INVESTIGATION INTO COCHLEAR IMPLANT PROCESSING STRATEGIES AND SIMULATIONS

A Bachelor Thesis
by

Floris Rotteveel

Student: Floris Rotteveel, s2898144, Bsc Biomedical Technology, third year
Supervisor: Dr. Etienne Gaudrain
Research group: dB SPL (prof. Deniz Başkent)
Submitted to the University of Groningen

August 2018

ABSTRACT

The main cause of hearing loss is the irreversible diminishing of functional inner and outer hair cells. The function of the cochlear implant (CI) is to bypass the destroyed hair cells and directly stimulate the auditory nerve at the right tonotopical locations. Cochlear implants are the most used neural implant to date. CI's directly stimulate the neurons of the auditory nerve inside the cochlea. Processing strategies, performed by the processor outside of the skull, are the methods of translating acoustic sounds into electrical signals used for stimulation. More recent strategies use novel features in coding to increase the signal-to-noise ratio and frequency resolution. While general aspects of cochlear implant processors are publicly known, the manufacturers do not release detailed information about the coding strategies used. As the coding strategies become more complex, it is becoming increasingly difficult to predict the output of the implant and how acoustic sounds are represented through the implant. Therefore, it is important to create an experimental setup on which input and output of CI's can be related to each other. The Nucleus MATLAB toolbox (NMT), by Cochlear Corp., allows the simulation of the Cochlear Nucleus cochlear implant. An investigation into the workings of NMT was conducted, and more light has been shed on the operations of this cochlear implant simulation. From the electrodograms plotted by NMT, the parameters of frequency allocation map and current levels as a function of sound pressure levels for targeted and adjacent channels can be extracted for a processing strategy. This report was written as a preparation for the design of an experimental setup to investigate the parameters of modern processing strategies.

TABLE OF CONTENTS

INTRODUCTION

Cochlear implants (CIs) are the most used neural implant to date. In the 1980s the first multi-channel cochlear implants were introduced, which provided the possibility of hearing multiple frequencies increasing speech perception significantly. Since then, new processing strategies have steadily increased the functionality of the cochlear implants. In 1995, a large group of CI recipients was able to attain 80 percent understanding of high-context sentences (NIH Consensus Conference, 1995). Today, it is common to see CI users reach normal performance in speech understanding in a silent environment.

Cochlear implantation is performed by placing the receiver under the skin and inserting the electrode array through the round window into the cochlea. The flexible electrode array assumes the shape of the cochlea. A reference electrode is usually placed outside of the cochlea for mono-polar coupling. This implant can stay in the recipient for decades. Because the processing is performed on the outside of the skull, the implant can still improve with new technology even after implantation. CIs directly stimulate the neurons of the auditory nerve inside the cochlea. The processors, placed outside of the skull, convert sound waves into electrical signals.

While general aspects of cochlear implant processors are publicly known, the manufacturers do not release detailed information about the coding strategies used. However, when studying perception through a cochlear implant, it is crucial to know what signal the implant delivers to the patient, and how these devices convert sound waves into different electrical pulses on multiple electrodes in the implant.

In the bachelor project connected to this thesis, an experimental setup will be created to be able to characterize the input/output signal through measurements, in order to further understand how the cochlear implant converts sound waves into electrical signals. This setup will consist of devices connected to the implant for measurement and MATLAB functions to generate the desired input and output formats. This thesis will serve as the introductory investigation before the experiments will be conducted.

This report is organized into two sections. First, a zoom-in approach is taken to introduce the subject of cochlear implant processing strategies by providing introductions into the anatomy of the hearing ear, hearing loss, cochlear implants, cochlear implant processing strategies, and the unknowns of the processing strategies. Secondly, an investigation will be conducted into the workings of a processing strategy simulation called Nucleus MATLAB Toolbox (NMT). This investigation shows part of the similarity between the simulation and real implants and supports the design of an experimental setup for investigating parameters of modern cochlear implant processing strategies.

ANATOMY OF THE HEARING EAR

The human ear can be divided into three sections: the outer, middle and inner ear. The outer ear is the pinna and the ear canal. The pinna is shaped to catch more sound coming from the front than from behind. The ear canal and middle ear are separated by the tympanic membrane. In the normal hearing ear, sound is passed on by the tympanic membrane to the middle ear. The vibrations cause movement in the three small bones of the middle ear called the malleus (hammer), incus (anvil) and stapes (stirrup). The footplate of the stapes is connected to the membrane of the oval window, which leads into the cochlea. The oval window forms the separation between the middle and inner ear. The cochlea is the organ which transforms the pressure variations into neural pulses. The cochlea is shaped like a snail shell with two and a half turns. The outer cavities are called the scala vestibulia and scala tympani, which houses the fluid called perilymph. Because the fluid is incompressible, any movement caused by the stapes on the oval window moves the fluid towards the counter-opening called the round window. The oval and round windows move in opposite phase. The outer cavities surround the inner cavity called scala media, which is filled with endolymph. The scala media is separated from the scala vestibuli with Reissner's membrane and separated from the Scala tympani with the basilar membrane (BM) (Alberti PW, 2006). The BM is a structure which runs along the length of the cochlea and has different properties depending on the location inside the cochlea. At the base it is narrow and stiff, whereas at the apex it is broad and flexible. Each frequency which can be heard has a certain location on the BM where the response is maximal. Because of this, the basilar membrane has a tonotopic allocation of sound waves. High frequencies stimulate more at the base and low frequencies at the apex (Wilson, 2008).

