amplitude on electrode 18. A recipient would only be able to discriminate between these two frequencies if such a small difference can be detected.

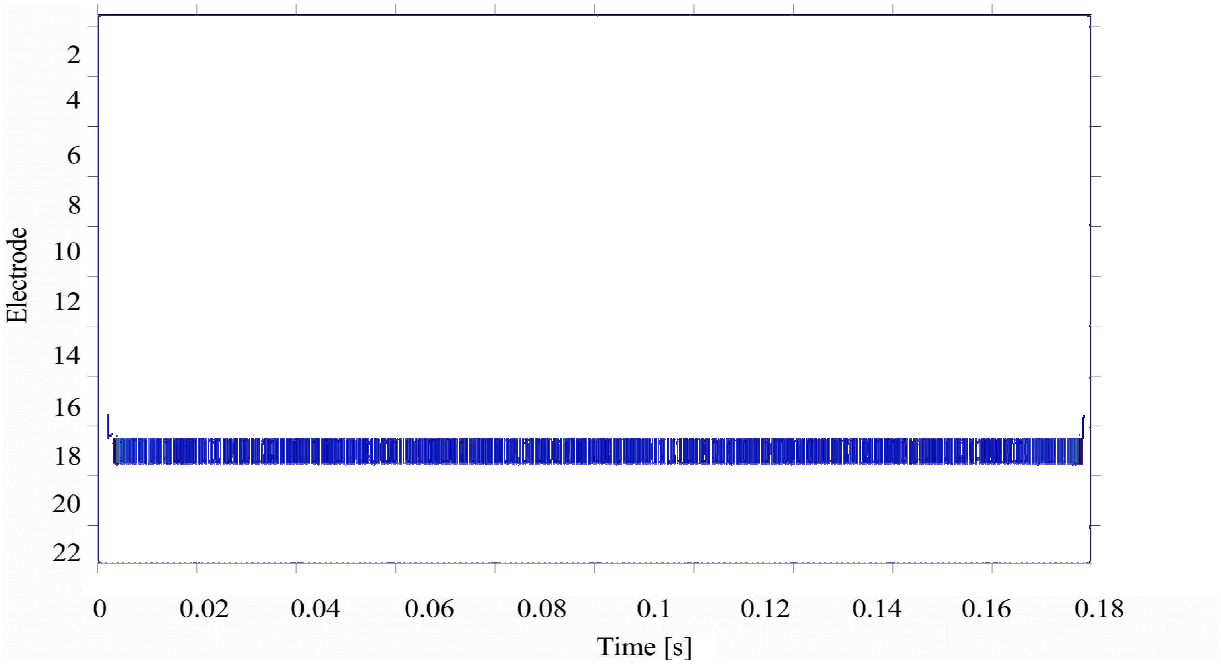

Figure 4.11 Electrodogram of the difference between 90 Hz and 91 Hz using ACE.

When comparing the electrodograms of a 90 Hz and a 91 Hz pure tone, processed with the travelling wave encoding strategy, a much bigger difference can be seen (Figure 4.12). Not only are larger amplitudes observed (colours approaching the red portion of the spectrum), but differences are also seen on electrodes 16 to 22. Some small differences in some of the lower electrodes are also observed.

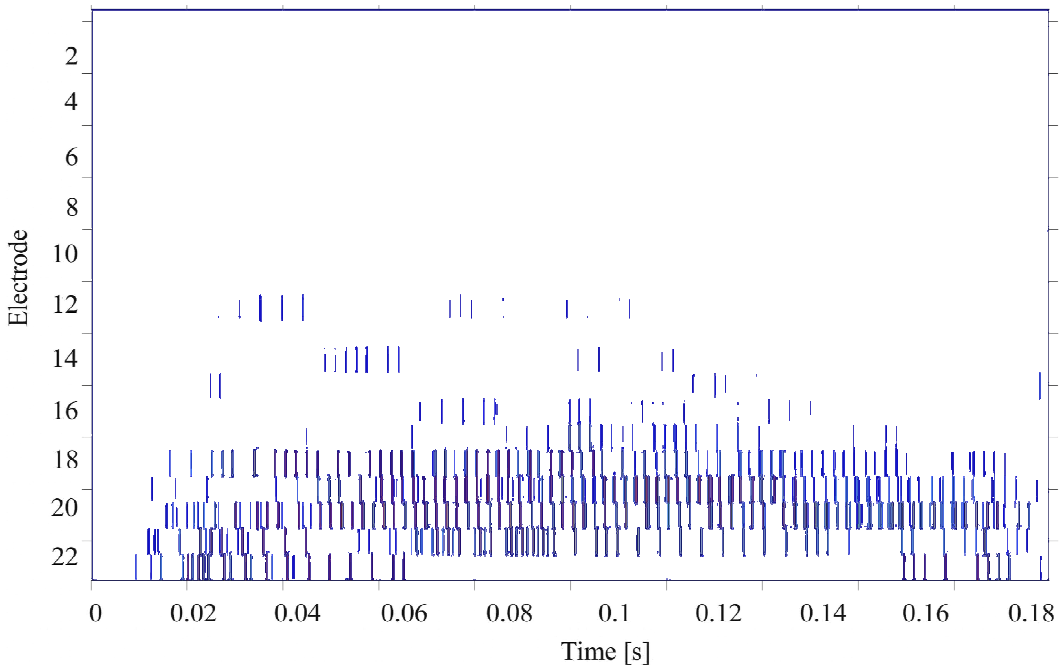

Figure 4.12 Electrodogram of the difference between a 90 Hz sound and a 91 Hz sound using the travelling wave encoding strategy.

If the analysed sound does not comply with Nyquest's criterion of being less than half the sub-sampling frequency, aliasing occurs even though the original sampling of 16 kHz was sufficiently fast. Such an electrodogram is shown in Figure 4.13. This is the reason for using two stimulation rates for the 100 Hz and 1 kHz sounds, 1 kHz and 2.4 kHz respectively as mentioned in Paragraph 4.2. An electrodogram when sampling at 2.4 kHz of a 1 kHz pure tone is shown in Figure 4.14.

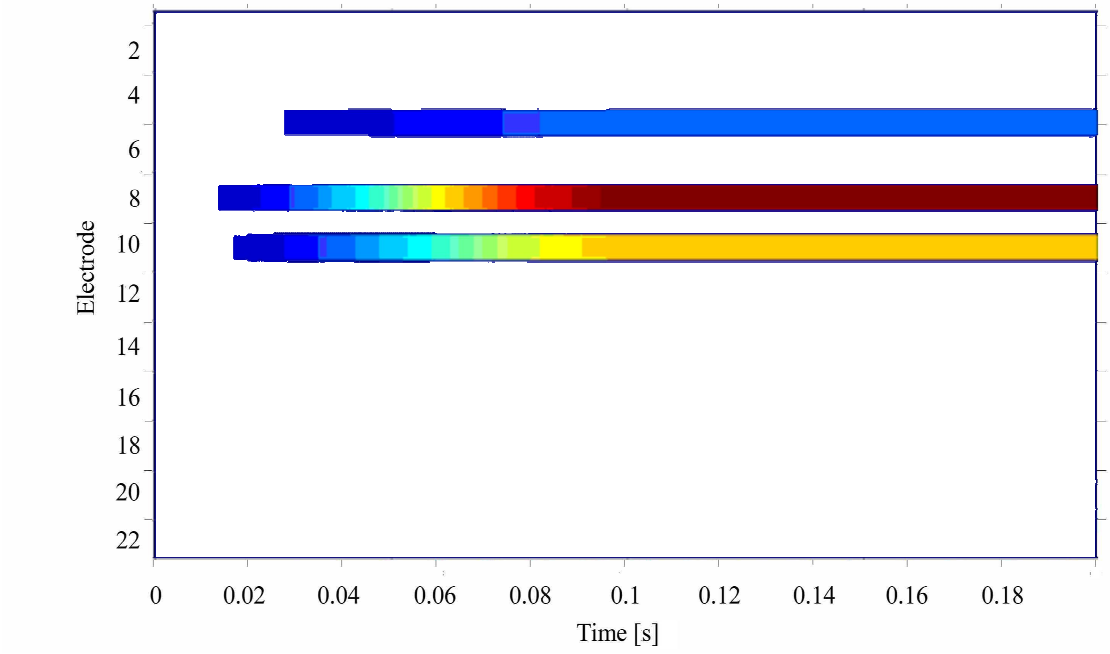

Figure 4.13 Electrodogram of a 1 kHz pure tone, sub-sampled at 1 kHz

An electrodogram when sampling at 2.4 kHz of a 1 kHz pure tone is shown in Figure 4.14.

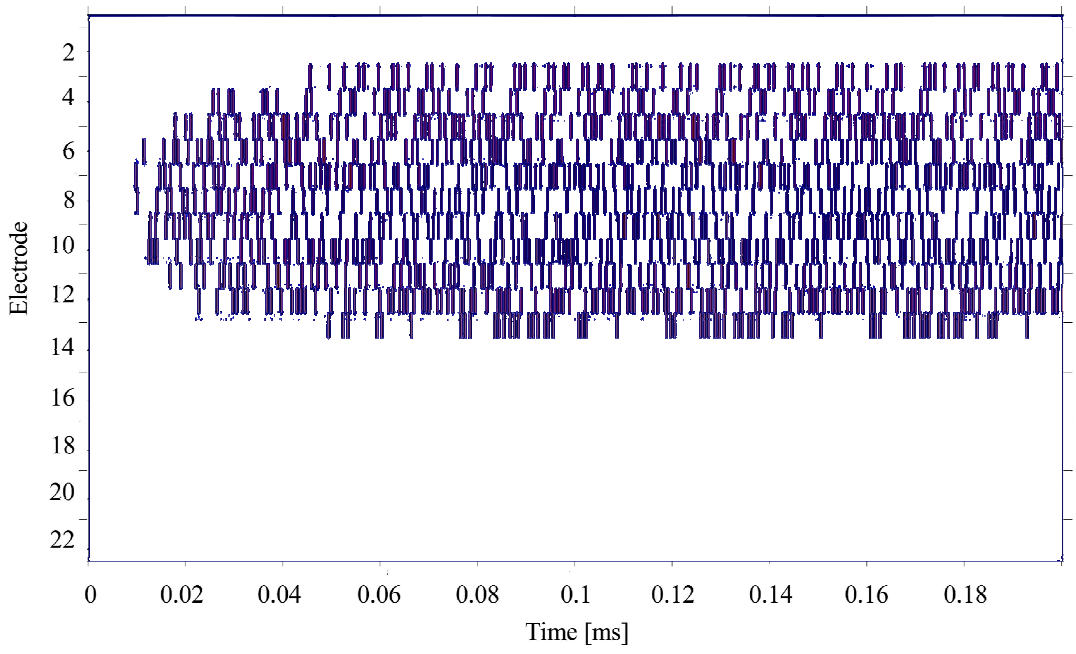

Figure 4.14 Electrodogram of a 1kHz pure tone, sub-sampled at 2.4 kHz

4.7 SUMMARY

This chapter focussed on the integration of the hydrodynamic model into the Nucleus Implant Communicator (NIC) toolbox in Matlab. This toolbox allowed stimuli produced using the travelling wave encoding strategy to be compared to stimuli produced using a current, commercial strategy (advanced combination encoder (ACE)). The NIC library and Matlab toolbox were discussed and the specific changes to the implementation of the model to allow integration with this toolbox were also mentioned. De-bugging and controls to eliminate the possibility of over-stimulation were also discussed. The advanced combination encoder (ACE) strategy was used as a baseline and only the modules that impacted the travelling wave encoding strategy were changed, so as to compare the two strategies as closely as possible. When the output from the two strategies were compared visually, it was concluded that the travelling wave encoding strategy appears to provide more information beneficial to frequency discriminating frequencies that lie close together than the advanced combination encoder (ACE) speech coding strategy. Following these results, it is therefore expected that recipients will have better pitch discrimination when presented with stimuli processed with the travelling wave encoding strategy.

