Understanding Hearing Through Cochlear Implants - A Simulation Model of the Complete Hearing Path

A thesis submitted for the degree of
Philosophy Doctor in System Engineering

Politehnica University Timisoara, Romania
Faculty of Automation and Computers

Artur KUCZAPSKI

Supervisor: Professor Dr. Ing. Gheorghe-Daniel ANDREESCU
Referends: …
…
…

2019

FOREWORD

This work is dedicated, with love, to my son Eduard, who was not gifted by birth with the ability to perceive the acoustic universe which surrounds us all. It was not our fault, as parents, and certainly, it was not his. My duty, as a father, is to help him in any way I can to overcome his disability. I began to inform myself in a field where I had no previous interest, but which became a major focus in my activity. I learned about cochlear implants (CIs) and I managed to provide the device and the needed medical attention. Nowadays my son can hear with the aid of a cochlear implant and can interact with us using sounds.

I observed that there is a considerable spread of performance among different users of cochlear implants. I wondered why. I learned about what different types of implants can offer, about the age of implantation in prelingual children, about the various [re]habilitation programs. Yet something was missing. I continued to learn about the way healthy people are hearing, trying to compare it with the way implanted people are hearing. I learned about the physiology of the acoustic nerve and how it passes the information to higher neurological structures; but this was not enough, so I tried to develop a model for the electric charge distribution inside the cochlea when stimulated by a cochlear implant.

Trying to understand the artificial stimulation of the acoustic nerves, I built a digital oscilloscope with twelve channels to observe the waveforms applied to the cochlea. The CI manufacturers are not happily disclosing such information, in their effort to protect their intellectual property, but this tool proved to be very helpful in understanding how the cochlear implants help deaf people hear. In a few cases, I was even able to help other parents by detecting and debugging some peculiar malfunctions. Most importantly, I was able to collect the data necessary to represent the artificial hearing brought by cochlear implants, which was the ultimate goal in understanding what sounds my son has to deal with.

By trial and error and starting from scratch, I reached a certain level of understanding and reached a certain number of conclusions for a better use of cochlear implants and for maximizing the results. I am a member of a more than 400 people self-helping group called *Asociatia Persoanelor cu Deficiente de Auz "Asculta Viata" ("Hear Life" Hearing Impaired Association)*. I was able to help others with my findings and now I want to validate what I learned through the prism of a PhD program and to offer my research to the public domain, so it can reach an even broader audience.

ACKNOWLEDGEMENTS

I cannot thank my PhD advisor Professor Dr. Ing. Gheorghe-Daniel ANDREESCU, enough for his exquisite support in putting together the results of my research. His solid experience gracefully transmitted to my work made this paper round-shaped and clear for reading and understanding.

Likewise, I owe thanks to my former teacher, the fitting specialist working with my son, and my dear friend, Lecturer Dr. Ing. Antonius-Nicolae STANCIU, who not only supported the habilitation of my son, but also initiated me in the technologies related to cochlear implants.

I also owe a big thanks to other parents of children wearing cochlear implants, for sharing their own experiences with me, which helped objectivize my findings and gave statistic relevance to my research.

I have to thank also my wife and kid for supporting me during my working long hours in a pretty undocumented field. Even my little son somehow understood that my efforts are for his benefit.

Least but not last, I feel that here is a good place to thank for my engineering background attained at the Politehnica University Timisoara, which made this further work possible.

TABLE OF CONTENTS

TABLE OF FIGURES

1 INTRODUCTION

Cochlear implants are implantable electronic prosthetic devices developed to restore hearing in patients with severe to profound hearing loss. Most frequently hearing loss is caused by damage or complete loss of the sensory hair cells of the cochlea, making it unable to convert sound vibrations to auditory nerve stimuli. Cochlear implants restore hearing by capturing sounds with an external microphone and converting them into electric impulses delivered directly to the hearing nerves. The delivered electric impulses stimulate the hearing nerve generating nerve impulses partly resembling the patterns created by the natural hearing process. Although the created hearing sensation is not natural and it has significantly inferior quality compared to the natural hearing, in most cases hearing is still restored to a degree which allows speech understanding in reasonable listening condition. Despite this limitation, the cochlear implant can still be considered the most successful and widespread bionic implant (Eshraghi, et al., 2012) (Hochmair, et al., 2007) (Blake S. Wilson, 2008) (Drennan & Rubinstein, 2008).

1.1 SHORT HISTORY OF COCHLEAR IMPLANTS

Although the development of modern cochlear implants was started in the 1970s, electrically induced hearing sensations were first reported in the early 1800s by Alessandro Volta, the developer of the first electric battery or electric pile. Studying the effects of the electric current over living bodies, Volta has inserted two low voltage electric probes connected to an electric pile into his ear canals and discovered that switching on and off the current through the probes, crackling like sound is perceived. The experiment was repeated in 1855 with alternating current by Duchenne de Boulogne, who experienced similar findings as Volta. In 1930, Wever and Bray discovered strong correlations between the electrical action potentials recorded from the cochlea, and the sound stimulus delivered to the ear during the recording (Hallpike & Rawdon-Smith, 1934). The findings of this experiment, called "Wever and Bray phenomenon" suggested that useful hearing could be restored if one could generate and deliver similar electric stimulation directly into the hearing nerves. (Eshraghi, et al., 2012)

The first direct electrical stimulation of the auditory nerve was performed in 1957, by André Djourno and Charles Eyriès. During the surgery of a patient with bilateral cholesteatoma and complete hearing loss due to radical mastoid surgery, they have implanted a special designed electrode close to the remaining of the auditory nerve, and an induction coil placed on the temporalis muscle and connected to the electrode. After the surgery, using electromagnetic coupling, they transmitted a signal from a microphone to the implanted electrode. The experiments showed that the patient was able to detect different sound intensities, but the frequency discrimination was poor and limited to frequencies below 1KHz. (A & C., 1957) (Eshraghi, et al., 2012).

The publication of Djourno and Eyriès's experimental results has attracted the attention of Dr. William House and Dr. John Doyle from Los Angeles, and in 1961 they started to implant the first patients. Two types of implants were used in these surgeries: single electrode implant and implant with five electrodes. They developed a surgical technique to insert the electrodes directly into the *scala tympani* through the round window of the cochlea. The performed surgeries lead to promising but limited outcomes. Patients were able to regain some marginally useful hearing, but due to the used materials, infections and electrode rejections were frequent bringing the procedures to a halt. (House WF, 1973) (House WF, 1976).

The first cochlear implant systems, usable by patients outside the laboratory, was developed by the University of California at San Francisco during the 1970s. They have created both single and multi-electrode cochlear implant systems, with body-worn external sound processing units (Wilson & Dorman, 2008).

After the 1970s, cochlear implants were continuously improved leading to better and better hearing restoration, going from simple awareness of sound to useful speech perception in listening scenarios with quiet and low ambient noise. Speech perception in difficult conditions, like concurrent speaking situations or high ambient noise, is still rather impaired among Cochlear Implant users. Music perception, especially tones and melody, is considered to be most difficult and even impossible for many patients (Wilson & Dorman, 2008).