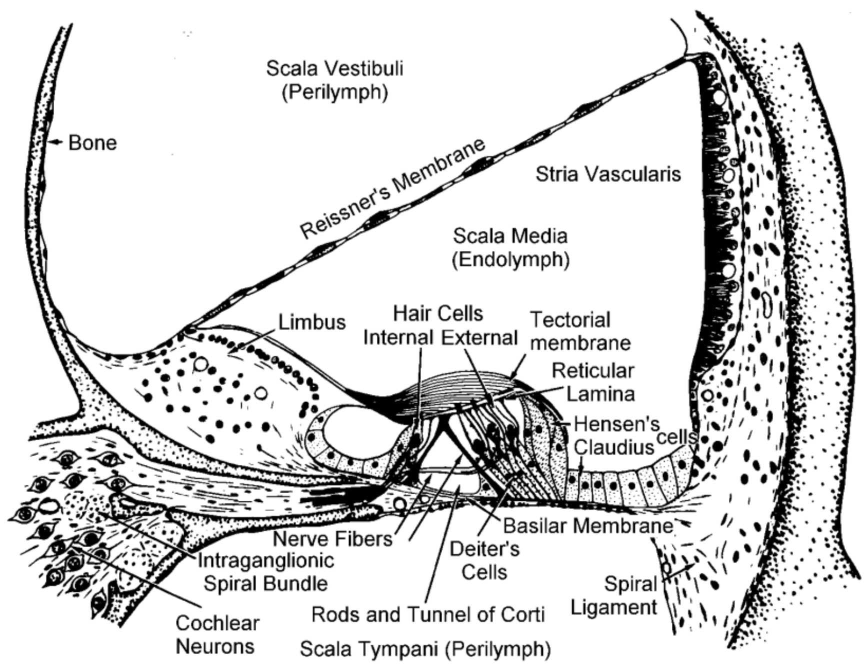

Figure 1, A cross section of the cochlea showing details of the scala vestibulia, scala media and scala tympani. (from Hallowell and Silverman, 1970)

On the basilar membrane sits the organ of Corti which has four rows of hair cells along the BM. The hair cells can be categorized into inner hair cells (IHC's) and outer hair cells (OHC's).

The normal ear has about 3,500 inner hair cells, 12.000 outer hair cells and about 30.000 spiral-ganglion cells (Nadol, Joseph B., 1993). The stereocilia of the hair cells protrude into the endolymph of the cochlea. Vibrations in the endolymph cause the stereocilia to move. The stereocilia are inter-connected via tip-links. Deflections cause the tip links to be stretched, opening stretch-sensitive ion channels. The endolymph is filled with potassium (K+). The potassium enters the hair cells via the ion channels, depolarizing the cells and thereby initiating the neural pulse. The potassium channels quickly close again, allowing for rapid pulse firing. Outer hair cells do not stimulate neurons, but act as amplifiers and sharpeners. The sound is amplified non-linearly, increasing the sensitivity to quiet sounds more than louder sounds. The sharpening effect improves pitch perception, which is particularly useful for the complex speech of humans (Wilson, 2008).

HEARING LOSS

The main cause of hearing loss is the absence of functional inner and outer hair cells. The degradation of OHC's increases the hearing threshold levels and decreases pitch perception resolution. The loss of functional IHC's causes profound deafness. Hair cells are delicate structures which can be destroyed by varying reasons. The most common causes of deafness are genetic defects, overexposure to loud sounds, infectious diseases and aging. The absence of functional IHC's causes the neurons to be disconnected from the cochlea. Mammals generally lack the ability to regrow hair cells, meaning any hearing loss is irreversible (Edge AS, Chen ZY, 2008). The function of the cochlear implant is to bypass the destroyed hair cells and directly stimulate the auditory nerve at the right locations (Wilson, 2008). Without the stimulation of inner hair cells, the most peripheral part of the spiral ganglion will degenerate. The cell body will survive however. Since the cell body is the targeted cite of stimulation of the cochlear implant, the degeneration of the peripheral part does not interfere significantly with the functionality of the CI (Hinojosa R, Marion M, 1983).