Chapter 5 EXPERIMENTS

5.1 INTRODUCTION

This chapter discuss the experimental procedure that was followed. It mentions the case profiles of the recipients that participated in the study and will focus specifically on electrodograms that show the stimulus pattern and levels of stimulation. The experimental setup of each experiment will be discussed and results will be shown.

5.2 OBJECTIVES

Experiments with the travelling wave encoding strategy aimed to evaluate the potential improvement in pitch discrimination compared to commercial strategies, as perceived by cochlear implant recipients. The hypothesis was that coding sound according to the way described in preceding chapters, should provide a cochlear implant recipient with improved pitch discrimination ability, compared to current, commercial speech coding strategies. Since pitch ranking might be an indicator of speech perception potential (Hanekom & Shannon, 1996), it warrants evaluating the extent to which alternative sound processing strategies (in this case the travelling wave encoding strategy) influence pitch ranking. As discussed in the preceding chapters, the travelling wave encoding strategy seems to convey more information about pure tones than the advanced combination encoder (ACE) strategy and may therefore aid recipients in discriminating smaller pitch differences.

Experiments were designed to assess the pitch discrimination ability of recipients with both a commercial strategy (advanced combination encoder (ACE)) and the travelling wave encoding strategy, applying the hydrodynamic model of the basilar membrane displacement. For this purpose, 21 pure tones around 100 Hz with 1 Hz intervals (90 Hz – 110 Hz) and 21 pure tones around 1 kHz with 10 Hz intervals (0.90 kHz – 1.10 kHz) were

pre-processed (signal processing as described Chapter 4) with the Nucleus Matlab Toolbox for stimulation with both advanced combination encoder (ACE) and the travelling wave encoding strategy. These frequency intervals were chosen following visual inspection of the results, as discussed in Chapter 4. If perfect (100%) discrimination was observed, the frequency resolution would have been increased, while too difficult tasks would have resulted in decreased frequency resolution. Frequency resolution at 1 kHz was decreased in anticipation of poorer pitch perception at higher frequencies, as is observed with normal hearing individuals. Prior to assessing the pitch discrimination results, it is important to eliminate other potential differences, such as loudness and duration of the stimuli. The latter is obtained by using identical inputs, leading to stimuli that is the same length.

5.3 EXPERIMENTAL PROCEDURE

The experimental procedure can be divided into three separate procedures, namely loudness balancing, pitch discrimination and pitch ranking. These will be discussed separately.

5.3.1 Loudness balancing procedure

The first part of each experiment was a loudness balancing procedure to ensure that, as far as possible, all the stimuli would be perceived at equal loudness. An adaptive procedure was used to accomplish this. Two stimuli were presented and the recipient had to indicate which of the two sounded louder.

The adaptive procedure is based on a standard zero-seek procedure that approaches the current level at which the test signal will be perceived as equally loud to the reference signal. If the recipient indicated, for example, the reference stimulus to be louder than the test stimulus, the test stimulus' level of stimulation will be increased until the test stimulus is perceived as being louder than the reference stimulus. At this point the

direction of the steps will be reversed to decrease the test stimulus loudness. After each reversal, the step size is halved and after the third reversal (in this application) the two sounds were considered to be balanced. Figure 5.1 shows the step size decreases following each reversal. The initial step size was 5% of the tested value, decreasing to 2.5% and 1.25% for the second and third reversal respectively. A typical step size of 1.25% would (for comfort-levels of between 180 and 200 current levels) be two to three current levels. This was accepted as adequate, since the recipients struggled to use smaller step sizes. The centre of the 21 stimuli, for 100 Hz and 1 kHz tone respectively, was selected as the reference with which all the other stimuli were compared. This ensured loudness balancing across all stimuli, i.e. one constant reference against which to compare and adjust all the stimuli.

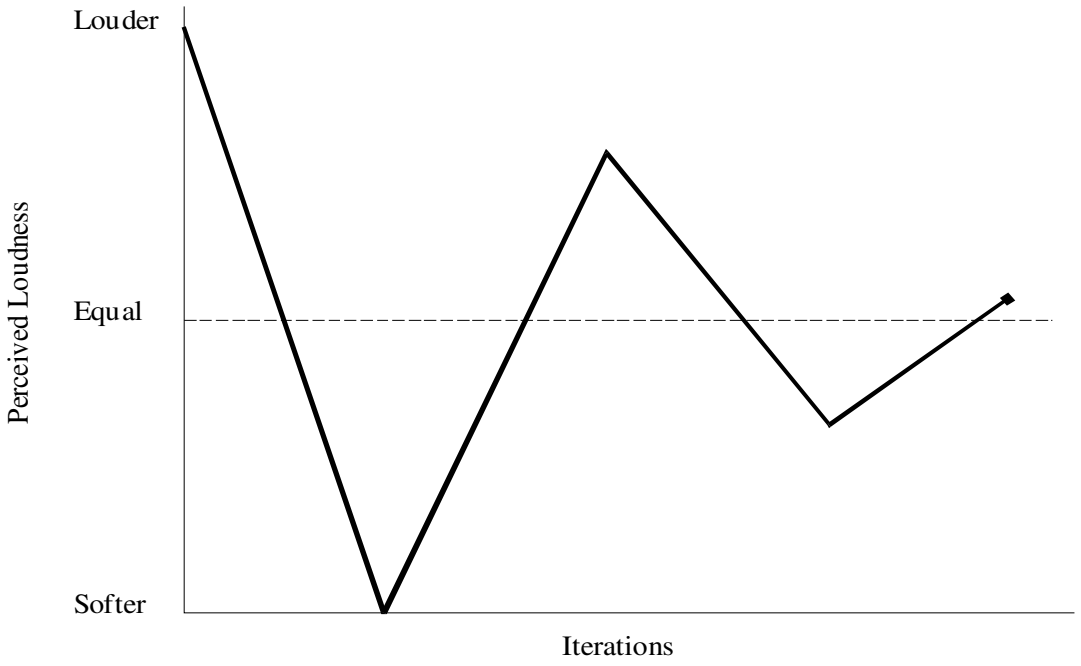

Figure 5.1 Adaptive loudness balancing procedure

5.3.2 Stimulus discrimination experiment

Since the objective is to compare the ability of the travelling wave encoding strategy to convey pitch information to that of commercially available strategies, the stimulus discrimination experiment was conducted as follows. Discrimination was evaluated with a three-interval forced-choice experiment. Three stimuli were presented, with two being the same and one being different. The recipient was asked to identify the 'odd one out'. This experiment did not attempt to connect a perceived pitch to the stimuli, but only to assess if the stimuli could be distinguished from one another.

Each of the 21 sounds was compared to all the others, resulting in 231 comparisons. The compared pairs that were compared with each other were randomly ordered and the 'odd one out' was randomly presented in one of the three intervals. Each recipient completed six repetitions of this experiment for each set of test stimuli. This was deemed to be sufficient, since the patterns that emerged have already stabilised after six repetitions. Only correct answers were counted and normalised to 1 by dividing the number of correct answers by the number of repetitions. A score of 0.0 indicates no discrimination, 0.3 indicates random discrimination and 1.0 represent perfect discrimination. These results can be displayed in a confusion matrix, as shown in Table 5.1. The reference stimuli are represented along the x-axis, while the test stimuli are represented along the y-axis. For example, in row 1 and column 3, sound 1 and 3 were correctly discriminated from each other in three of the six tries (i.e. 0.5). Along the diagonal of this example matrix, each stimulus was compared to itself, preventing any positive discrimination. As the 'distance' between two compared stimuli increases (upward and to the right), the probability of discriminating the two also increases.

Table 5.1 Confusion matrix - example scenario

Stim	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1	1	1	1	1	1	1	1
4	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1	1	1	1	1	1	1
5	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1	1	1	1	1	1
6	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1	1	1	1	1
7	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1	1	1	1
8	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1	1	1
9	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1	1
10	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1	1
11	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1	1
12	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1	1
13	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1	1
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8	1
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	0.8
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.7
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3

If the whole upper triangle approximates a random response, all the stimuli, even the two furthest apart, are impossible to discriminate between and the experiment might need to be adjusted so that presented stimuli are further apart. Conversely, if the only stimuli that are not perfectly identified are the ones where the stimulus is compared to itself, the stimuli may need to be presented at pitches that are closer together.

5.3.3 Pitch ranking experiment

Once the stimulus discrimination experiment is completed, it would be worthwhile to establish whether the different stimuli processed with the travelling wave encoding strategy have tonal qualities that distinguish them or whether another characteristic causes the perceptual difference. A two-interval forced-choice experiment was set up to perform

pitch-ranking on all 231 stimuli combinations. A recipient was presented with two stimuli at a time and was asked to choose the one that had the lower pitch. Compared pairs were randomly ordered and the one with a lower frequency was randomly presented in either the first or the second interval. Only correct answers were counted and normalised to by dividing the number of correct answers by the number of repetitions. A score of 0.0 showed 100% certainty on pitch reversal, i.e. the stimulus derived from a higher frequency had a perception of being lower in pitch. A score of 0.5 showed no pitch discrimination, while a score of 1.0 showed 100% certainty of correct pitch, i.e. the stimulus derived from the lower frequency had a perception of being lower in pitch. The results can also be displayed in a matrix form, similar to those of the discrimination experiment. An example of a pitch ranking matrix would look similar to Table 5.1. Stimuli that are far apart would be easy to pitch rank correctly, but as the stimuli become more similar, the score is expected to decrease towards 0.5 – with no difference in tonal quality on the diagonal.

In line with the hypothesis, it is expected that the pitch discriminaiton using the travelling wave encoding strategy would be superior to pitch ranking when using current commercial speech coding strategies. For the same input signals, the confusion matrix obtained with the travelling wave encoding strategy is therefore expected to contain a larger area that is correctly discriminated than one obtained from current commercial strategies.

5.4 BACKGROUND OF RECIPIENTS

The details of the recipients that took part in the experiments are given below. They are part of a group of recipients that regularly assist with experiments at the department.

Table 5.2 Recipient information.

Recipient ID	*S1*	*S9*
Age	*55*	*54*
Etiology	*Progressive loss*	*Progressive loss*
Implanted in tested ear	*6 months*	*3 years*
Implant	*Nucleus 24 Contour Advance*	*Nucleus 24 Contour Advance*
Speech Processor	*ESPrit 3G*	*ESPrit 3G*
Coding strategy	*ACE*	*ACE*
Channel stimulation rate	*900 / 500 Hz*	*900 Hz*
Number of maxima	*10*	*10*

During the experiments, a loaner SPrint speech processor was used, to ensure that no changes could be made to the recipient's own speech processor. Each recipient's details contained in their speech processors were used to obtain threshold and comfort levels (see paragraph 4.3 for an explanation of these levels) necessary for the implementation of the speech coding strategies.

5.5 ELECTRODOGRAMS

Electrodograms are three-dimensional plots, shown in two dimensions, depicting a specific electrode position on the y-axis against time on the x-axis. Each block represents one stimulus pulse, of which the colour depicts the amplitude of stimulation (the third dimension) on the specific electrode at a specific time. When comparing the electrodograms of tones around 100 Hz, processed with the advanced combination encoder (ACE) strategy (see Figures 5.2 and 5.3) almost no change is visible between the two electrodograms in Figures 5.2 and 5.3 even though the stimuli are 10 Hz apart. This was expected since the filter bands at the low frequencies have a maximum resolution of 125

Hz (see Figure 1.4). The only difference between 100 Hz and 90 Hz would be the amplitude of the output of the second filter.