Several studies were performed to understand the relation between the perceived sounds and the location and patterns of electric stimulation. It was observed that the place of the stimulation plays an important role in the pitch of the perceived sounds. Electric stimulation in the basal region of the cochlear duct induces high pitched sound perception, while stimulating the apical region of the cochlear duct induces low pitched sound perception. Besides the stimulation location, the rate of the stimulation impulses can influence the perceived pitch too. When stimulation rates below 500 Hz are used, the rate of the stimulation determines the pitch of the perceived sound (Schnupp, et al., 2011). According to these experiments, there are two main mechanisms to analyze and identify perceived sounds:

- **Spatial cues** – the positions along the basilar membrane stimulated by the sound
- **Temporal cues** – the rate of the nerve firings created by sound

The single electrode cochlear implants were based on transmitting the sound information only relying on temporal cues, while multi-electrode cochlear implants mostly rely on spatial cues. In case of typical multi-electrode cochlear implants, the sound recorded by the microphone is split, using bandpass filters, into as many frequency bands as many electrodes the implant has. The amplitude envelope of the filtered signals is used to modulate the intensity of the stimulation levels of each individual electrode. In most cases the stimulation rate is fixed, and temporary cues are not utilized (Wilson & Dorman, 2008) (Hochmair, et al., 2007) (Schnupp, et al., 2011) (Somek, et al., 2006) (Moctezuma & Tu, 2011) (Choi & Lee, 2012). Today only a few coding strategies exist that use both spatial and temporal cues (Harczos, et al., 2013) (Nogueira & Buechner, 2012) (Chen & Zhang, 2006).

Hearing restoration levels and quality widely vary from patient to patient and from one cochlear implant type to the other. Perhaps the most important factors to influence the success of the hearing restoration are:

- The number and placement of the stimulating electrodes
- The stimulation strategy used to convert sound in electric stimuli (Moctezuma & Tu, 2011) (Choi & Lee, 2012) (Somek, et al., 2006).
- The ability of the patient to adapt and learn to interpret the artificially generated hearing sensation.

The last point in the list above, is also the most difficult to predict and control. It was observed that, even with identical cochlear implants, each patient will benefit differently.

There are three major categories of implant recipients with specific recovery expectations:

- **Patients with postlingual hearing loss.** These (mostly adult) patients have lost their hearing, well after they learned to speak. In this situation, the brain has already learned to interpret the nerve signals of the cochlea. After implantation, in most cases, these patients will regain useful hearing soon after implantation, although they will report robotic sounds and mediocre to poor hearing quality.

- **Patients with prelingual or congenital hearing loss, with late implantation.** Due to the hearing loss, these patients didn't learn to hear and speak in their childhood, and because of the late implantation, later than the age 7, it is highly unlikely that useful hearing will be restored. It is suggested (Ramsden, 2002) that the high plasticity of the brain in early childhood is required to learn to interpret the complex signals of the hearing nerves.

- **Patients with prelingual or congenital hearing loss, with early implantation.** In these cases, the cochlear implant is implanted in early childhood, between ages 1 and 2. The outcome of these interventions is highly varied, but typically after long-lasting hearing and speech therapies, useful hearing is established with showing significantly higher performance and quality compared to the patients with postlingual hearing loss. Some of these patients perform so well, that they manage to learn to play musical instruments.

In all of these categories, the success is conditioned by the right IC parameter fittings and attended hearing and speech therapy sessions, and even after the most successful recoveries, the hearing in complex auditory scenarios is highly inferior compared to healthy natural hearing.

Researchers have studied the nature of the perceived sounds, and based on auditory models and reports of postlingually deafened cochlear implant recipients,

several models and software algorithms have been developed, to synthesize the sounds supposedly perceived by the cochlear implant users. (Mahalakshmi & Reddy, 2012) (Chilian, et al., 2011) (Loebach, 2007).

Such *auralization methods* and systems can be very useful in various areas of the cochlear implant research and utilizations:

- Objectively compare expected hearing quality with existing sound coding strategies;
- Test bench for development of new coding strategies;
- Improve fitting procedures;
- Help speech therapists to understand hearing through CI.

The existing auralization methods provide a good indication of how the sounds are perceived by mid-performing, mostly postlingually deafened cochlear implant users, but they fail to predict the hearing quality of peak performing patients.

1.2 Research Objectives and Structure of the Thesis

The *main objective* of the present work is *to develop a reliable simulation model to predict, quantify and demonstrate the expected hearing performance of cochlear implant users* in a way which is easily comprehensible not only by tech professionals, but also by non-technical oriented people like speech-therapists, psycho-therapists, and maybe most importantly, by relatives of the patients.

It is my considered opinion that the most appropriate method to demonstrate the hearing quality of a hearing-impaired person is to reproduce the sounds they perceive and replay it to normal-hearing persons. This is why the proposed *target is to develop and implement a novel auralization method of the perceived sounds which transforms the electrical output of a cochlear implant into sounds* relaying not only on the physiology of the ear, but also *mimicking the brain ability to learn and adapt to new stimulation patterns*.

Also, I had the opportunity to study and follow the evolution of more than 300 children with cochlear implants. These children are patients of the fitting specialist Dr. Ing. Antonius STANCIU, and they undergo periodic hearing tests and fittings. All these patients are MED-EL cochlear implant users, and all their audiograms and fittings are archived within the fitting software's database.

As a *secondary objective* of the thesis, I have proposed to explore the possibilities to help these patients by *providing various tools to aid the fitting process and troubleshooting of defects using ideas and methods developed as a spin-off* during the research efforts invested in the main objective of the thesis.

Figure 1 depicts the structure of the present thesis reflecting the duality of the research objective. The thesis starts with an introductory chapter (Chapter 1) consisting of a *short overview of the history of the cochlear implants*, and the presentation of the *research objectives*.