INTRODUCTION TO COCHLEAR IMPLANTS

Cochlear implants are composed of two parts. The internal, implanted, electrode-device and the external device. The external part, more commonly called the "processor", is connected magnetically to the internal part. The processor consists of a microphone, battery pack, and a speech processor. The internal device is implanted in the skull of the patient, just above and behind the ear. This device stays in the patient for decades. It consists of a receiver, a reference electrode and an intra-cochlear electrode array. Two important biological components of the signal chain are the auditory nerve and the auditory cortices of the CI user. The level of integrity of the nerves and cortices are essential for the proper functionality of the CI and can often be a bottleneck to the system. Sound waves are picked up by the microphone and processed by the speech processor. The processed signal is transmitted via magnetic induction to the receiver. The receiver, powered by the transmitter, generates electrical pulses from the processed signal and passes them on to the intra-cochlear electrode. Here, the electrical pulses are used to stimulate the auditory nerve directly. Multiple electrodes are used in the cochlea, so that multiple frequencies can be heard by the recipient. Present day cochlear implants use between 12 and 22 electrodes. A certain amount of current spread is found, when electrodes are stimulated. Current spread is the distribution of current over a part of the basilar membrane, instead of staying on the targeted electrode. The endolymph is a conductive fluid which fills the whole scala media. When one electrode is stimulated, the current is spread across the cochlea. This causes adjacent neural fibers to be stimulated. Because of this effect, increasing the number of electrodes does not necessarily produce higher frequency resolution. (Padilla et al., 2017) To prevent the charging of the cochlea, a current pulse always needs to be returned. The pulse can be returned on the adjacent electrodes, which is called bipolar coupling, or via the reference electrode, just outside the cochlea. This is called monopolar coupling and is used in most present-day systems (Wilson, 2008).

COCHLEAR IMPLANT PROCESSING STRATEGIES

Processing strategies are the ways of translating acoustic sounds into electrical signals which can be used for stimulation. The general idea of all strategies is to subdivide the frequencies into bands using a bank of band pass filters. The signal is then compressed non-linearly to ensure that the wide range of environmental sound is playable on the CI. The signal of each channel is directed to each electrode. The electrode numbering can be base-to-apex or apex-to-base, depending on the brand of the CI processor. Base-to-apex means the low-number channels have the higher frequencies. Most strategies use non-simultaneous electrode stimulation. This means that the electrodes fire sequentially, one after the other. As described later in this chapter, some strategies do stimulate simultaneously. The most common processing strategy was for a long time continuous interleaved sampling (CIS) and is available on all implant systems. Other strategies are n-of-m, spectral peak (SPEAK) and advanced combination encoder (ACE). These strategies select a certain number of channels for each frame of stimulation. Only the channels with the highest amplitudes are sent to the electrodes. N-of-m and ACE use a set number of channels. With SPEAK, this number can vary per frame depending on the composition of the input sound. However, SPEAK has a lower rate of stimulation. The channel-selection feature allows for omission of unnecessary background noises and could therefore increase the signal-to-noise (S/Ns) ratio (Wilson, 2008).

More recent strategies are SpecRes and SineEx, which make use of current steering. With current steering, tonotopic locations in between two electrodes can be stimulated by simultaneously delivering current to two adjacent electrodes. These strategies could overcome the limitation caused by the number of electrodes in the CI. However, currently these strategies do not significantly increase CI performance (Nogueira et al., 2009).

An even more recent strategy is MP3000. It is similar to n-of-m in that it evaluates the n channels with the highest magnitude, but MP3000 takes the n highest leveled channels relative to the determined masking level. MP3000 uses only 4 to 6 channels per frame and has a longer battery life compared to older strategies (Buechner et al., 2011).

UNKNOWNS OF PROCESSING STRATEGIES

Although a lot is known about modern CI processing strategies, manufacturers still keep important information classified. The first strategies described above are somewhat outdated and most manufacturers now use further developed strategies. Cochlear Corp for example uses the MP3000 strategy in their devices. The workings of the new features of MP3000 are as of yet unknown. As the coding strategies become more complex, it is becoming increasingly difficult to predict the output of the implant, and how acoustic sounds are represented through the implant. Therefore, it is important to create an experimental setup on which input and output of CI's can be related to each other.

NUCLEUS MATLAB TOOLBOX AS A SIMULATION METHOD

The Nucleus MATLAB toolbox (NMT), by Cochlear Corp., allows the simulation of the Cochlear Nucleus cochlear implant. Different processing strategies, such as CIS and ACE, can be used to process audio files and generate electrodograms, plots which show the current levels over time for all channels. Because NMT was released and used by a manufacturer of CIs, it can be assumed that the simulation is sufficiently accurate for comparison.

For the purpose of this project, different audio files and their respective electrodograms and current level–acoustic loudness curves will be generated. These graphs will then be compared to the data acquired from the same input using a real cochlear implant in the experimental setup. Understanding the level of similarity between the NMT simulation and the actual implants will be valuable information for further research in cochlear implants. When it is known that the simulation is accurate enough, experiments can be conducted using the simulation instead of the physical implant. The simulation experiments can be conducted in a fraction of the time of the physical experiments, Also, parameters can be manipulated at will without the risk of breaking the device. Before being able to investigate the level of similarity between NMT and an implant, an investigation into the working of NMT needs to be conducted.

INVESTIGATION INTO THE WORKINGS OF NMT

Only the ACE strategy will be tested, because this is the standard processing strategy which makes use of the channel selection feature. The ACE strategy is also available for CIs of brands other than Cochlear, but under different names. SpecRes, SineEx or the MP3000 strategy will not be looked at, because NMT does not allow for the simulation of these strategies.