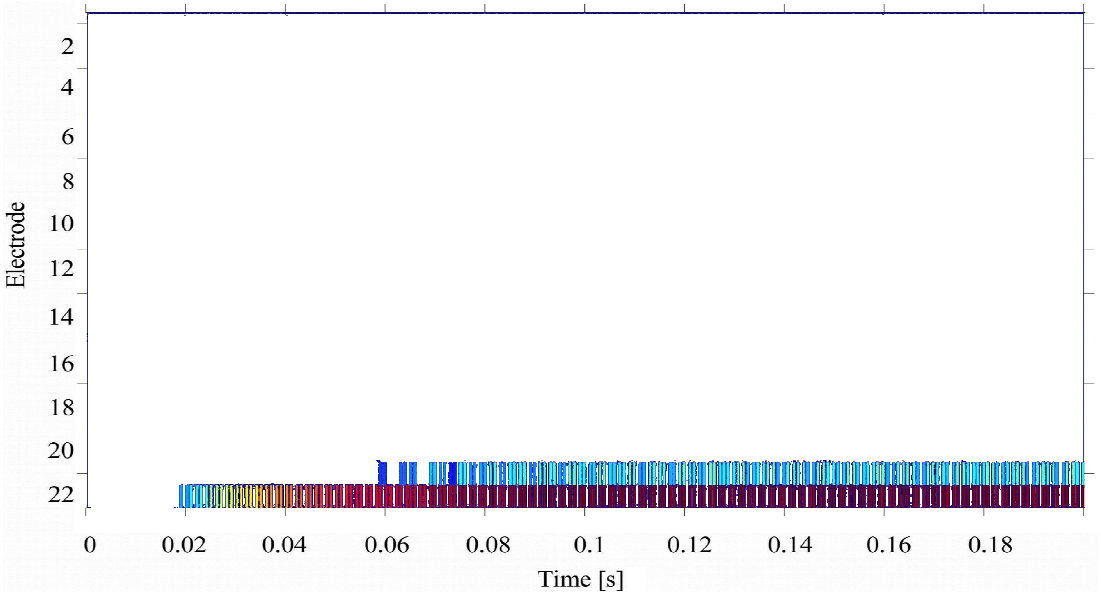

Figure 5.2 Electrodogram of a 90 Hz tone processed with the advanced combination encoder (ACE) speech processing strategy. Two electrodes are activated, with an amplitude difference.

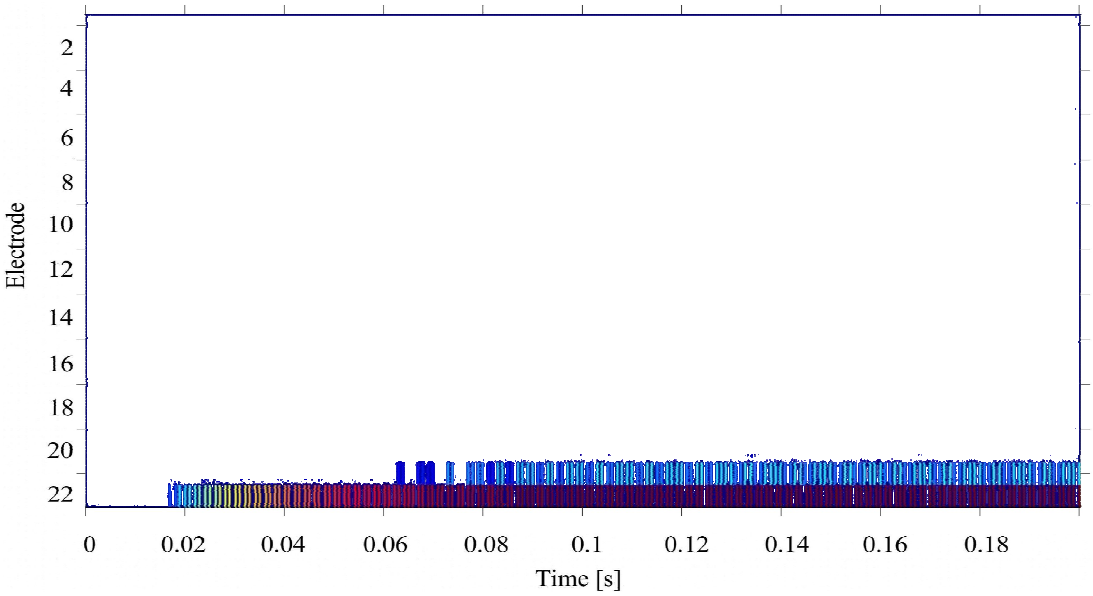

Figure 5.3 Electrodogram of a 100 Hz tone processed with the advanced combination encoder (ACE) speech processing strategy. Virtually identical to the 90 Hz electrodogram.

If the same two tones are processed through the travelling wave encoding strategy, the visual discrimination is much easier (see Figure 5.4 and 5.5). Temporal information

(tracking of the 100 Hz tone through the periods of negative basilar membrane deflection) and the spatial information (the progression of the travelling wave towards the point of maximal deflection) allow the stimuli to be distinguished easily. The slowing down of the travelling wave can be seen in Figure 5.4 as a lag of the onset of each burst of stimulation relative to the yellow arrow (straight line for constant velocity) as the travelling wave approaches the apex (electrode 22).

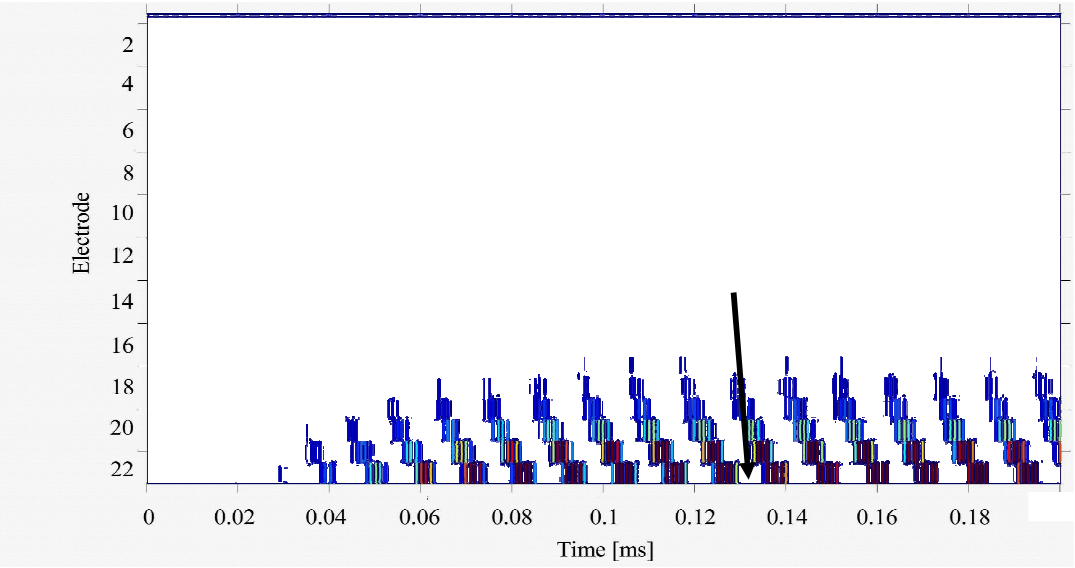

Figure 5.4 Electrodogram of a 90 Hz tone processed with the travelling wave encoding strategy. The travelling wave can be seen moving from the base (low electrode numbers) to the apex (high electrode numbers). The arrow shows constant speed, i.e. the slowing down of the travelling wave can be seen.

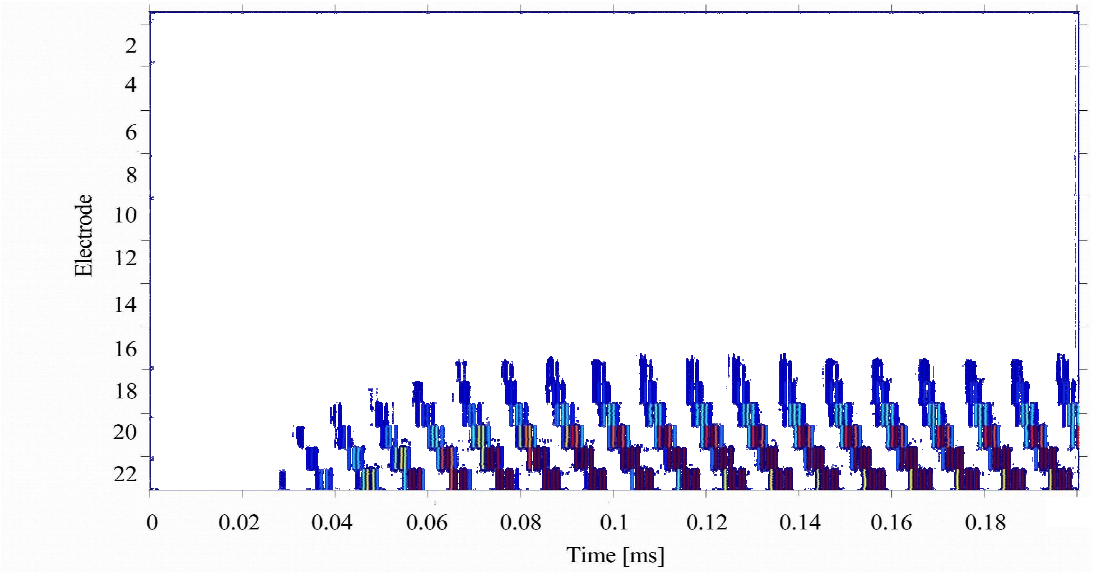

Figure 5.5 Electrodogram of a 100 Hz tone processed with the travelling wave encoding strategy. Visual discrimination between 90 Hz and 100 Hz is possible due to an increased number of activated electrodes and reduced period between pulses, i.e. one extra pulse in 0.1 ms.

The ease with which frequencies close to each other can be visually distinguished from the electrodograms generated by the travelling wave encoding strategy, compared to those generated by advanced combination encoder (ACE), suggests that cochlear implant recipients will also be able to distinguish between these sounds when using the travelling wave encoding strategy.

The sensation of the 100 Hz stimulus should also be perceived as higher pitched when compared to the 90 Hz stimulus due to the increased stimulation on electrodes closer to the base of the cochlea that is normally associated with higher pitch tones (Lesser & Berkley 1971). The recipient should also perceive a tonal quality from the periodic peaks of the travelling wave (McKay & McDermott 1996). The effect that the slowing down of the travelling wave has on pitch percept as it approaches the point of maximal deflection has, as far as is known not been assessed in the literature.

5.6 RESULTS

The results of the three experiments are shown below. The two recipients' results will be reported individually.

5.6.1 Loudness balancing

As can be seen in Figure 5.6, the result of the loudness balancing procedure showed an average change from the initial comfort-level values used ranging from 0.8% to 3.8% across all the processed pure tones and standard deviations ranging from 0.7% to 8.6%.