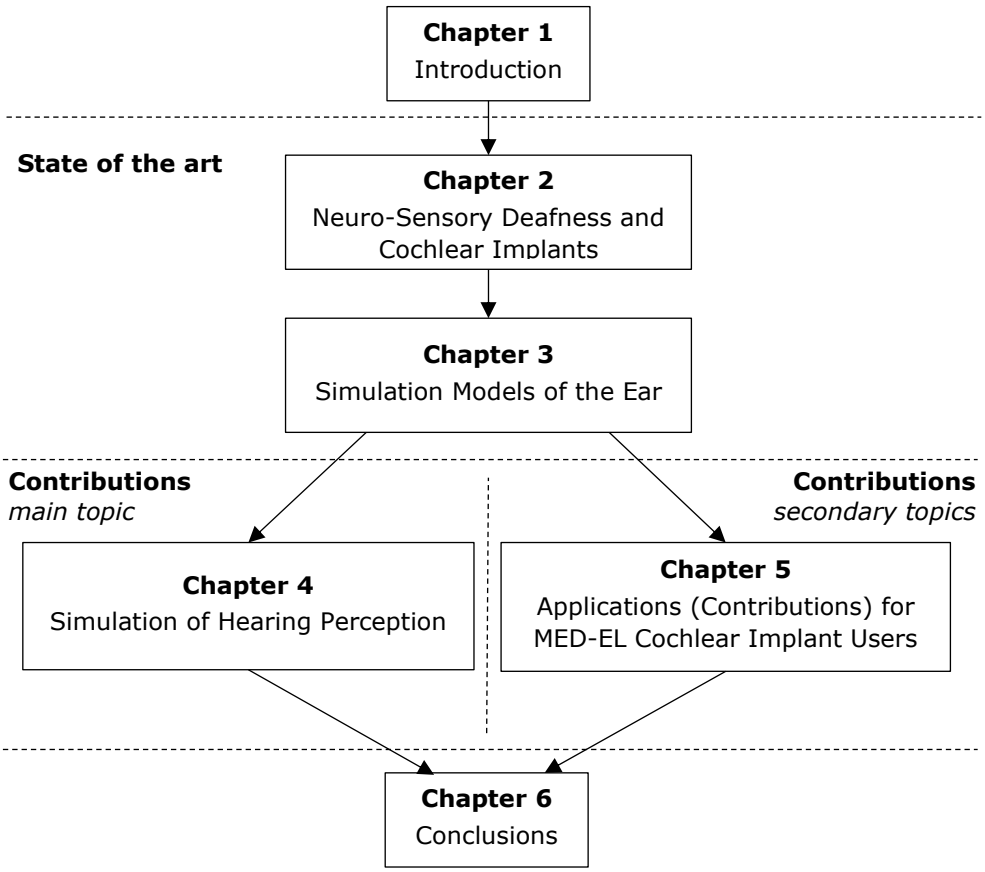


Figure 1: Thesis structure

The second part of the thesis, describing *state of the art* relevant to the proposed research objectives, is divided into two chapters.

In the first chapter of this part of the thesis (Chapter 2) the *anatomical structure of the ear and the neuro-sensory deafness* is presented including *relevant aspects of the mechanisms of hearing*. In the same chapter, a *review of the existing cochlear implants and their stimulation strategies* is given.

In the second chapter of this part of the thesis (Chapter 3), the *mathematical models necessary to the simulation of the ear* are introduced. These models are focusing on the main structures of the inner ear: the basilar membrane, hair cells and auditory nerves firing model.

The third part of the thesis is reserved for the original contributions.

The *contributions to the main research objective* are described in Chapter 4. First, a *simplified model of the cochlear* is introduced and implemented, followed by a *comparison between the existing auralization models and the newly proposed auralization model*. The *implementation of the newly proposed auralization model is detailed and experimental results* are explained.

The *contributions to the secondary research objectives* are presented in Chapter 5, introducing *MED-EL's cochlear implant system* and describing *five different original contributions* related to it:

- Statistics of typical stimulation levels
- Computer-assisted fitting of cochlear implants – Tracking the Effective Stimulation Threshold
- Case studies of Fitting evolution
- Interfacing with Med-El Cochlear Implant Processors
- Intra-cochlear current flow model

Finally, Chapter 6 concludes the thesis presenting the summary of the contributions and comparing them to the original scope of the thesis.

2 NEURO-SENSORY DEAFNESS AND COCHLEAR IMPLANTS

2.1 THE EAR

The ear is the organ of hearing and balance. It is responsible for translating mechanical vibrations and body accelerations into nerve signals that are interpreted by the brain. The ear is composed of three delimited parts named according to their positions, outer, middle and inner ear (Figure 2). Each of these parts is responsible to convey and transform sound vibrations through different mediums. The *outer ear* is responsible to channel and to amplify the vibration of the air through the ear canal directly to the tympanic membrane or the ear eardrum. This membrane is delimiting the outer and the *middle ear* and it translates the air propagated vibrations into bone propagated vibrations by connecting to the auditory ossicles of the middle ear. There are three ossicles connected, the malleus, the incus and the stapes and they connect the tympanic membrane and the membrane of the oval window situated on the cochlea, which is the organ of the inner ear. Sound vibrations of the air are poorly transferred into a liquid medium; therefore, the main role of the ossicles is to facilitate the vibrations to enter the liquid filling the cochlea, acting as an impedance adapter. Another role of the ossicles is to limit the amplitude of the transmitted vibrations. This is achieved by two small muscles attached to them, the tensor-tympany and the stapedius, which can constrain the movement of the ossicles when necessary. Finally, the *inner ear* transforms the vibrations of the contained liquid into nerve signals transmitted to the brain through the auditory nerves. The anatomy of the ear and the propagation of the sound vibrations are shown in Figure 2 (Moctezuma & Tu, 2011).

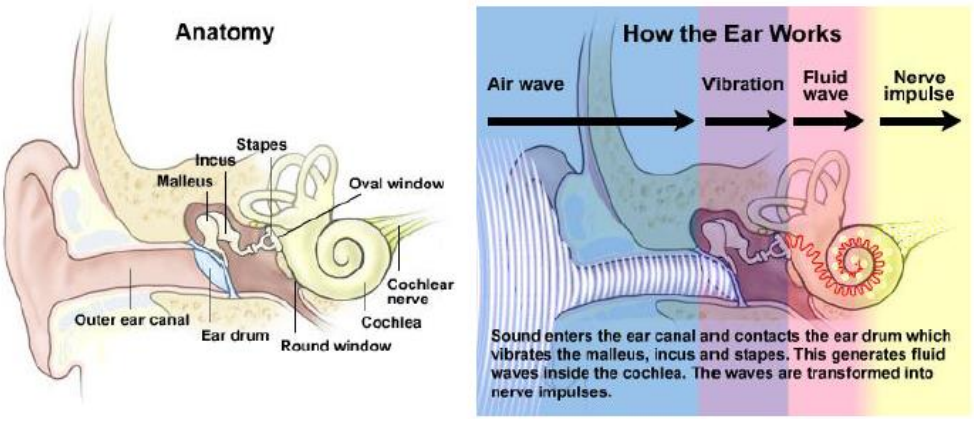


Figure 2: The anatomy (left) and the physiology (right) of the normal ear (Moctezuma & Tu, 2011)

The *cochlea* is a spiral-shaped bone cavity of the inner ear filled with liquid, named the cochlear duct (Figure 4). The cochlear duct is longitudinally divided by the basilar membrane creating two canals communicating only at the end of the cochlear duct through the helicotrema. The first canal is the vestibular duct (Scala Vestibuli) and it is starting at the oval window, where the sound vibrations are entering the liquid contained within. The vibrations travel through the vestibular duct, alongside the basilar membrane, up to the helicotrema where they are transferred to the second duct, the tympanic duct (Scala Tympani) and conveyed back to the beginning of the cochlear duct, where the vibrations are leaving the cochlea through the round window.