Parameters

The most important parameters which can be altered in NMT are: T-level, M-level, phase duration, inter phase gap, and input dynamic range. T-level and M-level describe the threshold levels in current level for the quietest and loudest sounds respectively which will be processed. Phase duration is the temporal length of a single pulse. Inter phase duration is the time in between two opposite pulses. Input dynamic range is the difference, in dB, between the quietest and loudest which the CI will present at any given time. This range will shift upward or downward depending on the loudness of the environment.

Processes

Figure 2 shows the general signal flow through NMT for three coding strategies. The audio signal is sub-divided into different channels using a bank of filter bands. The output, in the form of a frequency-time matrix (FTM), is compressed with a Loudness-Growth Factor (LGF). The n most prevalent channels are selected. Then, the FTM output is converted into a channel-magnitude sequence and the pulses are mapped so that the pulses fire sequentially.

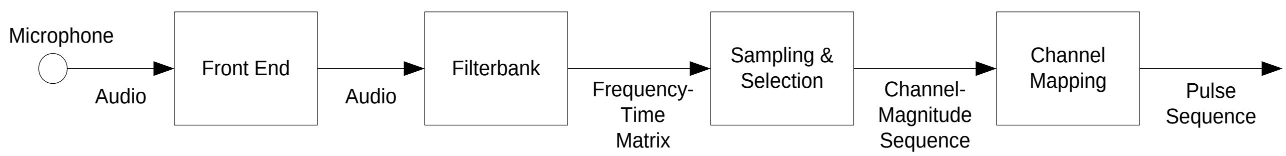

Figure 2, Signal flow in SPEAK, ACE and CIS strategies in NMT. (Swanson BA, Mauch H, 2006)

Characteristics

Figure 3 shows the amplitude compression of magnitudes performed by the LGF. The input magnitudes of range 0 to 200 dB are compressed into a magnitude range of 0 to 1, mapped to threshold and comfort level respectively. Input magnitudes above 150 dB are all compressed to 1.

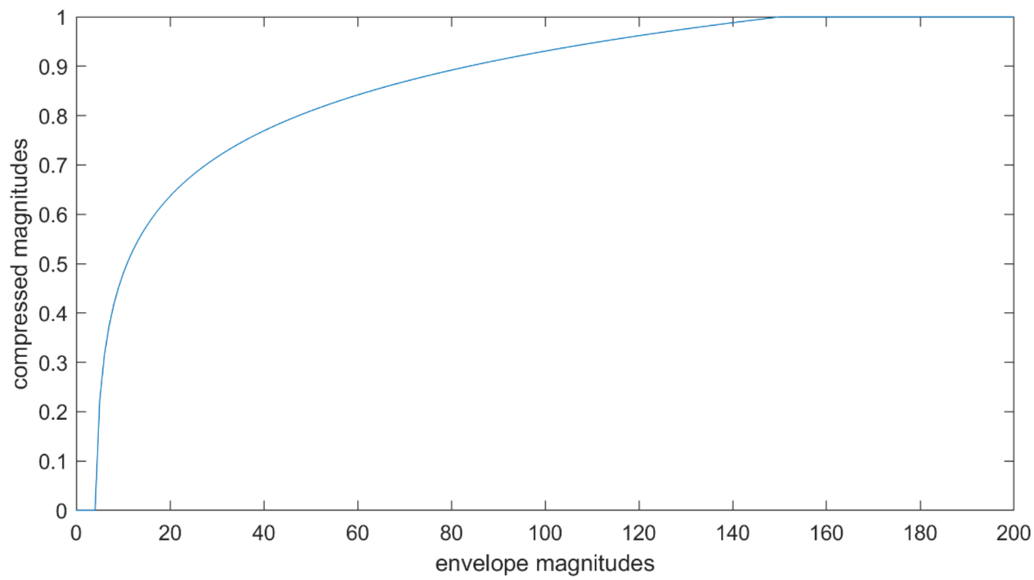

Figure 3, Loudness-Growth Factor compression of magnitudes.(Swanson BA, Mauch H, 2006)

The allocation of frequencies over the different channels of the ACE processing strategy is visualized with a channel-frequency plot (Figure 4). The channel order is base-to-apex, tonotopically speaking high to low frequencies. In figure 4, with higher frequencies, larger frequency bands are used.