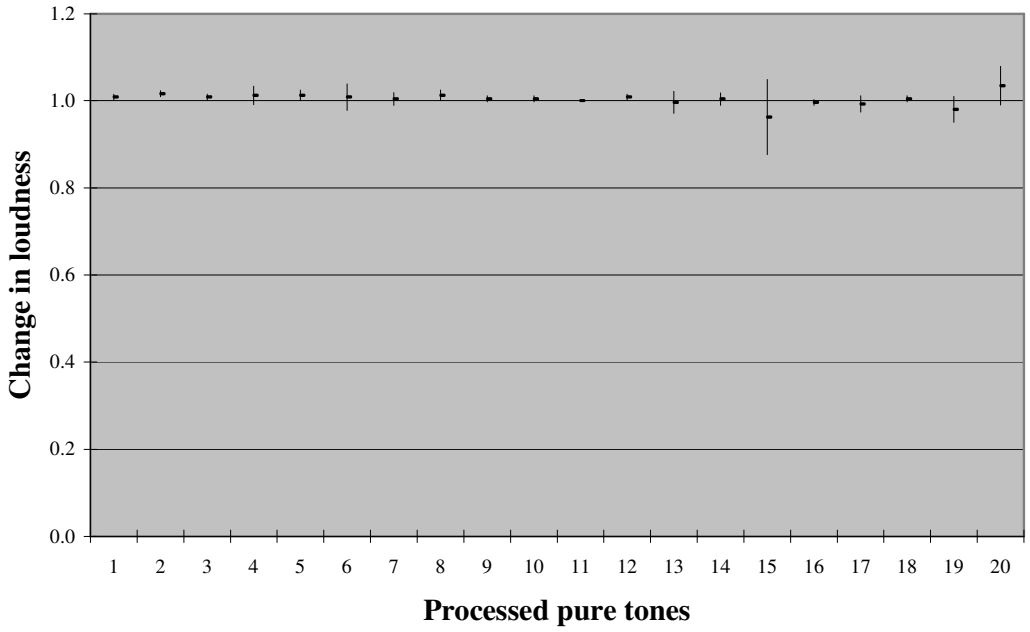

Figure 5.6 An example of loudness balance outcome for S1, showing small deviations from original values. The horizontal axis shows each of the 21 stimuli, derived from 21 different pure tones. The vertical axis shows how the loudness of each stimulus was adjusted in loudness to be perceived as 'equally loud' in comparison to the reference sound (either 100 Hz or 1 kHz – pure tone 11).

The results of the loudness balancing suggests that all the sounds are very similar in loudness and also want to suggest that loudness differences did not influence the discrimination experiments.

5.6.2 Discrimination experiment

The discrimination experiment results are shown in confusion matrices below. In Table 5.3, the discrimination confusion matrix of S1 is shown for the advanced combination encoder (ACE) strategy, after six repetitions of the discrimination experiment were completed. The matrix show the normalised outcome of an experiment with both the columns and rows representing stimuli derived from pure tones, as indicated. Due to the three-interval-forced-choice nature of the experiment, 0.3 indicates random discrimination and 1.0 a constantly correct discrimination. To quantify the data presented in the confusion matrices, a discrimination index D is introduced. D is the mean of each sound's mean discrimination score, i.e. the column and the row for 99 Hz is averaged to obtain the mean discrimination score for 99 Hz with each of the 21 sounds. All 21 sounds' means are averaged to obtain D.

Compared to Tables 5.1 and 5.4 there appears to be no discrimination of stimuli in Table 5.3, D is close to the 0.33 random discrimination level. Even the stimuli furthest apart (stimuli derived from 90 Hz and 110 Hz respectively) could not reliably be distinguished from each other.

Table 5.3 S1 advanced combination encoder (ACE) 100Hz Discrimination Confusion Matrix. D= 0.36 ± 0.04

Hz	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110
90	0.3	0.2	0.5	0.3	0.2	0.2	0.5	0.3	0.5	0.5	0	0.5	0.5	0.2	0.3	0.2	0.3	0.2	0.7	0.2	0.5
91	0	0.2	0	0.7	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.2	0.3	0.3	0.7	0.3	0.3	0.2	0.5	0.7	0.3
92	0	0	0.3	0.3	0.5	0.2	0.3	0.2	0.7	0.2	0.5	0.3	0.5	0.5	0.3	0.5	0.5	0.2	0.5	0.3	0.2
93	0	0	0	0.5	0.2	0.3	0.5	0.3	0.3	0.3	0.5	0.2	0.7	0.3	0.5	0.5	0.2	0.3	0.5	0.7	0.2
94	0	0	0	0	0.2	0.7	0.5	0.3	0	0.7	0.2	0.3	0.5	0.8	0.2	0.3	0.7	0.3	0.3	0.7	0.2
95	0	0	0	0	0	0.3	0.3	0.2	0.2	0.3	0.5	0.2	0.3	0.5	0.2	0.5	0.5	0.3	0.5	0.5	0.7
96	0	0	0	0	0	0	0.2	0.3	0.2	0.3	0.2	0.5	0.3	0.7	0	0.7	0.2	0.3	0.7	0.3	0.5
97	0	0	0	0	0	0	0	0.2	1	0.3	0.2	0.3	0.5	0.3	0.3	0	0.3	0.7	0.3	0.3	0.2
98	0	0	0	0	0	0	0	0	0	0.3	0.5	0.3	0.2	0.2	0.3	0.3	0.3	0.3	0.2	0.5	0
99	0	0	0	0	0	0	0	0	0	0.5	0.3	0.2	0	0.7	0.3	0	0.3	0.3	0.5	0.3	0.2
100	0	0	0	0	0	0	0	0	0	0	0.7	0.5	0.3	0	0.3	0	0	0	0.3	0.3	0
101	0	0	0	0	0	0	0	0	0	0	0	0.2	0.3	0.3	0.2	0.7	0.5	0.3	0.5	0.7	0
102	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.7	0	0.2	0.3	0.7	0.3	0.3	0.2
103	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.3	0.5	0.5	0.3	0.7	0.3	0.7
104	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0.7	0.7	0.3	0.3	0.7	0.5
105	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.3	0.2	0.5	0.3	0.2
106	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.2	0.2	0.2	0.7
107	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0.5	0.3	0.5
108	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.7
109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0
110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3

Table 5.4 shows the higher level of discrimination obtained by S1 (also using six repetitions) with the travelling wave encoding strategy, with D = 0.8 indicating an average of 80% correct discrimination of the pure tones. This equates to an improvement of 40% on discrimination when using the travelling wave encoding strategy. The stimuli used to obtain both the results in Tables 5.3 and 5.4 were computed from identical input signals, as mentioned in Chapter 4.

Table 5.4 S1Travelling wave encoding strategy 100Hz Discrimination Matrix. D= 0.77 ± 0.05

Hz	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110
90	0.5	1	0.3	1	1	1	1	1	1	1	1	0.8	1	1	1	0.8	1	1	1	1	1
91	0	0.3	0.3	0.3	0.7	0.7	0.8	0.7	0.7	0.8	1	1	1	1	1	1	1	1	1	1	0.8
92	0	0	0.3	0.5	0.2	0.7	1	0.5	0.8	1	0.8	1	1	1	0.8	1	1	1	1	1	1
93	0	0	0	0.3	0.7	0.8	0.5	0.8	1	0.8	0.8	1	1	1	1	0.8	1	1	1	1	1
94	0	0	0	0	0.5	0.3	0.7	0.3	0.5	0.8	1	1	1	0.7	1	1	1	1	1	1	0.5
95	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	1	0.8	1	1	1	0.8	0.8	1	1	1
96	0	0	0	0	0	0	0.5	0.3	0.5	0.2	0.7	1	1	1	1	1	1	1	0.8	1	0.5
97	0	0	0	0	0	0	0	0.2	0.5	0.5	0.8	0.8	0.8	0.8	1	1	1	1	1	1	0.2
98	0	0	0	0	0	0	0	0	0.5	0.2	0.8	0.8	1	0.8	0.8	1	1	1	1	1	0
99	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	1	1	1	1	0.8	1	1	1	0.8
100	0	0	0	0	0	0	0	0	0	0	0.5	0.8	0.7	0.8	0.7	1	1	0.7	1	0.8	0.8
101	0	0	0	0	0	0	0	0	0	0	0	0.7	0.7	0.7	0.3	0.3	0.8	0.5	0.3	0.8	0.8
102	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.7	0.5	0.8	0.7	0.3	0.5	0.8	0.8
103	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.3	0.3	0	0.7	0.7	0.5	1
104	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0.2	0.5	0.2	0.3	0.5	1
105	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.5	0.5	0.2	0.5	1
106	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.7	0.5	0.5	1
107	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.3	1
108	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0	1
109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	1
110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3

When compared to Table 5.3, Table 5.5 shows better results obtained from S9 for the same experiment (although only four repetitions were used) when using the advanced combination encoder (ACE) strategy generated stimuli for sounds around 100 Hz, with D = 0.6. Visually, it does not seem to form such a distinct upper-triangle like Table 7.6.

Table 5.5 S9 advanced combination encoder (ACE) 100 Hz Discrimination matrix. D= 0.62 ± 0.06

Hz	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110
90	0.8	0.3	0	0.8	1	1	0.3	0.3	0.3	0.5	1	0.8	1	1	0.3	1	0.8	1	0.8	1	1
91	0	0.8	0	0.3	1	0.8	1	0.3	0	0.5	0.5	1	0.8	1	0.8	0.8	0.8	0.5	0.8	0.8	0.8
92	0	0	0.8	0.3	0.8	0.8	0.8	0.5	1	0.5	1	0.8	1	1	0.3	0.3	1	0.5	0.8	0.5	0.5
93	0	0	0	0.5	0.5	0.5	0.5	0.5	0.8	0.8	0.8	1	0.5	0.8	0.8	0.8	0.5	0.5	0.8	0.5	0.3
94	0	0	0	0	0.3	0.5	1	1	0.8	1	0.5	0.5	0.3	0.8	0.8	1	0.8	0.8	0.8	1	0.5
95	0	0	0	0	0	0	0.8	1	0.8	1	0.8	0.5	0.5	0.5	0.8	0.5	0.8	0.5	0.5	0.3	0
96	0	0	0	0	0	0	0.5	0.3	0.8	0.5	1	1	1	0.8	0.8	1	1	1	1	1	1
97	0	0	0	0	0	0	0	0.5	0.8	0.3	1	1	0.5	0.8	0.3	0.8	1	0.3	1	0.3	1
98	0	0	0	0	0	0	0	0	0	0	1	0.5	1	1	0.5	0.3	0.5	0.5	0.8	0.5	1
99	0	0	0	0	0	0	0	0	0	0.8	1	1	0.5	1	0	0.5	0.8	0.5	0.3	0.5	0.8
100	0	0	0	0	0	0	0	0	0	0	0.5	0.3	0.3	0.3	0.5	0.8	0.3	0.5	0.3	0.5	0.5
101	0	0	0	0	0	0	0	0	0	0	0	0.3	0.3	0.3	0.3	0.5	0.3	0.8	0.3	0.8	0.3
102	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0	0.8	1	0.3	0.5	0.8	0	0.3
103	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	1	0.8	0	0.8	1	0.8	0.5
104	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0.5	0.8	0.3	0.8	0.8	1
105	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0.3	0.5	0.8	0.5
106	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.3	0.5	0.5
107	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.5	0.8
108	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.8	0.8
109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5
110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Like Table 5.4, Table 5.6 also shows a higher level of discrimination when compared to Table 5.5 for S9, with D = 0.7. S9 improved her discrimination by 10% when using the travelling wave encoding strategy and the upper-triangle in the matrix looks more pronounced.