At each point of the basilar membrane, a part of the vibrations is transferred through the basilar membrane, from the vestibular duct to the tympanic duct. The frequency of the short-circuited vibrations is dependent on the position on the basilar membrane. The basal part of the basilar membrane, situated close to the round and oval windows, only permits the transfer of high-frequency vibrations, around 20 kHz, while at the apical part close to the helicotrema only low-frequency vibrations, around 20 Hz, are transferred. Thus, the basilar membrane, similar to a prism, effectively decomposes the vibration into frequency components (Schnupp, et al., 2011). Figure 3 shows the approximate locations of the frequencies transferred through the basilar membrane.

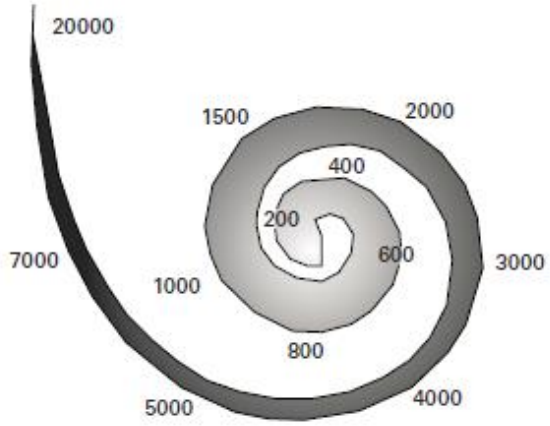


Figure 3: Approximate best frequencies of various places along the basilar membrane, in Hertz (Schnupp, et al., 2011)

On top of the basilar membrane, other organs and structures are situated (Figure 4):

- Scala media: a narrow duct filled with liquid, between Scala Vestibuli and Scala Tympani;
- The *organ of Corti*: situated in the Scala media, that consists of the Tectorial membrane, the outer hair cells and the inner hair cells connected to the termination of the auditory nerves.

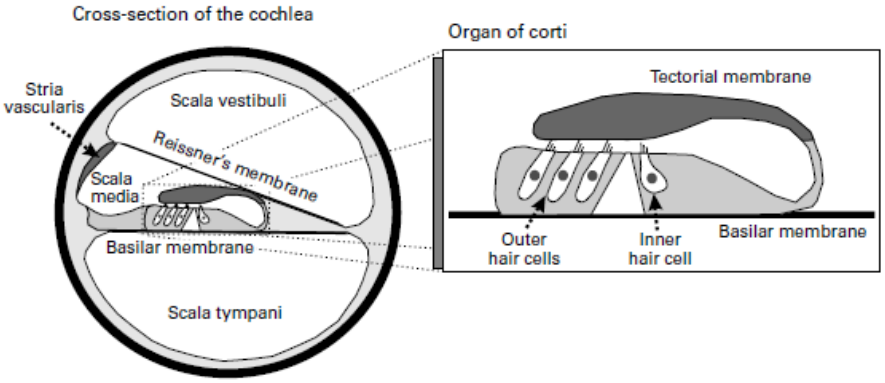


Figure 4: Cross section of the cochlear duct and detailed schematic structure of the basilar membrane and organ of Corti (Schnupp, et al., 2011)

The hair cells are named after the apical modifications grown at the top of the hair cells, the stereocilia, which are tiny hair-like fiber formations. These stereocilia react to the slightest displacement of the basilar membrane causing electric-polarization of the hair cell. Auditory nerve fibers are only connected the inner hair cells, and when due to the stereocilia movement and the cell polarization, the nerve action potential is reached, nerve impulses are generated. The nerve impulses are conveyed to the brain, processed, and perceived as sounds. The outer hair cells are not connected to the auditory nerves, instead, they are reacting to the vibrations by contractions and thus generating more displacement of the basilar membrane, effectively amplifying the stimulation of the inner hair cells (Schnupp, et al., 2011).

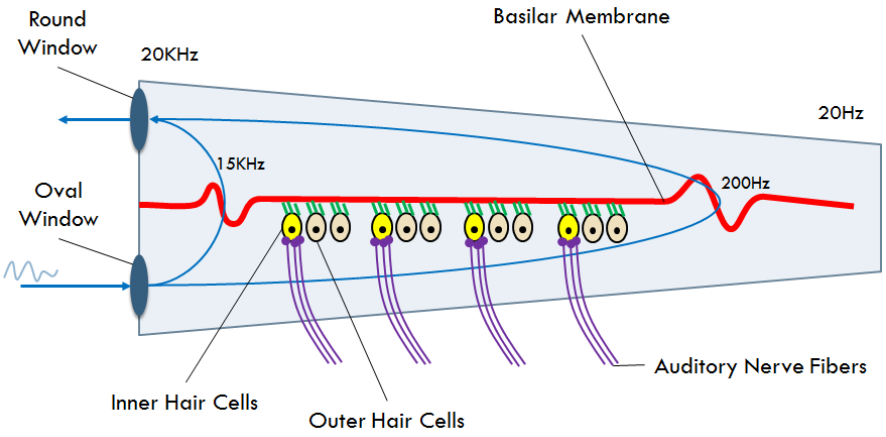


Figure 5: Simplified representation of the cochlea, basilar membrane, hair cells and auditory nerve fibers

Figure 5 shows a simplified and unfolded representation of the cochlea. In the image, a composite sound vibration is applied to the oval window, combining a 15 kHz and a 200 Hz pure tone sound signals. The displacement of the basilar membrane is schematically depicted at the two positions corresponding to the composing frequencies, and the propagation path and direction of the vibrations are marked. A few of the outer hair cells, and the inner hair cells connected to the auditory nerve fibers are depicted too.

Because of the frequency decomposition of the vibrations, the position of the stimulated nerve provides good information of the approximate frequencies of the stimulating vibrations, however, the nerve stimulation intensity and the stimulation patterns are important to sound perception and recognition too.

It has been observed that the hair cell membrane voltage, and implicitly the intensity of the stimulation, increases with the amplitude of the vibrations, and in case of low and mid frequencies, the polarization voltage oscillates in sync with the frequency of the stimulating vibration. (Schnupp, et al., 2011)

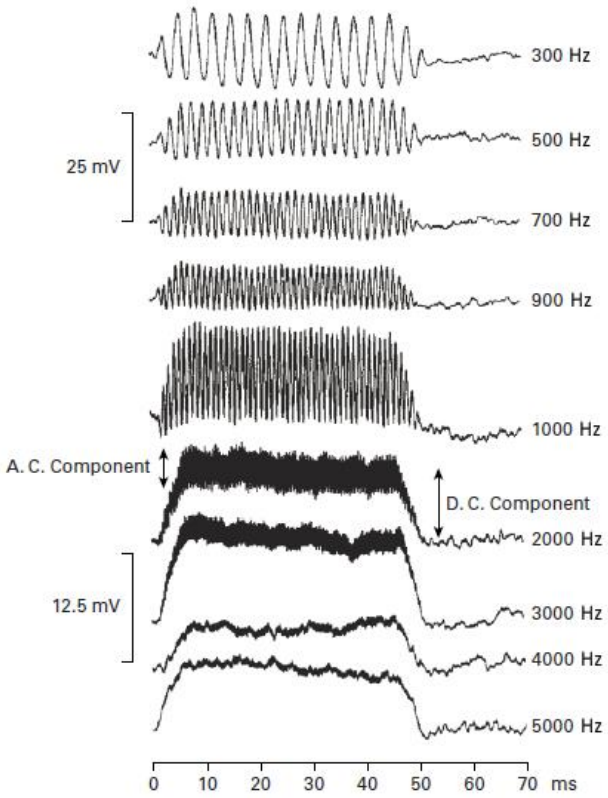


Figure 6: Changes in hair cell membrane voltage at different stimulation frequencies from 300 Hz to 5000 Hz (Schnupp, et al., 2011)