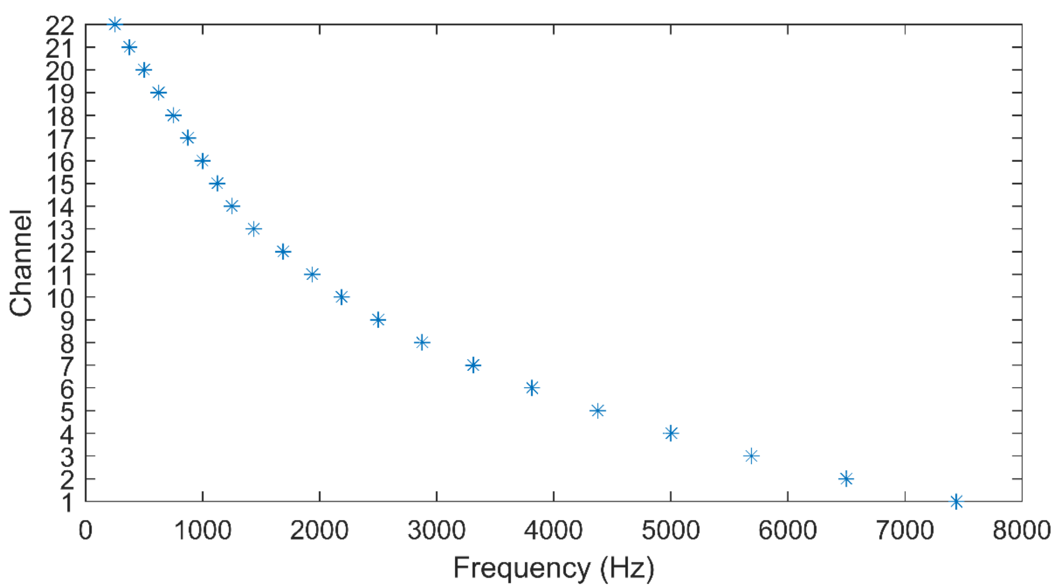

Figure 4, Channel-frequency plot of the ACE strategy using the default number of channels (22). The data points represent the middle of the frequency band of each channel.

To evaluate whether the frequencies of figure 4 are indeed mapped to those channels, an audio file was created which played each of the frequencies from figure 4 for 0.5 seconds one after the other. The loudness factor (value 0 to 1) was variable. Figure 5 shows the electrodograms of this audio file with loudness (L) values of 0.02 and 1. L = 0.02 was found to be the smallest value to produce results. The edges of the tones are smoothed using the `cosgate` function to prevent ticks at the beginning and end of each tone.

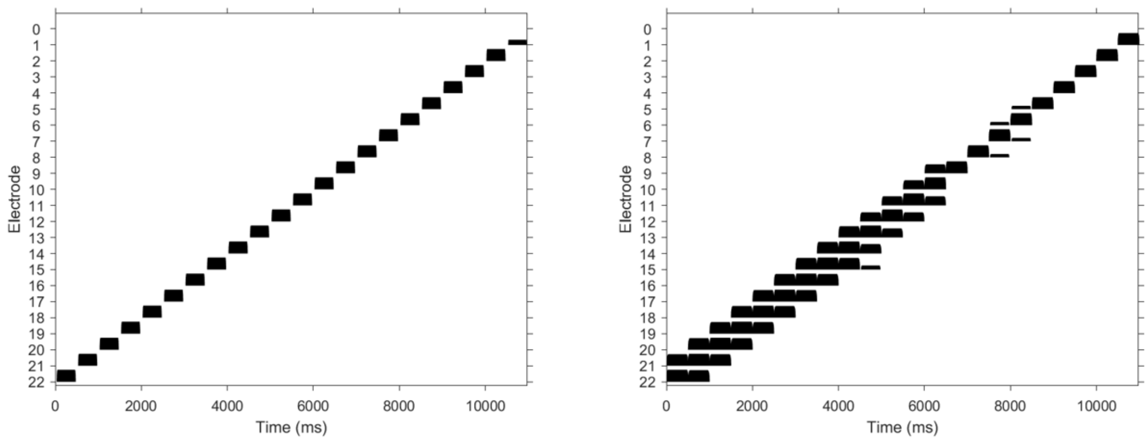

Figure 5, Electrodograms of audio files with characteristic frequencies, loudness of 0.02 (left) and 1 (right).

From figure 5 it can be concluded that with louder input audio, more adjacent channels other than the desired channel are stimulated. It was expected that for the middle frequency of a channel, only that channel would be excited for every loudness. Because of the stimulation of undesired channels at L = 1, a channel-frequency plot should be made for different loudness values. Figure 6 shows the measured channel-frequency plot for L = 0.02 and 1.