Table 5.6 S9Travelling wave encoding strategy 100 Hz Discrimination Matrix. D= 0.67 ± 0.04

Hz	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110
90	0.6	0.1	0.4	0.4	0.6	0.6	0.6	0.7	0.9	1	1	0.7	1	0.7	1	1	1	1	1	1	0.7
91	0	0.1	0.3	0.6	0.1	0.4	0.4	0.6	0.4	0.9	1	0..9	0.9	1	0.9	1	1	1	1	0.7	0.6
92	0	0	0.1	0.4	0.3	0.3	0.4	0.7	0.3	0.7	0.9	1	1	1	1	1	0.9	1	1	0.9	1
93	0	0	0	0.1	0.3	0	0.9	0.7	0.9	0.4	0.7	0.6	1	1	1	1	1	1	1	1	0.9
94	0	0	0	0	0.6	0.3	0.6	0.4	1	1	0.9	0.7	1	1	1	1	1	0.9	0.9	0.7	0.9
95	0	0	0	0	0	0.6	0.3	0.1	0.9	0.7	0.7	0.7	0.9	1	1	1	1	0.9	1	1	1
96	0	0	0	0	0	0	0.3	0.3	0.3	0.6	0.7	0.6	0.9	1	1	0.9	1	0.7	0.9	1	0.9
97	0	0	0	0	0	0	0	0.1	0.1	0.3	0.7	1	1	0.9	1	0.7	0.7	0.9	1	1	1
98	0	0	0	0	0	0	0	0	0.1	0.4	0.1	0.7	1	0.7	0.6	0.7	1	1	0.6	1	0.4
99	0	0	0	0	0	0	0	0	0	0.4	0.6	0.6	0.4	0.7	0.9	1	0.6	0.7	0.9	1	0.3
100	0	0	0	0	0	0	0	0	0	0	0.4	0.6	0.4	0.7	0.7	0.4	0.9	0.7	0.7	0.9	0.4
101	0	0	0	0	0	0	0	0	0	0	0	0.4	0.1	0.1	0.7	0.4	0.7	0.4	1	0.3	0.6
102	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.3	0.4	0.3	0.3	0.3	1	0.4	0.4
103	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9	0.1	0.4	0.7	0.4	0.7	0
104	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.6	0	0.4	0.6	0.6	0.7
105	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.1	0.4	0.3	0.9
106	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.1	0.4	0.7	0.6
107	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0	0.6	0.9
108	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.7
109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	1
110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Discrimination for the higher frequency pure tones were almost identical for S1 when comparing the advance combination encoder (ACE) strategy and the travelling wave encoding strategy (see Table 5.7 and Table 5.8), with D = 0.8 in both instances. Both experiments used six repetitions. The discrimination compares well with the discrimination at 100 Hz for S1, although the frequency steps are 10 Hz apart. The travelling wave encoding strategy does not seem to improve discrimination for S1 at frequencies around 1 kHz.

Table 5.7 S1 advanced combination encoder (ACE) 1 kHz Discrimination matrix. D= 0.77 ± 0.05

kHz	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.90	0.3	0.7	0.5	0.7	0.3	0.5	0.5	1	1	0.8	1	1	1	1	0.8	1	1	1	1	1	0.8
0.91	0	0.5	0.3	0.5	0.7	0.2	0.7	0.7	0.5	0.8	1	0.8	1	1	1	1	1	1	1	1	0.8
0.92	0	0	0	0.7	0.2	0.5	0.7	0.7	0.8	1	1	1	1	1	0.8	1	1	1	1	1	1
0.93	0	0	0	0.5	0.3	0.7	0.7	0.7	0.7	0.8	1	0.7	0.8	1	1	1	1	1	1	1	0.7
0.94	0	0	0	0	0.3	0.5	0.2	0.7	0.3	0.8	1	1	1	1	1	1	1	0.8	0.8	1	1
0.95	0	0	0	0	0	0.3	0.5	0.5	0.7	0.7	0.8	1	0.8	0.8	1	1	1	1	1	0.8	0.7
0.96	0	0	0	0	0	0	0.5	0.3	0.3	0.7	0.5	0.7	1	1	1	1	1	1	0.8	1	0.2
0.97	0	0	0	0	0	0	0	0	0.5	0.3	0.8	1	1	1	1	0.8	1	1	1	1	0.5
0.98	0	0	0	0	0	0	0	0	0	0.8	0.5	1	1	1	1	1	1	1	1	0.8	0.2
0.99	0	0	0	0	0	0	0	0	0	0	0.5	0.8	0.7	0.8	0.8	1	1	1	1	1	0.7
1.00	0	0	0	0	0	0	0	0	0	0	0.7	0	0.8	0.7	0.7	0.8	1	1	1	1	0.2
1.01	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.2	0.7	0.7	1	1	1	1	0.7
1.02	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.8	0.5	0.7	0.8	1	1	1	1
1.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.3	0.7	0.7	0.8	0.8	1	0.8
1.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.2	0.5	0.8	0.7	0.8	1
1.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0.5	0.7	0.7	1
1.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.7	0.2	0.3	0.8
1.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.5	0.5	1
1.08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.3	0.8
1.09	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	1
1.10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5

Table 5.8 S1Travelling wave encoding strategy 1 kHz Discrimination matrix. D= 0.80 ± 0.08

kHz	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.90	0.2	1	0.5	1	1	1	0.8	1	1	0.8	1	1	1	1	1	1	1	1	1	1	1
0.91	0	0.2	0.7	0.8	0.8	0.8	0.8	0.8	1	0.8	1	1	1	1	1	1	0.8	1	1	1	1
0.92	0	0	0.2	0.8	0.5	0.2	1	0.5	1	1	0.8	0.8	0.7	1	1	1	0.8	1	1	1	1
0.93	0	0	0	0.2	0.5	0.2	0.3	0.7	1	0.5	0.8	0.7	0.8	1	0.8	1	1	1	1	1	0.5
0.94	0	0	0	0	0.2	0.2	0.5	0.5	1	0.7	0.2	0.7	0.8	1	1	0.8	1	1	0.8	0.8	0.7
0.95	0	0	0	0	0	0.2	0.3	0.3	1	0.5	0.7	0.5	0.7	0.8	1	1	0.8	1	1	1	0.8
0.96	0	0	0	0	0	0	0.2	0.2	1	0	0.5	1	0.3	1	1	1	0.8	0.8	0.8	1	0.7
0.97	0	0	0	0	0	0	0	0.5	0.8	0.8	0.3	0.7	0.7	1	1	0.7	0.8	1	1	1	0.3
0.98	0	0	0	0	0	0	0	0	0.2	1	0.8	0.7	1	1	1	1	1	1	1	1	0.8
0.99	0	0	0	0	0	0	0	0	0	0.5	0.5	0.7	0.8	1	1	0.7	1	1	1	1	0.8
1.00	0	0	0	0	0	0	0	0	0	0	0.3	0.3	0.7	1	0.8	1	0.8	0.8	1	1	0.7
1.01	0	0	0	0	0	0	0	0	0	0	0	0.8	0.8	0.8	1	0.7	0.7	1	1	1	0.7
1.02	0	0	0	0	0	0	0	0	0	0	0	0	0.5	1	0.7	0.8	0.5	0.7	1	0.7	0.3
1.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.7	0.8	1	0.8	1	1	0.8
1.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.3	0.2	0.7	0.8	1	0.8
1.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0.7	0.5	0.8	1	0.8
1.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	1	0.8	1
1.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.5	0.8	1
1.08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.8	0.8
1.09	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1.10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Comparing Table 5.9 and Table 5.10, the travelling wave encoding strategy seems to give worse discrimination for S9 with D = 0.4 when compared to the advance combination

encoder (ACE) strategy, with D = 0.8 (Although only three repetitions were obtained for Table 5.9 and six for Table 5.10) The discrimination score for Table 5.10 approaches the 0.33 random-score.

Table 5.9 S9 advance combination encoder (ACE) 1 kHz Discrimination matrix. D= 0.76 ± 0.05

kHz	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.90	0.3	0.7	0	0	0	0.3	0.7	1	1	1	1	1	1	1	0.7	0.7	1	1	1	1	1
0.91	0	0.7	0	0.7	0.7	0.7	0.3	1	1	1	1	1	1	1	1	1	1	1	1	1	0.7
0.92	0	0	0.7	0	0.3	1	0.7	1	1	1	1	1	1	1	1	1	1	1	1	1	1
0.93	0	0	0	0.3	0.3	0.3	0.7	0.7	1	1	1	1	1	1	1	1	1	1	1	1	1
0.94	0	0	0	0	0.3	0.3	1	1	1	1	1	1	1	0.7	1	1	1	1	1	1	1
0.95	0	0	0	0	0	0	0.7	0.3	1	1	1	1	1	1	0.7	1	1	1	1	1	1
0.96	0	0	0	0	0	0	0.3	0	0.3	0.7	1	1	0.7	1	1	1	1	1	1	1	0.7
0.97	0	0	0	0	0	0	0	0.3	0.3	0.7	0.7	0.7	0.7	1	1	1	1	0.7	1	1	0.7
0.98	0	0	0	0	0	0	0	0	0.3	0	0.3	0.7	0	0	0.7	1	1	1	1	1	0
0.99	0	0	0	0	0	0	0	0	0	0.3	0	0.7	0.7	0.7	1	1	1	0.7	1	1	0.3
1.00	0	0	0	0	0	0	0	0	0	0	0.3	0	0.7	0.3	0.7	1	1	1	0.7	1	0.7
1.01	0	0	0	0	0	0	0	0	0	0	0	0.3	0.3	0.3	0.7	1	1	1	1	1	0.3
1.02	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0.7	0.7	0.7	1	1	1	0.3
1.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.7	0.7	0.7	1	0.7	1	1
1.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0	0	1	0.3	0.3	1
1.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.7	0.7	0.7	0.7
1.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.7	0.7	0.3	1
1.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.7	0.3	1
1.08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.7	0.3
1.09	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1.10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 5.10 S9 Travelling wave encoding strategy 1 kHz Discrimination matrix. D= 0.42 ± 0.05

kHz	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.90	0.3	0.6	0.9	0.6	0.3	0.3	0.3	0.6	0.4	0.6	0.4	0.6	0.4	0.4	0.4	0.4	0.4	0.3	0.4	0.4	0.4
0.91	0	0.6	0.4	0.4	0.6	0.4	0.3	0.4	0.4	0.3	0.4	0.4	0.4	0.4	0.3	0.4	0.4	0.4	0.4	0.3	0.4
0.92	0	0	0.4	0.4	0.4	0.3	0.3	0.6	0.1	0.6	0.4	0.4	0.4	0.4	0.4	0.4	0.3	0.4	0.4	0.3	0.1
0.93	0	0	0	0.3	0.6	0.6	0.3	0.1	0.7	0.9	0.4	0.1	0.3	0.3	0.3	0.3	0.6	0.4	0.3	0.3	0.3
0.94	0	0	0	0	0.3	0.1	0.6	0.7	0.6	0.3	0.4	0.3	0.1	0.6	0.3	0.1	0.4	0.4	0.4	0.4	0.6
0.95	0	0	0	0	0	0.6	0.7	0.9	0.4	0.9	0.7	0.9	0.3	0.6	0.1	0	0.4	0.4	0.4	0.4	0.6
0.96	0	0	0	0	0	0	0.4	0.7	0.7	0.9	0.1	0.3	0.4	0.6	0.3	0.1	0.3	0.3	0.4	0.4	0.6
0.97	0	0	0	0	0	0	0	0.6	0.4	0.4	0.1	0.4	0.3	0.4	0.3	0.4	0.4	0.4	0.3	0.3	0.6
0.98	0	0	0	0	0	0	0	0	0.6	0.4	0.1	0.6	0.3	0.4	0.3	0.3	0.3	0.4	0.4	0.4	0.3
0.99	0	0	0	0	0	0	0	0	0	0.3	0.6	0.3	0.4	0.4	0.4	0.4	0.1	0.6	0.3	0.3	0.1
1.00	0	0	0	0	0	0	0	0	0	0	0	0.6	0.4	0.6	0.3	0.4	0.3	0.3	0.4	0.4	0.7
1.01	0	0	0	0	0	0	0	0	0	0	0	0.7	0.4	0.6	0.4	0.3	0.1	0.4	0.4	0.4	0.4
1.02	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.9	0.4	0.4	0.4	0.4	0.3	0.4	0.7
1.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.4	0.6	0.3	0.4	0.4	0.4	0.9
1.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6	0.3	0.1	0.4	0.4	0.4	0.7
1.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.4	0.4	0.1	0.4	0.9
1.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.4	0.3	0.4	0.9
1.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.1	0.4	1
1.08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.4	0.9
1.09	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.9
1.10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3

Comparing the results from the 100 Hz stimuli with the results from the 1 kHz stimuli it can be noted that the travelling wave encoding strategy seems to give improved discrimination over the advanced combination encoder (ACE) strategy when processing low frequency sounds for both S1 and S9, with a more pronounced benefit seen in S1. S1 showed an improvement from D = 0.4 to D = 0.8 and S9 showed an improvement from D = 0.6 to D = 0.7.