Figure 6 shows the results of an experiment, where the hair cell membrane voltage was recorded, while the stimulation was 80dB pure tone sounds of different frequencies. It can be observed that at 300 Hz stimulation, the recorded voltages are closely resembling the stimulation waveform with a very low DC (continuous) component. Increasing the frequencies, the DC component increases.

Up to 1 kHz, the oscillation amplitude is maintained, and the recorded waveform still resembles the stimulation waveform. Between 1 kHz and 4 kHz, the amplitude of the oscillations within the recorded voltages is decreasing, and above 4 kHz the polarization does not contain the component of the stimulation frequency anymore.

Because the probability of the nerve firing depends on the polarization of the hair cell, below 4 kHz the nerve firing rate matches the exact stimulation frequency, carrying valuable information regarding the exact composition of the perceived sound. Furthermore, the nerve impulses are phase locked to the stimulation signal.

Figure 7 depicts the recording of an experiment showing the phase lock effect at 100 Hz. It can be observed that the distances between impulses are always equal or an integer multiple of the stimulation signal period and the nerve impulse always happened at the same phase.

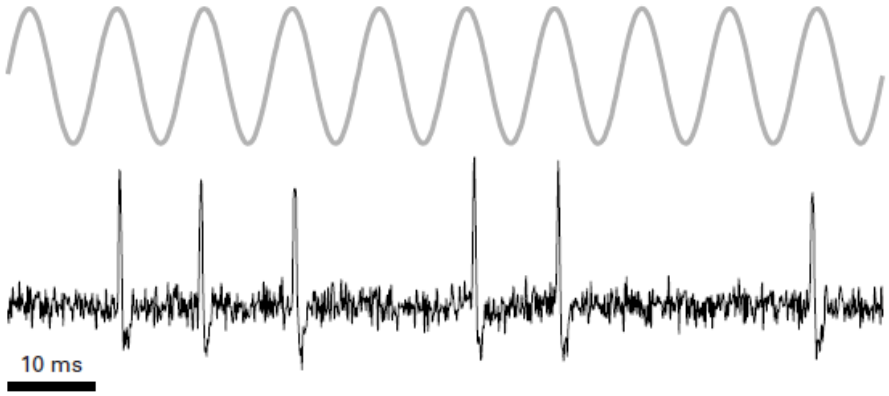


Figure 7: Comparison between stimulation waveform and recorded nerve impulses at 100 Hz (Schnupp, et al., 2011)

Based on the structure of the cochlea and the experimental observations, it was concluded that the ear uses two main methods to encode information into nerve impulses (Mahalakshmi & Reddy, 2012) (Schnupp, et al., 2011):

- **Spatial cues / Place Coding** – the positions along the basilar membrane stimulated by the sound;
- **Temporal cues / Rate Coding** – the rate of the nerve firings created by the sound.

Place coding is obtained by the basilar membrane's ability to decompose the frequency spectrum and to stimulate only nerves related to composing frequencies of the stimulus. This type of coding is available to the whole spectrum of hearable frequencies, but the exact frequencies are not transmitted, and therefore the pitch detection accuracy provided by place coding is limited.

Temporal coding is the method used to encode very accurate information regarding the pitch of the perceived sound. According to the experiments, this coding only works for frequencies below 4 kHz.

Supposedly (Schnupp, et al., 2011), to detect the sound pitch, the generated nerve impulses are further processed in the brain, using an autocorrelation like method, to detect periodicities of the sound which are not present as explicit frequency components. This autocorrelation mechanism can explain why vibrations emitted by certain musical instruments are missing the fundamental frequency components, but they are still perceived with the desired pitch.

The mechanisms used by the ear to convert sound stimulus into nerve firing patterns are extensively studied and well understood, however, what are the *exact mechanisms used by the brain to analyze and perceive meaningful sounds is yet to be discovered*.

2.2 Introduction to Cochlear Implants

Deafness or hearing loss is a medical condition in which a person has reduced or inexistent capability to perceive sound vibration of the air and translate them into meaningful sound perception. There are many ear pathologies which can lead to partial or complete hearing loss, however, most frequently the hearing loss is caused by the destruction of the stereocilia or hair cells. This condition is known as sensorineural hearing loss.

When the stereocilia of the inner hair cells are destroyed, even though the vibrations are perfectly transmitted to the inner ear, the cochlea becomes incapable of translating sound vibrations into nerve firings, therefore profound hearing loss is installed. In most vertebrates, hair cells are incapable of regeneration. The only known solution for restoring the hearing sensation is through direct artificial electrical stimulation of the remaining auditory nerve fibers. During the last decades, many experiments were conducted to identify useful stimulation methods (Choi & Lee, 2012) (Blake S. Wilson, 2008) (Chen & Zhang, 2006) (Zeng, et al., 2015). The results of these experiments were translated into practical cochlear implant devices, designed to at least partially reproduce the nerve impulses generated by natural healthy hearing.

In the case of modern cochlear implants, stimulation impulses are delivered by an electrode array introduced into the cochlear duct, consisting of 12 to 24 individual electrodes positioned along the basilar membrane. The electrode array insertion can be executed through the round window of the cochlea, or through a hole drilled in the wall of the cochlear duct (cochleostomy). In all cases of commercially available

cochlear implants (Choi & Lee, 2012) (Blake S. Wilson, 2008) (Hochmair, et al., 2007), *each electrode is responsible with the delivery of stimulation impulses corresponding to an assigned frequency range*. Due to the positioning of the electrodes, distinct groups of auditory nerve fibers are stimulated mimicking the place coding strategy of the cochlea. In case of natural hearing, the basilar membrane produces a continuous decomposition of the sound vibration frequencies transmitted to thousands of hair cells stimulating tens of thousands of auditory nerve fibers (Meddis & Lopez-Poveda, 2010). In the case of cochlear implants, the resolution of place coding is limited by the number of electrodes. It was shown by experiments (Hochmair, et al., 2015) (Smith, et al., 2013) (Noble, et al., 2013), that at least 8 electrodes are required to provide useful speech understanding for non-tonal languages. In the case of music or tonal languages, like Chinese, place coding does not provide sufficient information. Music is perceived mostly as noise, and intonations and the pitch required to understand tonal languages are lost (Wang, et al., 2016) (Falcone & Bhatti, 2011) (Lin, et al., 2011)].