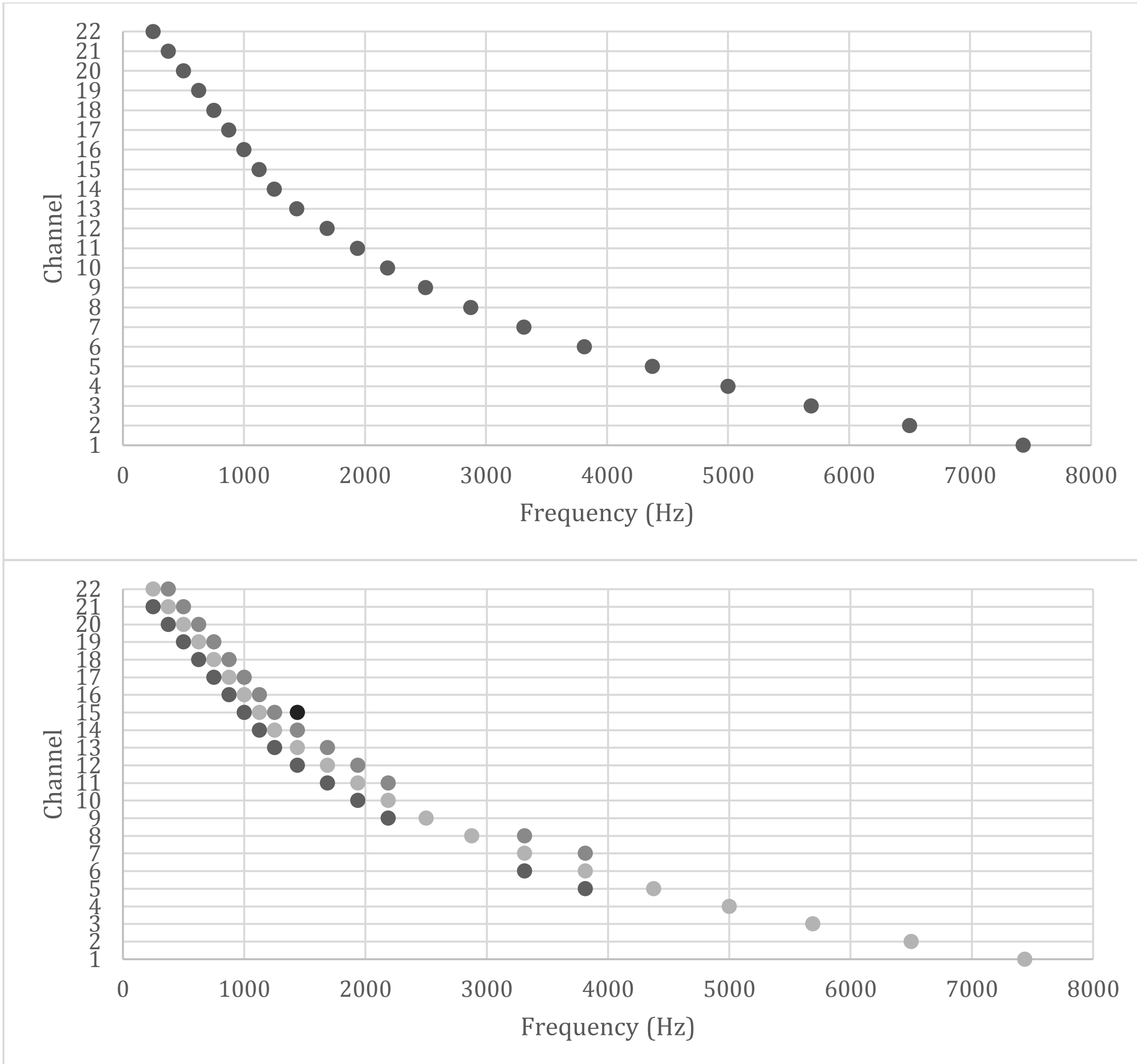

Figure 6, channel-frequency graphs for loudness (L) of 0.02 and 1.

Figure 6 shows the increase in the number of stimulated channels when the loudness is increased. In general, it can be said that for the more apical channels (higher channel number, lower frequencies), more adjacent channels are activated.

<u>Relation between sound pressure and current levels</u>

In CI's, louder input sounds are translated into higher current levels delivered to the electrodes. To visualize this relation, the current level can be plotted as a function of the acoustic level in dB.
First, the dB scale can be calculated from sound pressure levels (SPL) in Pascal using the following formula.

$$L_p = 20 * log_{10}\left(\frac{P_{rms}}{P_{ref}}\right) dB$$

Here, L_p is the sound pressure level in dB. P_{rms} is the pressure level of the signal in Pascals, and P_{ref} is the reference pressure level in Pascals. For P_{ref} a value of 1 was used.

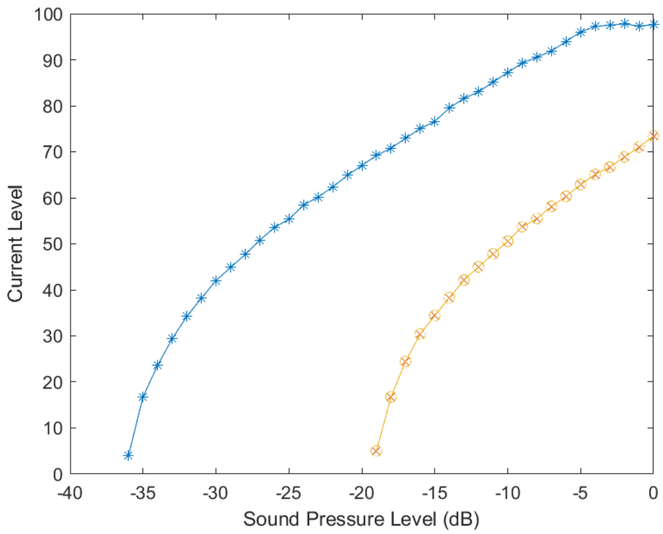

Figure 7, Current level - dB plot of channels 10 (yellow crosses), 11 (blue stars), and 12 (yellow circles), when channel 11 is stimulated. The sound pressure levels shown are relative dB's going up to 0.

Figure 7 shows the relation between sound pressure level and current levels when channel 11 is stimulated. Figure 6 showed that at certain frequencies, adjacent channels were also stimulated. This can also be seen in figure 7. Only above a certain SPL do the adjacent channels get stimulated. The current level of the adjacent channels seems to increase compared to the targeted channel. Around -20 dB the difference in current level is about 65, at 0 dB the difference is about 25. The intensity of the input signal into the brain can be manipulated by three factors: the temporal pulse length, pulse magnitude and the quantity of stimulated nerve fibers (Namasivayam, 2004). Most likely, the latter strategy is used here. The intensity of the audio signal is translated to the auditory nerve by increasing the current through the adjacent electrodes, thereby increasing the number of nerve fibers stimulated.

