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.