It is very difficult to increase the number of electrodes, partly due to the reduced diameter of the cochlear duct diameter, but mostly due to the interferences between the electrodes floating in highly conductive liquid filling the cochlea. The cochlea spiral geometry is a limiting factor too, constraining the depth of the electrode insertion. Depending on the manufacturer, electrode array lengths vary between 25mm and 31.5 mm, the length of the human cochlear duct being around 35 mm.

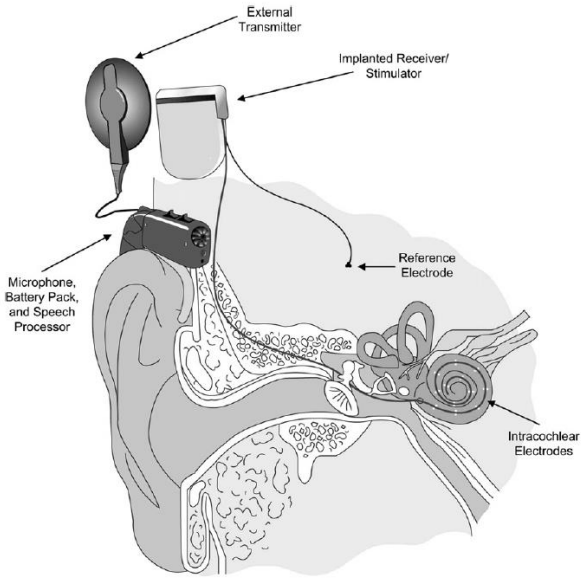


Figure 8: Components of a cochlear implant system (Wilson & Dorman, 2008)

Figure 8 depicts the typical parts and their positioning of a *cochlear implant system*. The system is composed of four main parts: the external sound processor; the external transmitter; the internal receiver and stimulator; and the electrode array.

The external processor is usually situated behind the ear or it can be body-worn. It contains one or more microphones to pick up the sounds, a battery pack to power the processor, and the processor itself to analyze the sounds picked up by the microphone and to compose the stimulation impulse trains to be delivered to the cochlea. The impulses generated by the speech processor are transmitted to the cochlear implant by the external transmitter part. The external transmitter uses an electromagnetic coupling to transfer the generated impulses and to provide electric power to the implanted receiver of the cochlear implant system. The implanted receiver is surgically placed in a deepening carved in the temporal bone of the skull and is situated right under the skin. The internal receiver delivers the received impulses through the electrode array.

As of now, there are only a few companies worldwide to provide cochlear implants, each providing a variety of electrode arrays and stimulation strategies: Cochlear Ltd. (Australia) with 24 electrodes, MED-EL (Austria) with 12 pairs of electrodes, Advanced Bionics (USA) with 16 electrodes, Oticon (France) with 20 electrodes, Nurotron (China) with 26 electrodes, and a very cheap yet low complexity cochlear implant is developed in India by DRDO (Defense R&D Organization) (Moctezuma & Tu, 2011) (Mahalakshmi & Reddy, 2012) (Zeng, et al., 2015) (Ghildiyal, 2016)

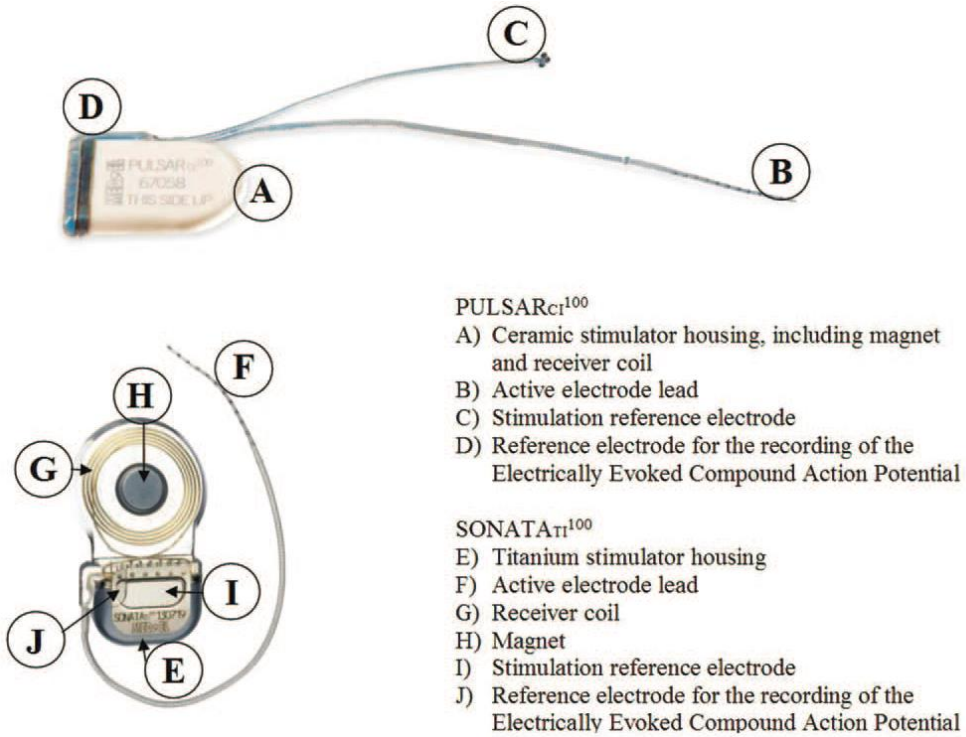


Figure 9: PULSAR and SONATA cochlear implants developed by MED-EL (Hochmair, et al., 2007)

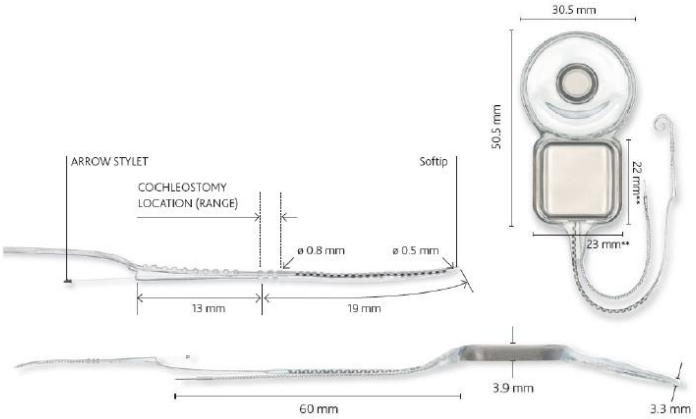


Figure 10: Nucleus 5 cochlear implant developed by Cochlear (Moctezuma & Tu, 2011)

Three different cochlear implant models are shown in Figure 9 and Figure 10, two from the MED-EL and one from Cochlear.

Despite the large differences in the number of electrodes, in terms of the hearing performance, all companies achieve about similar hearing performances, showing that simply increasing the number of electrodes does not provide hearing quality improvement.