When higher frequency sounds are processed the travelling wave encoding strategy seems to give equal or worse discrimination ability to the recipient when compared to the advanced combination encoder (ACE) strategy. S1 obtained D = 0.8 using both strategies and S9 obtained D = 0.8 using advanced combination encoder (ACE) and D = 0.4 when using travelling wave encoding strategy. These results are summarised in Table 5.11.

Table 5.11 Discrimination indices for both 100 Hz and 1 kHz for S1 and S9 using six repetitions of 231 comparisons.

Frequency of stimuli	Recipient	D for advanced combination encoder (ACE)	D for travelling wave encoding
100 Hz	S1	0.36 ± 0.04	0.77 ± 0.05
	S9	0.62 ± 0.06 *	0.67 ± 0.04
1 kHz	S1	0.77 ± 0.05	0.80 ± 0.08
	S9	0.76 ± 0.05 **	0.42 ± 0.05

* Only four repetitions used

** Only three repetitions used

With a sample size of two recipients, proper statistical analysis could not be done on the experimental data, although some trends was noticed. The travelling wave encoding strategy seems to improve the discrimination of 21; 1 Hz separated pure tones around 100 Hz to about 80% correct on average across all 21 sounds with about 10% variance across the sounds. This suggests that there are some additional information included in the travelling wave encoding strategy that allows users to discriminate pure tones around 100 Hz more accurately.

The travelling wave encoding strategy does not seem to improve discrimination of pure tones around 1 kHz (or might even decrease discrimination, as with S9). This suggests that the additional information presented to the recipient for pure tones around 100 Hz is lost at the higher frequency of 1 kHz. These results will be further discussed in Chapter 6.

5.6.3 Pitch ranking experiment

Stimuli presented for pitch ranking experiments were processed only with the travelling wave encoding strategy, since the nature of stimulation is different from current strategies and the improved discrimination shown in Table 5.4 and Table 5.6 (when compared to Tables 5.3 and 5.5) does not necessarily carry pitch information. Since the advanced combination encoder (ACE) strategy rely on the tonotopicity of the cochlea alone and good pitch ranking shown in literature (Hanekom & Shannon 1996), pitch ranking experiments was not conducted with the advance combination encoder (ACE) strategy.

The pitch ranking experiments were also reported in Tables, similar to the discrimination experiments, with values as discussed in paragraph 5.3.2. D is computed in the same way as for the discrimination experiments (see paragraph 5.6.2) and indicates the accuracy with which the recipient can rank the pitch of the various sounds. The results of the pitch ranking experiment, using the 100 Hz stimuli, are shown in Tables 5.11 and 5.12 for S1 and S9 respectively. Both shows good pitch ranking with D = 0.8 or 80 % correct pitch ranking.

The pitch ranking ability looks similar to the discrimination found in Tables 5.4 and 5.6 (i.e. the sounds that was correctly discriminated was also correctly pitch-ranked) and indicates that both recipients could correctly pitch-rank different stimuli when pure tones processed with the travelling wave encoding strategy differed by 2 – 3 Hz. An interesting exception is the repeatable pitch reversal (values below 0.5) that was found when comparing 110 Hz with tones from 101 Hz to 109 Hz in both recipients.

Table 5.12 S1Travelling wave encoding strategy 100 Hz Pitch Ranking matrix. D= 0.78 ± 0.08

Hz	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110
90	0.2	0.5	0.5	0.8	0.7	0.8	1	1	0.7	1	1	1	1	1	1	1	1	1	1	1	1
91	0	0.5	0.5	0.5	0.3	0.8	0.8	0.7	1	1	0.8	1	1	1	1	1	1	1	1	1	1
92	0	0	0	0.2	0.5	0.3	0.7	0.8	0.7	1	1	0.8	1	1	1	1	1	1	1	1	1
93	0	0	0	0.5	0.5	0.7	0.7	0.5	1	0.7	1	1	1	1	1	1	1	1	1	1	1
94	0	0	0	0	0.7	0.5	0.8	0.5	0.3	0.8	1	1	1	1	1	1	1	0.8	1	1	0.8
95	0	0	0	0	0	0.3	0.8	0.8	0.5	0.8	0.7	0.8	1	1	1	0.8	1	1	1	1	0.7
96	0	0	0	0	0	0	0.5	0.7	0.7	0.7	0.7	1	1	1	1	1	1	1	1	1	0.8
97	0	0	0	0	0	0	0	0.5	0.8	0.7	0.8	0.8	1	1	1	1	1	1	1	1	1
98	0	0	0	0	0	0	0	0	0.5	0.7	0.8	0.7	0.8	1	1	1	1	0.8	1	1	0.5
99	0	0	0	0	0	0	0	0	0	0.8	0.7	0.7	1	0.8	1	1	1	1	1	1	1
100	0	0	0	0	0	0	0	0	0	0	0.2	0.8	1	1	0.8	1	0.8	1	0.7	0.8	0.7
101	0	0	0	0	0	0	0	0	0	0	0	0.3	0.2	0.8	0.5	1	0.7	0.7	0.8	0.5	0
102	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.3	0.7	0.5	0.8	0.7	0.8	0.2
103	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0.3	0.7	0.8	0.7	0.5	0
104	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.7	0.8	0.7	0.3	0
105	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0.8	0.3	0.7	0.3	0
106	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.5	0.3	0.8	0
107	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.7	0.3	0.2
108	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	0.3	0.2
109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0
110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7

Table 5.13 S9 Travelling wave encoding strategy 100 Hz Pitch Ranking matrix. D= 0.78 ± 0.07

Hz	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110
90	0.3	0.4	0.9	0.7	0.4	0.4	0.6	0.9	1	0.7	1	1	1	0.9	1	0.9	1	1	1	0.9	1
91	0	0.4	0.9	0.6	0.6	0.9	1	0.4	0.7	1	1	1	1	0.9	1	1	1	1	1	1	1
92	0	0	0.7	0.3	0.4	0.4	0.9	0.4	0.4	1	1	1	1	1	0.9	0.9	1	1	1	1	0.9
93	0	0	0	0.7	0.6	0.7	0.7	0.9	0.7	0.9	0.6	1	1	1	1	1	1	1	1	1	1
94	0	0	0	0	0.3	0.6	0.3	0.7	1	0.9	1	0.9	1	1	1	1	1	0.7	0.9	1	1
95	0	0	0	0	0	0.4	0.6	0.9	0.6	0.9	0.9	1	1	1	1	0.9	1	1	1	1	1
96	0	0	0	0	0	0	0.6	0.6	0.9	1	1	0.9	0.7	1	1	0.9	1	0.9	1	1	0.9
97	0	0	0	0	0	0	0	0.3	0.9	0.7	0.9	1	0.9	1	0.9	1	0.9	1	1	1	1
98	0	0	0	0	0	0	0	0	0.7	0.9	0.7	0.7	1	1	1	1	1	1	1	0.9	0.7
99	0	0	0	0	0	0	0	0	0	0.3	0.3	0.6	0.9	0.6	0.9	1	0.9	1	0.9	1	0.6
100	0	0	0	0	0	0	0	0	0	0	0.4	0.4	0.7	0.9	1	0.9	0.7	1	1	1	0.6
101	0	0	0	0	0	0	0	0	0	0	0	0.3	0.7	0.7	0.7	1	1	0.9	0.9	1	0.1
102	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0.7	0.7	0.9	0.9	0.3	1	0.9	0
103	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.4	0.6	0.9	0.9	0.7	0.4	0
104	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.3	0.6	0.4	0.9	0.9	0.3
105	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6	0.1	0.6	0.6	0.7	0
106	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.7	0.6	0.1	0.7	0.1
107	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.6	0.3	0.6	0
108	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4	0.9	0
109	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
110	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.4

The stimuli generated from pure tones in the range around 1 kHz were also used during pitch ranking experiments. Results are displayed in Tables 5.13 and 5.14. These results indicate poor pitch quality of stimuli with many pitch reversals (D = 0.4 and 0.2 for S1 and S9 respectively) and frequent inability to do pitch ranking when using the travelling wave encoding strategy.

Table 5.14 S1 Travelling wave encoding strategy 1 kHz Pitch Ranking matrix. D= 0.37 ± 0.15

kHz	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.90	0.5	1	0.5	1	1	1	1	0.5	1	0.5	0.5	1	0.5	1	1	0	0.5	0	0	0	0.5
0.91	0	0.5	0.5	0	0.5	0	0	0	0.5	0	0	0.5	0	1	0	0	0	0	0	0	0
0.92	0	0	0.5	1	0	0	0	0	0.5	0.5	0.5	0.5	0	0	0	0	0	0	0	0	0.5
0.93	0	0	0	1	0.5	0.5	0.5	0.5	0.5	0	0.5	0	0	0.5	0.5	0	0	0.5	0	0	1
0.94	0	0	0	0	0.5	0	0.5	0	1	1	0.5	1	0	0.5	1	0	0	0.5	0	0	0
0.95	0	0	0	0	0	0.5	0	0.5	1	0	0.5	1	0	0.5	1	0	0	0.5	0	0	0.5
0.96	0	0	0	0	0	0	0.5	0.5	1	0.5	0.5	0.5	0.5	1	0.5	1	0	0	0	0	1
0.97	0	0	0	0	0	0	0	1	1	1	1	0	0	1	0	0.5	0.5	0	0	0	0.5
0.98	0	0	0	0	0	0	0	0	1	0	1	0	0	1	0.5	0	0	0	0	0	0
0.99	0	0	0	0	0	0	0	0	0	1	0.5	1	0	1	1	0	0.5	0	0	0	1
1.00	0	0	0	0	0	0	0	0	0	0	1	1	0.5	1	0	0.5	0	0	0	0	0.5
1.01	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0.5	0.5	1	0	0	0	0	0
1.02	0	0	0	0	0	0	0	0	0	0	0	0	0.5	1	1	1	0.5	0	0	0	1
1.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0.5
1.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0	0.5	0	0	0.5
1.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	1	0	0	0	1
1.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0	0	1
1.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.5	0	1
1.08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
1.09	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	1
1.10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5