<u>Output of the Process function</u>

To process a wavefile in NMT, first the parameters are set, using for example p = ACE_map. Next, the wavefile can be processed with `q = Process(p, filename)`. The output of this function is a structure with six fields: `q.electrodes`, `q.current_levels`, `q.modes`, `q.phase_widths`, `q.phase_gaps`, and `q.periods`, where q is the name of the structure. The first two fields are vectors, the last four are single value parameters. `q.electrodes` is a vector of the stimulated channel numbers over time. The wavefile length is divided up into frames. In each frame, the channel numbers with the highest magnitude are placed sequentially, with some idle time in between each channel. The vector `q.current_levels` describes the magnitude of the pulse for each channel specified in `q.electrodes`. Logically, `q.electrodes` and `q.current_levels` are of the same length. `q.modes` describes the stimulation mode, however only a monopolar mode are available in this simulation. The value `q.modes = 103` is used for this

monopolar mode. `q.phase_widths` describes the width of a pulse in number of samples. `q.phase_gap` describes the length, in samples, between the positive and negative pulse. `q.periods` is the time in microseconds between the beginning of one pulse and the beginning of the next (Swanson BA, Mauch H, 2006).
Using `Plot_sequence(q)`, an electrodogram can be plotted from this structure `q` which shows the activated channels over time and the magnitude of the excitation. Figure 8 shows and example of an electrodogram.

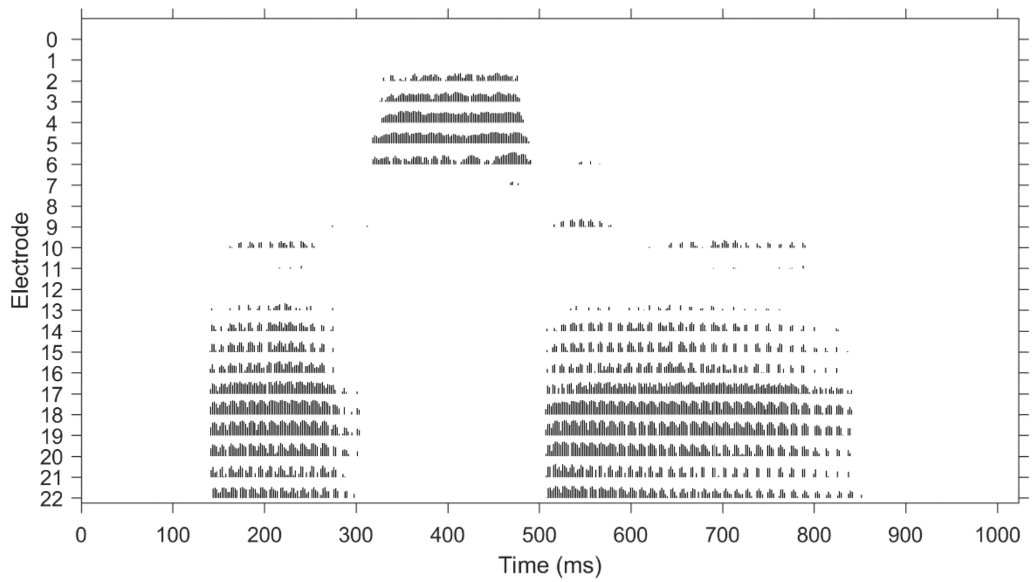

Figure 8, Electrodogram of the word 'asa', spoken by a male adult.

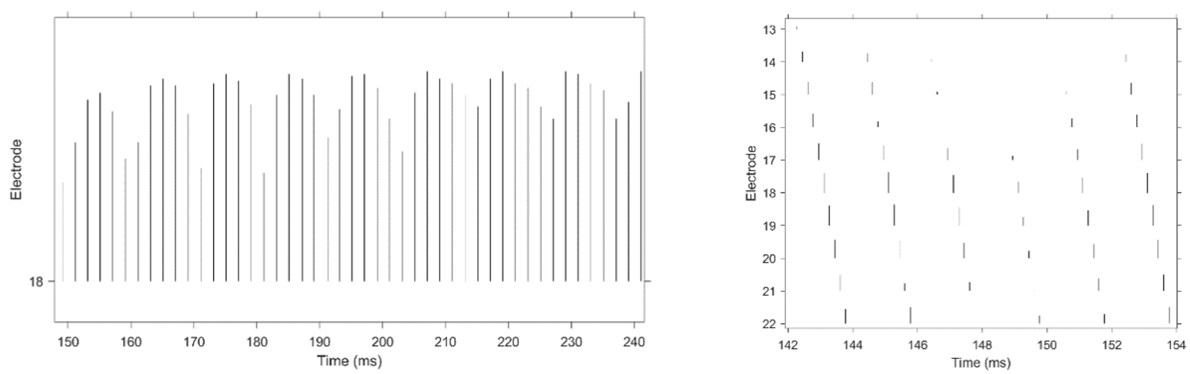

Figure 9, Zoomed in view of the 'asa' electrodogram. Left shows a single electrode and its pulse magnitudes, right shows 10 electrodes firing sequentially.