To increase hearing performances, researches were conducted to develop new stimulation strategies making use of both place and temporal coding, and therefore providing stimulation patterns closer to the normal hearing (Fayad, et al., 2008) (Chen & Zhang, 2006) (Harczos, et al., 2013). Promising results were obtained, and some of the strategies are already implemented in commercially available implant systems (Hochmair, et al., 2007).

In Figure 11, the ideal insertion procedure of Cochlear's Contour Advance™ curved electrode array is demonstrated. In the case of curved electrode arrays, after the insertion, the electrodes are closely wrapped around the central axe (Modiolus) of the cochlea in order to better focus the stimulation current to the targeted auditory nerve groups. This approach enables the reduction of the distance between the electrodes allowing a higher total number of electrodes. The drawback of this design is that the electrode array length is limited and only the first half of the cochlea can be covered. To reduce the insertion related traumatisms, an insertion stylet is used. The electrode array body is based on a silicon tube with a closed tip, allowing the insertion of the stylet to fully or partially straighten it. At the beginning of the insertion, the stylet is completely inserted (a) followed by a gradual extraction of the stylet with the concomitant insertion of the electrode array (b, c). At the end of the insertion, the

stylet is completely removed and the electrode array is firmly wrapped around the modiolus.

Because of the partial coverage of the cochlea, only the area corresponding to the mid and high frequencies is stimulated, resulting in high-pitched sound perception. According to postlingually deafened patient reports, usually after long utilization of such an implant the brain manages to adapt to the new stimulation patterns and gradually reduces the perceived pitch. In our opinion, this is a good indication of the brain's ability to adapt to various types of non-natural stimulation patterns.

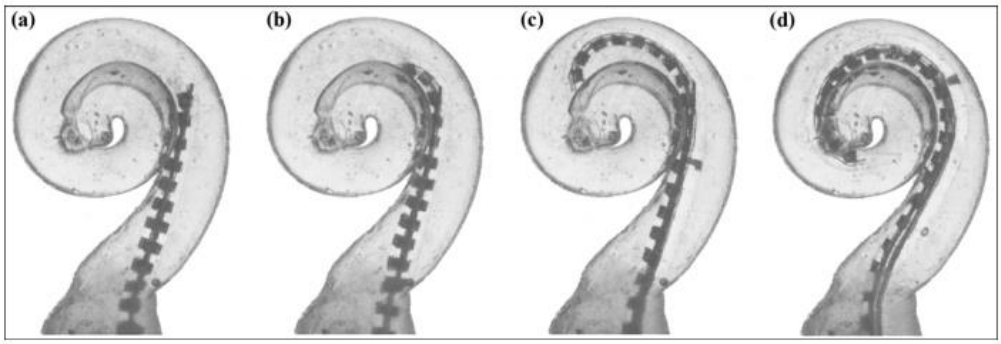


Figure 11: Ideal insertion of the Cochlear Advance ™ curved electrode array, using insertion stylet (Rebscher, et al., 2008)

Another electrode design approach is shown in Figure 12. In this case, the aim is to provide full cochlear implant coverage with the cost of a higher electrode to nerve distance. In this case, a straight electrode is used to allow a deeper insertion by guiding the electrode along the exterior wall of the cochlear duct. Because of the higher distance between the electrodes and the nerves, to induce higher current spread in the cochlea, a lower number of individual electrodes is used. However, to achieve better defect tolerance each electrode is doubled, providing the same stimulation impulses on both sides of the electrode array.

(a) (b)

Figure 12: MED-EL Standard Electrode array with (a) partial insertion and (b) full insertion (Dhanasingh & Jolly, 2017)

Although the number of stimulation electrodes is half compared to the Cochlear curved electrode, the design studies show that, due to the full cochlear coverage, it does not impact the quality of the perceived sounds (Hochmair, et al., 2007) (Hochmair, et al., 2015) (Taitelbaum-Swead, et al., 2005).

2.3 STIMULATION STRATEGIES

The most critical factor in determining the hearing quality is the stimulation strategy used to transform sounds into electric impulses. As it was already presented, there are two main coding strategies employed in the natural hearing process: Place coding; and Rate coding. Each cochlear implant provider has developed their own sound preprocessing and electric stimulation strategies, but all strategies are heavily relying on place coding. Figure 13 provides a *general overview and classification of the existing stimulation strategies*. The stimulation strategies are divided in 2 major groups: Strategies focusing on Coarse features and strategies focusing on Fine features. Strategies focusing on coarse features aim to provide good speech understanding, while the strategies focusing on fine features are more general approaches trying to improve general hearing quality including music perception.

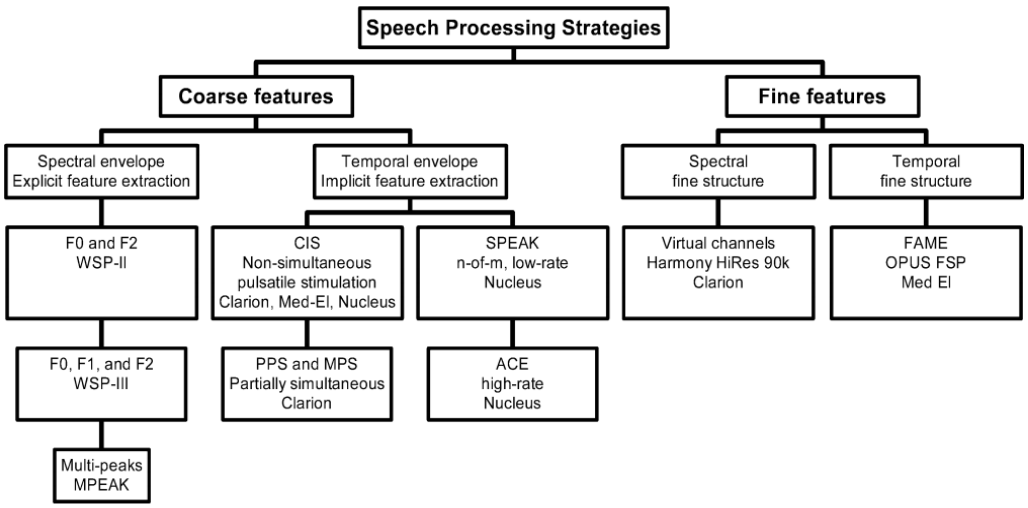


Figure 13: Classification of stimulation strategies (Zeng, 2008)

According to the manufacturers listing, the following signal processing strategies are used in commercially available cochlear implants:

Company	Stimulation Strategies
Cochlear	ACE – Advanced Combination Encoder CIS – Continuous Interleaved Sampling MP3000 / PACE – Psychoacoustic ACE SPEAK – Spectral Peak Strategy
MED-EL	HDCIS – High Definition CIS FSP/FS4/FS4P – Fine Structure Processing
Advanced Bionics	CIS – Continuous Interleaved Sampling HiRes/HiRes120 – High-Resolution Strategy MPS – Multiple Pulsatile Sample
Oticon	Crystalis/MPIS – Main Peak Interleaved Sampling
Nurotron	CIS – Continuous Interleaved Sampling APS – Advanced Peak Selections Symphony - Virtual channel strategy

Figure 14: Processing strategies used by manufacturers (Wilson & Dorman, 2008) (Choi & Lee, 2012) (Somek, et al., 2006) (Zeng, et al., 2015)

In the following, the most widespread stimulation strategies will be described.