Table 5.15 S9 Travelling wave encoding strategy 1 kHz Pitch Ranking matrix. D= 0.21 ± 0.11

kHz	0.90	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	1.00	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.09	1.10
0.90	0.5	0.2	0.8	0.7	0	0	0	0.3	0	0.3	0	0.2	0	0	0	0	0	0	0	0	0
0.91	0	0.8	0.3	0	0.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.92	0	0	0.8	0.2	0	0	0	0.3	0	0.2	0	0	0	0	0	0	0	0	0	0	0
0.93	0	0	0	0.2	0.5	0.3	0.3	0.7	0.5	1	0	0	0	0	0.3	0	0.2	0	0	0	0.2
0.94	0	0	0	0	0.5	0	0.8	0.8	0.3	0.7	0.5	0	0	0.7	0	0.3	0	0	0	0	0.2
0.95	0	0	0	0	0	0.7	0.7	1	0.5	0.7	0.7	1	0	0.5	0	0	0	0	0	0	0.5
0.96	0	0	0	0	0	0	0.5	1	1	1	0	0.2	0	0.3	0	0	0	0	0	0	0.3
0.97	0	0	0	0	0	0	0	0.8	0.5	0.8	0	0.3	0	0	0	0	0	0	0.3	0	0.5
0.98	0	0	0	0	0	0	0	0	0.7	0.3	0	0.2	0	0	0	0	0	0	0	0	0
0.99	0	0	0	0	0	0	0	0	0	0.5	0.5	0	0	0	0	0	0	0	0	0	0.5
1.00	0	0	0	0	0	0	0	0	0	0	0.3	0.5	0.3	0.3	0	0	0	0	0	0	0.7
1.01	0	0	0	0	0	0	0	0	0	0	0	1	0.2	0.8	0	0.3	0	0	0	0	0.7
1.02	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.7	0	0.2	0.3	0	0	0	1
1.03	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0	0.5	0	0	0	0	1
1.04	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.8	0.3	0	0	0	0	1
1.05	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.8	0	0	0	1
1.06	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.3	0.2	0	1
1.07	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0	0	1
1.08	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.5	0.3	1
1.09	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	1
1.10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

The pitch ranking experiments seems to correlate with the suggestion made at the end of paragraph 5.6.2, namely that the pure tones around 100 Hz, when processed with the travelling wave encoding strategy seems to carry additional (pitch) information that assist recipients when trying to discriminate pure tones. Pure tones around 1 kHz, when processed through the travelling wave encoding strategy does not seem to carry pitch information (D = 0.4 and 0.2) and does not improve the discrimination of recipients when listening to these sounds.

5.7 SUMMARY

This chapter discussed the experimental procedure that was followed. It also focussed specifically on electrodograms that show the stimulus pattern and levels of stimulation. The experimental setup of each experiment was discussed and tabled results were shown. The travelling wave encoding strategy enabled the two recipients to discriminate and

correctly pitch-rank stimuli derived from pure tone differing by 2 – 3 Hz (or 80% across 21 sounds), centred around 100 Hz. The travelling wave encoding strategy did not enable recipients to discriminate better at frequencies centred around 1 kHz.

Chapter 6 DISCUSSION

6.1 INTRODUCTION

This chapter discuss the results obtained during the experimental sessions. Results is discussed in general terms, followed by discussion of and comparison between trends from current speech coding strategies and the travelling wave encoding strategy, based on the hydrodynamic model. The effectiveness of using the discussed hydrodynamic model as a basis for speech coding strategies is discussed, together with some recommendations for improvements and further investigations using the same building blocks.

6.2 LOUDNESS BALANCING PROCEDURE

Results of loudness balancing experiments show that the stimuli loudness, determined from the recipient's threshold- and comfort-levels as set during their previous programming session, was close to the final adjustment, seen in the small offset. An interesting, but expected, result was a slight increase in the average overall comfort-levels of 0.5 current levels (+0.3% ± 1.8%) at the lower stimulation rates around 100 Hz and a slight decrease in the average overall comfort-levels of -0.4 current levels (-0.2% ± 3.6%) at higher stimulation rates around 1 kHz. Such rate-dependant changes can be attributed to the auditory system perceiving increased stimulation rates as increased loudness.

Some outliers were seen (e.g. see Table 5.6) that could not readily be explained by visual inspection of the electrodograms, electrode position or initial frequency that it was derived from, were observed. Even the outliers were comparatively small (5% in the example above), thereby confirming the stability and accuracy of the MAPs.

The procedure followed to balance the loudness of the different sounds gives a fast way (balancing of the 21 sounds could be done in 12 – 15 minutes) to balance the loudness of two sounds with minimum effort from the recipient. Other studies used more involved procedures (Henry, McKay. et. al., 2000), but with both strategies using overlapping electrodes for the various sounds and repeatability of the balancing with a standard deviation of 1% for virtually all sounds, it is concluded that this procedure lead to adequate balancing of the stimuli (within 1%) using a fast adaptive protocol. It is the author's opinion that discriminating sounds by such small loudness differences is unlikely.

6.3 DISCRIMINATION OF FREQUENCIES

Discrimination of pure tones around 100 Hz, processed with the travelling wave encoding strategy seems to be superior when compared to discrimination of the tones processed with advanced combination encoder (ACE) strategy (see Table 5.11). Improvements to discrimination of pure tones, however, seems to be more pronounced in the low frequency range of the spectrum, since almost identical discrimination trends are observed for tones around 1 kHz processed with advanced combination encoder (ACE) and the travelling wave encoding strategy respectively. The worse D seen for S9 when comparing the travelling wave encoding strategy with the advanced combination encoder (ACE) strategy might be due to the smaller numbers used in these two experiments.

When listening to the sounds processed with the travelling wave encoding strategy, both recipients had better discrimination around 100 Hz (could typically discriminate sounds that are 4 Hz apart) than towards the outer edges, i.e. 90 Hz and 110 Hz (typical discrimination of sounds separated by 8 Hz) (see Tables 5.4 and 5.6). Upon inspection of the initial pure tones and visual inspection of the stimuli, no apparent reason for this

uneven distribution of discrimination could be found. It can be noted though, that a similar-looking distribution was found in some of the advance combination encoder (ACE) strategy responses at 1 kHz (see Tables 5.7 and 5.9). It is hypothesised that there might be some boundary effects that make discrimination on the boundaries more difficult when conducting these experiments.

As shown in Paragraphs 6.1 and 6.2, the travelling wave encoding strategy seems to induce improved pitch discrimination at the lower frequencies. If the information that is used by the recipient is linked to the rectifying action of the inner hair cells, then this is in good agreement with results from Russel and Sellick (1978), as well as a more recent study where pitch discrimination of a modulated carrier was limited to below 400 Hz in one recipient (McDermott and McKay, 1996). The possible information carried by the slowing down of the travelling wave could also be more difficult to isolate at higher frequencies.

6.4 PITCH RANKING

Pitch ranking results of the stimuli processed with the travelling wave encoding strategy at 100 Hz show that the different stimuli are not only different but, in general, confirm that stimuli are associated with specific tonal qualities. Stimuli derived from higher frequencies sounds higher in pitch, while those derived from lower frequencies are perceived to have lower pitches.

Pitch ranking trends appear to be similar to those from discrimination experiments. The area where good discrimination is possible (low frequency range) also allows good pitch ranking (at low frequencies). This pitch ranking is not always, e.g. Table 5.12 and 5.13,

for the 110 Hz column, where 0 indicates a perceived pitch reversal. In these two experiments there were some uncertain pitch ranking just below 100 Hz, becoming reversed pitch above 100 Hz. The pitch reversal for these sounds could not be explained by inspecting the input pure tones or by visually inspecting the stimuli for these sounds. There might be a rate-dependant aliasing-effect induced in the recipient or it might be a function of the sub-sampling frequency of 1 kHz. The pitch ranking matrix for stimuli derived from 1.1 kHz and processed with the travelling wave encoding strategy, shows poor pitch ranking and even reverse pitch ranking (Tables 5.14 and 5.15). These reversal of pitch might again be due to some aliasing-effect due to the higher stimulation rates.

It is concluded that the travelling wave encoding strategy provides improved pitch discrimination at frequencies around 100 Hz for the two recipients evaluated with a 20% - 40% increase in discrimination when compared to current commercially available speech coding strategies. No such improvement is seen at 1 kHz for these recipients. The increased latency of the stimuli (due to the slowing down of the travelling wave) and the half-wave rectifying effect of the inner hair cells might be the reason for the improvement in discrimination at lower frequencies and the lack of improvement at the higher frequencies.

Chapter 7 CONCLUSION

In this chapter, the study will be analysed and discussed in light of the hypothesis, objectives and problem stated in the first chapter.

7.1 SUMMARY OF THE WORK

A suitable hydrodynamic model was found in literature. It represented the motion inside the cochlea of both the fluid and the basilar membrane closely and constants used were linked to physical properties of the cochlea. The differential equations describing the pressure distribution of the fluid on the basilar membrane and the subsequent movement of the basilar membrane were solved numerically to predict the position of the basilar membrane for each discrete time increment. Being able to process any input signal (not just single-frequency sinusoids) make this model suitable for speech processing as well as processing of other complex sounds. A balance between complex, highly accurate models (e.g. three-dimensional models, including macro- and micro-mechanics inside the cochlea and active sharpening through the outer hair cells) and less accurate models (e.g. one-dimensional models using transmission line theory) was found in two-dimensional hydrodynamic models that include some third-dimension effects, without considering micro-mechanics of the cochlea or active elements.

The chosen model was implemented in Matlab. Two sets of differential equations were used. One describes the distribution of the pressure on the basilar membrane, due to the movement of both the stapes and oval window, and its subsequent movement. The second set of differential equations describes the acceleration of the basilar membrane due to the distributed pressure. These equations were solved successively, using a discrete equivalent for the acceleration. The next position of the basilar membrane was computed using basilar membrane positions from previous time steps and the new acceleration and velocity values computed.

A simplified model of the hair cell transfer function was used to act as a 'rectifier' of the basilar membrane movement, increasing the likelihood of the neurons firing on negative deflections. Since the implementation of this model was in a cochlear implant system, the number of sites along the basilar membrane that is relevant to the implementation was much less than the amount needed for accurate simulation – only 22 electrodes along the length of the cochlea were required. Furthermore, the current Nucleus cochlear implant limits the total stimulation rate to 14 400 pulses per second, which could be e.g. 1000 pulses per second on 14 electrodes or 2 400 pulses per second on 6 electrodes. Because of these restraints, the output of the model was sub-sampled in both space and time to 22 discrete sites along the basilar membrane at a total rate of 14 400 samples per second.

Pure tones were used as input for the speech processing (see Figure 1.6). Using Cochlear's NIC software, the rest of the coding was done in a similar way to the current advanced combination encoder (ACE) coding strategy (except for using a linear loudness growth) and presented to recipients.

The implementation of the model as a possible speech processing strategy was evaluated using pure tone stimuli. Pitch discrimination of pure tones processed with the advanced combination encoder (ACE) strategy (Figure 1.5) and the travelling wave encoding strategy (Figure 1.6) was compared. Pre-processing simple stimuli for use in an experimental setup suggested the potential usefulness of the travelling wave encoding strategy, since it was not necessary to implement the strategy in real time within the constraints of the speech processor (Henry, McKay. et. al., 2000).

An adaptive loudness-balancing procedure was performed to ensure the removal of perceived loudness differences from the different stimuli so that discrimination would not be linked to loudness. A discrimination experiment using 21 pure tones in 1 Hz increments, around 100 Hz, established the ability of the recipient to distinguish between each of the 21 tones. A similar experiment was performed with 21 pure tones in 10 Hz increments, around 1 kHz. The outputs of these experiments were confusion matrices with triangular forms - perfect discrimination for tones far apart and no discrimination for tones very close together.