Zoomed in, the electrodogram to the left shows to be single vertical lines of different magnitudes at a constant interval. The lines start at a certain channel, the length of the line shows the current level. Because the ACE strategy was used for this electrodogram with `n = 12`, only the twelve most prevalent channels can be activated in each frame. In the bachelor project connected to this thesis, a MATLAB function will be written which converts files from an oscilloscope into the NMT form described above in this section.

DISCUSSION

More light has been shed on the workings of the cochlear implant simulation. Equipped with the knowledge of the anatomy and physiology of the human ear, cochlear implants, cochlear implant strategies and its simulation, an experimental setup can be made to investigate the contemporary processing strategies.

Because only the ACE strategy was used in the investigation into the workings of NMT, caution is advised when translating these findings to other strategies. For CIs used in patients, the parameters described in the parameters paragraph are altered to suit the recipient's preferences. For example, the threshold and comfort levels are measured and altered in the processor to avoid painful stimulations. In this investigation, only default parameters were used. NMT is a simulation of the Cochlear Nucleus processor and other manufacturers will use different processing methods. It is important to note that differences can be found because of different brands.

CONCLUSION

Cochlear implants are the best functioning neural implant to date. The level of performance of CIs allows recipients to regain the ability to communicate with others. With a steady increase in functionality of the technology since the 1980's the cochlear implant continues to advance. At this moment, the best advances are booked in the progression of the processing strategies. Manufacturers of cochlear implant systems keep the specifics of their new processing strategies largely confidential. With increasing complexity of the strategies, it is important for research to better understand the coding methods employed in modern cochlear implants. Although it only allows older processing strategies, the Nucleus MATLAB Toolbox might be a useful tool for designing an experimental setup for investigating the parameters of modern cochlear implant processors.

The Nucleus MATLAB Toolbox can simulate the workings of a cochlear implant. The processing consists of the same sub-processes as a real CI by Cochlear. The loudness-growth factor plot displays a square-root compression of the input magnitudes. The channel-frequency allocation of NMT shows smaller frequency bands for the lower frequencies and larger frequency bands for the higher frequencies. Increasing the loudness of the input signal results in more current output by targeted channel, as well as the current output of adjacent channels. As can be seen in figure 6, the relative difference in current level between the targeted and adjacent channels decreases when the loudness is increased. This effect increases the quantity of stimulated neurons around the targeted location, thereby increasing the perceived loudness. The output of NMT is an electrodogram, visualizing the output of a real implant. From the electrodograms plotted by NMT, several parameters of processing strategies can be extracted. Plotting a frequency sweep as done in figure 5 shows the frequency allocation of the processing strategy. In the case of figure 5, the frequency allocation was already known, so the middle frequencies of each channels could be neatly plotted. When investigating a new processing strategy, a frequency sweep with more pitches could be used to infer the frequency allocation from the electrodogram. Varying the loudness factor of the input shows the possible increase in current level of the targeted and adjacent electrodes. Zoomed in, the electrodogram shows the pulse rate and the n number of channels activated in each frame.

REFERENCES

Alberti, P. (2001). The Anatomy and Physiology of the Ear and Hearing. In: B. Goelzer, C. Hansen and G. Sehrndt, ed., *Occupational exposure to noise: evaluation, prevention and control*, 4th ed. [online] pp.51 - 62. Available at: http://www.who.int/occupational_health/publications/occupnoise/en/ [Accessed 21 Jun. 2018].

Davis, H. and Silverman, S. (1970). *Hearing and deafness*. 3rd ed. New York, etc.: Holt, Rinehart and Winston.

Edge, A. and Chen, Z. (2008). Hair cell regeneration. *Current Opinion in Neurobiology*, 18(4), pp.377-382.

Hinojosa, R. and Marion, M. (1983). Histopathology of profound sensorineural deafness. *Annals of the New York Academy of Sciences*, 405(1 Cochlear Pros), pp.459-484.

Nadol, J. (1993). Hearing Loss. *New England Journal of Medicine*, 329(15), pp.1092-1102.

Namasivayam, A. (2004). Cochlear Implant Technical Issues: Electrodes, Channels, Stimulation Modes and more.

NIH consensus conference. Cochlear implants in adults and children. (1995). *JAMA: The Journal of the American Medical Association*, 274(24), pp.1955-1961.

Nogueira, W., Litvak, L., Edler, B., Ostermann, J. and Büchner, A. (2009). Signal Processing Strategies for Cochlear Implants Using Current Steering. *EURASIP Journal on Advances in Signal Processing*, 2009(1).

Padilla, M., Stupak, N. and Landsberger, D. (2017). Pitch ranking with different virtual channel configurations in electrical hearing. *Hearing Research*, 348, pp.54-62.

Swanson, B. and Mauch, H. (2006). *Nucleus Matlab Toolbox 4.20 - software user manual*.

Wilson, B. (2008). Cochlear implants: Current designs and future possibilities. *The Journal of Rehabilitation Research and Development*, 45(5), pp.695-730.

Wilson, B. and Dorman, M. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research*, 242(1-2), pp.3-21.