As sufficient discrimination was achieved, pitch-ranking experiments were performed. It was established that the new stimuli are not only different from each other, but can also be associated with specific a tonal qualities.

7.2 DISCUSSION OF RESEARCH QUESTIONS AND HYPOTHESIS

The primary research question addressed by this study was: Can spectral information presented to cochlear implant recipients be improved by incorporating more information regarding the travelling wave?

The hypothesis was that: Using a hydro-dynamic model of the basilar membrane, solved numerically, to act as encoder for sound, spectral information (pitch) can be more accurately perceived by cochlear implant users than using current strategies. This hypothesis was tested by inclusion of such a model as a speech coding strategy, pre-processing sound and presented to recipients in a range of experiments. Pitch discrimination experiments were done, comparing the above with current strategies using

pure tones alone. Following these, pitch ranking experiments was done to establish the tonal quality of the stimuli.

The results of the experiments done showed the hypothesis to be correct in the tested subjects, but only at 100 Hz. These experiments should be repeated with more subjects to achieve statistically significant results, however. The results showed the strength of the hydrodynamic model, a trend to discriminate better at 100 Hz compared to 1 kHz.

Current shortcomings of implementing the hydrodynamic model are its computing intensiveness and its inability to encode higher frequency information.

7.3 RESEARCH CONTRIBUTION

A solution to a time domain hydrodynamic model of the basilar membrane movement was found and this high resolution (in both frequency and time) model was converted to a practical implementation in a cochlear implant speech processor. A more advanced basilar membrane model as speech processing algorithm in cochlear implant speech processors was implemented. This model incorporates the travelling wave on the basilar membrane and improves the coding of both temporal pitch and the effect of the deceleration of the wave as it approaches its characteristic frequency in the information presented to the cochlear implant recipient.

It also provides a basic model for implementing more advanced models of the outer, middle and inner ear (and hair cell function) for future research. The modular approach of the Nucleus Matlab toolbox makes the additions created to run the experiments ideal building blocks for future research, as well as being easily compatible with additional building blocks that may add on to this proposed speech coding strategy.

It was shown that the fine temporal and spatial information contained in the travelling wave improved pitch perception in cochlear implantees.

7.4 IMPLICATION FOR COCHLEAR IMPLANTS

It might be possible for current users of the Nucleus cochlear implant to improve their pitch perception of especially the low frequencies, used in tonal languages and music, if this hydrodynamic model could be refined and implemented as a real-time speech coding strategy.

Given improved pitch discrimination in a cochlear implant recipient, Hanekom and Shannon (1996) suggest possible improvements in speech recognition. If this hydrodynamic model could be implemented as a real-time speech coding strategy, the implication for cochlear implant recipients might be an increase in their speech recognition.

7.5 FUTURE RESEARCH

Suggestions for further research are summarized below.

- A model should be developed that includes micromechanical elements, basilar membrane non-linearities (outer hair cell active elements) and fluid viscosity with output in the time domain, applicable to any input signal. From a literature survey, it appears that no current model satisfies the above criteria. Such a model could result in improved pitch discrimination at lower loudness and an improved input dynamic range to be implemented in a cochlear implant speech processor. These improvements have been seen in more recent models that incorporate some of these finer details (De Boer & Nuttal 1999).

- It would be beneficial to investigate whether the effect of coupling between basilar membrane movement and the inner hair cell is more related to the velocity of basilar membrane than its absolute position at high frequencies. The current assumption is that the inner hair cells act as a simple half-wave rectifier and the likelihood and amplitude of neural firing increases with increasing deflection of the basilar membrane.
- Given the above question, it could be valuable to investigate the effects of combining the travelling wave encoding strategy with, for example advanced combination encoder (ACE), to process frequencies above 500 Hz or 1 kHz while the travelling wave encoding strategy process the lower frequencies.
- Reducing the computational complexity or time to process incoming sound using the travelling wave encoding strategy is crucial to evaluate its function in a 'take home' experimental setup. The first suggested steps are to identify the redundant elements in the travelling wave encoding strategy, originating from the hydrodynamic model and experiment with ways to obtain the same results without needing high resolution in both the time and space domain.
- The importance of the phase changes, i.e. deceleration of the travelling wave close to the point of maximal deflection of the basilar membrane, can be evaluated by comparing pitch discrimination and pitch ranking with a simplified model that does not account for the deceleration of the wave as it reaches its point of maximal deflection.
- The importance of the travelling wave, i.e. the fine temporal structure, can be evaluated by modulating the advanced combination encoder (ACE) stimuli with the period of the pure tones used and thus including the effect of the inner hair cell transfer function that was included in the travelling wave encoding strategy.

REFERENCES

Allen, J. B., 1977. Two-dimensional cochlear fluid model: new results. *Journal of the Acoustical Society of America*, 61, 110-119.

Allen, J. B. and Sondhi, M. M., 1979. Cochlear macromechanics. *Journal of the Acoustical Society of America*, 6, 123-132.

Allen, J. B. and Neely, S. T., 1992. Micromechanical models of the cochlea. *Physics Today*, 40-47.

Böhnke, F. and Arnold, W., 1998. Nonlinear mechanics of the organ of corti caused by dieters cells. *IEEE Ttransactions on Biomedical Engineering*, 45, 1227 – 1233.

Böhnke, F.and Arnold, W., 1999. 3D - finite element model of the human cochlea including fluid-structure couplings. *ORL*, 61, 305 – 310.

De Boer, E. and Nuttall, A. L., 1999. The "inverse problem" solved for a three-dimensional model of the cochlea. III Brushing-up the solution method. *Journal of the Acoustical Society of America*, 105, 3410 – 3420.

Duifhuis, H., 1988. Cochlear macromechanics. In: Edelman, ed. *Auditory function: Neurobiological bases of* hearing. 1988, 131 – 155.

Hanekom, J. J.and Shannon, R. V., 1996. Place pitch discrimination and speech recognition in cochlear implant users. *Die Suid-Afrikaanse Tydskrif vir Kommunikasieafwykings*, 43, 27 – 40.

Hartmann, R., Tönder, N. and Klinke, R., 1998. A Versatile System for the Generation and Development of Speech Coding Strategies in Cochlear Implants. *IEEE Transactions on Biomedical Engineering*, 45, 773 – 782.

Henry, B. A., McKay, C. M., McDermott, H. J. and Clark, G. M., 2000. The relationship between speech perception and electrode discrimination in cochlear implantees. *Journal of the Acoustical Society of America*, 108, 1269 – 1280.

Janssen, M. J. A., Segal, A. and Viergever, M. A., 1978. Finite element solution of a two-dimensional cochlear model. *Journal of the Acoustical Society of America*, 64, S133.

Johnstone, B. M., Pattuzzi, R and Yates, G. K., 1986. Basilar membrane measurements and the travelling wave. *Hearing Research*, 22, 147 – 153.

Lesser, M. B., Berkley, D. A., 1971. Fluid mechanics of the cochlea. Part 1. *Journal of Fluid Mechanics*, 51, 497 – 512.

Lighthill, J., 1981. Energy flow in the cochlea. *Journal of Fluid Mechanics*, 106, 149 – 213.

Loizou, P. C., 1998. Mimicking the human ear. *IEEE Signal Processing Magazine*, 101 – 130.

Loizou, P. C., Dorman, M. and Fitzke, J., 2000. The effect of reduced dynamic range on speech understanding: Implications for patients with cochlear implants. *Ear and Hearing*, 21, 25 – 31.

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, 1622 – 1631.

McKay, C. M. and McDermott, H. J., 1996. The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites. *Journal of the Acoustical Society of America*, 100, 1081 – 1092.

Moore, B. C. J., B. R., 1986. The role of frequency selectivity in the perception of loudness, pitch and time. In: B. C. J. Moore, ed. *Frequency selectivity in hearing*, Academic Press, London, 1986, 251 – 308.

Moore, B. C. J., Alcatrana, J. I. and Glasberg, B. R., 2002. Behavioural measurement of level-dependent shifts in the vibration pattern on the basilar membrane. *Hearing Research*, 163, 101-110.

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

Neely, S. T. and Kim, D. O., 1986. A model for active elements in cochlear biomechanics. *Journal of the Acoustical Society of America*, 79, 1472 – 1480.

Pattuzzi, R. and Sellick, P. M., 1983. A comparison between basilar membrane and inner hair cell receptor potential input-output functions in the guinea pig cochlea. *Journal of the Acoustical Society of America*, 74, 1734 – 1741.

Peterson, L. C. and Bogert, B. P., 1950.A dynamical theory of the cochlea. *Journal of the Acoustical Society of America*, 22, 369 – 381.

Pickles, J. O., 1986. The role of frequency selectivity in the perception of loudness, pitch and time. In: B. C. J. Moore, ed. *Frequency selectivity in hearing*, Academic Press, London, 1986, 51 – 122.

Preyer, S. and Gummer, A. W., 1996. Nonlinearity of mechanoelectrical transduction of outer hair cells as the source of non-linear basilar membrane motion and loudness recruitment. *Audiology, Neuro-Otology*, 1, 3 – 11.

Prosen, C. A. and Moody, D. B., 1991. Low-frequency detection and discrimination following apical hair cell destruction. *Hearing Research*, 57, 142 – 152.

Rhode, W. S., 1971. Observation of vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *Journal of the Acoustical Society of America*, 65, 1218 – 1231.

Rose, J. E., Hind, J. E., Anderson, D. J. and Brugge, J. F., 1971. Some effects of stimulus intensity on response of the auditory nerve fibers in the squirrel monkey. *Journal of Neurophysiology*, 34, 685 – 699.

Ruggero, M. A., 1994. Cochlear delays and traveling waves: comments on 'experimental look at cochlear mechanics'. *Audiology*, 33, 131 – 142.

Russell, I. J. and Sellick, P. M., 1978. Intracellular studies of hair cells in the mammalian cochlea. *Journal of Physiology*, 338, 179 – 206.

Sellick, P. M., Pattuzzi, R. and Johnstone, B. M., 1982. Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. *Journal of the Acoustical Society of America*, 72, 131 – 141.

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

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

Sondhi, M. M., 1978. Method for computing motion in a two-dimensional cochlear model. *Journal of the Acoustical Society of America*, 63, 1468 – 1477.

Steele, C. R. and Taber, L. A., 1979. Comparison of wkb and finite difference calculations for a two-dimensional cochlear model. *Journal of the Acoustical Society of America*, 65, 1001 – 1006.

Steele, C. R. and Taber, L. A., 1979. Comparison of wkb and experimental results for three-dimensional cochlear models. *Journal of the Acoustical Society of America*, 65, 1007 – 1018.

Steele, C. R. and Lim, K., 1999. Cochlear model with three-dimensional fluid, inner sulcus and feed-forward mechanism. *Audiology and Neuro-Otology*, 4, 197 – 203.

Uppenkamp, S., Fobel, S. and Patterson, R. D., 2001. The effects of temporal asymmetry on the detection and perception of short chirps. *Hearing Research*, 158, 71 – 83.

Vetterli, M., 1991. Wavelets and filter masks for discrete-time signal processing. *Wavelets and their applications*, 17 – 52.

Von Bekesy, G., 1947. The variation of phase along the basilar membrane with sinusoidal vibrations. *Journal of the Acoustical Society of America*. 19, 452 – 460.

Wan, W. and Fan, C., 1991. A new mathematical model for the auditory system. In: Proceedings of the International Conference on Circuits and Systems. Shenzen, China, 13 – 15.