2.3.1 Feature Extraction Strategies

Feature extraction strategies are speech perception oriented, aiming to provide necessary information for useful speech perception. They are specially designed to extract one or more key features of the speech, in order to facilitate easy recognition of the words and sentences. Both place and rate coding methods are employed in these strategies, although natural hearing perception is not targeted. Typically, extracted features are:

- **Fundamental Frequency (F0)** – The base periodicity of the speech signal. Can be determined only in the case of voiced sounds, especially vowels.
- **Formants (F1, F2) –** First and second peaking harmonics of F0, if applicable.
- **Envelope –** envelope of the original signal or a frequency component

2.3.1.1 F0/F1/F2 Strategy

In speech perception, the most important information to present are the fundamental frequency of the speech (F0), and the first two formant frequencies (F1 and F2). Figure 15 shows the implementation of the F0/F1/F2 strategy. The input from the microphone is conveyed to two bandpass filters (BPF) and a lowpass filter (LPF), through an automatic gain control (AGC) circuit.

The low pass filter, with the cutting frequency set to 270 Hz, is used to extract the audio signal frequency band, where the speech fundamental frequency (F0) is mostly situated. F0 is determined with a zero-crossing detector (ZCD) and it is used to modulate the rate of the stimulation pulse generator.

The electrodes are divided in two groups dedicated to the first and second formants. To determine the frequency of F1 and F2, a similar setup is used as in the determination of F0. The signal is filtered through two bandpass filters, with a range between 280 Hz and 1 kHz for the F1, and between 1 kHz and 4 kHz for F2. In addition to the ZCDs, two envelope detectors are connected to the output of the bandpass filters. Using the frequencies of F1 and F2, one electrode is selected from each group and the stimulation pulse is generated considering the magnitude of the detected envelopes (A1, A2). The rate of the stimulation is determined according to the F0 frequency (Mahalakshmi & Reddy, 2012).

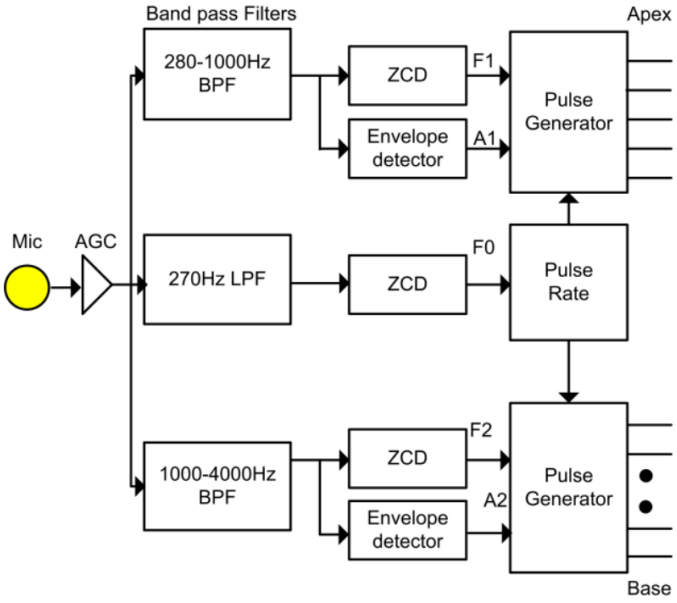


Figure 15: Block diagram of the F0/F1/F2 stimulation strategy (Mahalakshmi & Reddy, 2012)

In this setup, the following features are transferred through stimulation:

1. Frequency of F1 and F2 – using place coding
2. Amplitude of F1 and F2 – using stimulation intensity
3. Frequency of F0 – using rate coding

2.3.1.2 MPEAK Strategy

The F0/F1/F2 strategy emphasizes low-frequency information, which is useful for vowel identification, but the lack of information regarding high-frequency components makes consonant recognition very difficult (Mahalakshmi & Reddy, 2012). In order to improve the high-frequency representation, the MPEAK strategy was developed.

In Figure 16, representing the block diagram of the MPEAK strategy implementation, besides the components of the F0/F1/F2 strategy, three new channels are added. These three channels are built using band-pass filters (BPF) and envelope detectors in order to present high-frequency information to the cochlea, employing place coding. In case of the high-frequency channels, the extracted envelopes are used to modulate the output of a pulse generator with quasi-random intervals of an average rate of 250 pulses per second stimulation rates.

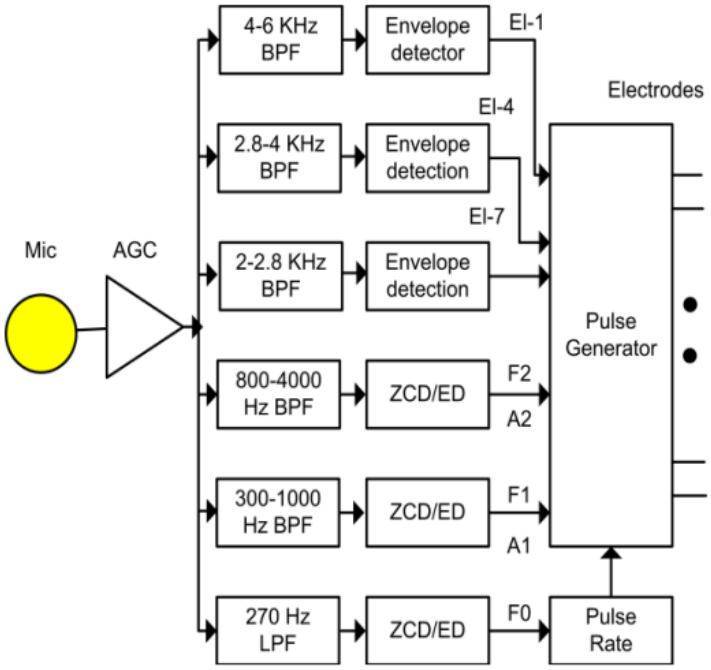


Figure 16: Block diagram of the MPEAK stimulation strategy (Mahalakshmi & Reddy, 2012)

2.3.2 Variations of N-of-M Strategies

N-of-M strategies are a family of basic encoding strategies which divides the audio frequency band in M channels, each channel being assigned to an electrode of the cochlear implant, and the stimulation is performed in impulse cycles. First, the audio signal is filtered into M frequency bands; at each cycle, the amplitude of the current envelope of each frequency band is determined. The first N bands with the highest amplitude are selected, and impulses with corresponding amplitudes are generated at the corresponding N electrodes. The impulses are sent individually in a sequence, in order to avoid channel interferences. The amplitude of the impulses is determined using a non-linear, typically logarithmic mapping curve. This strategy aims to present rich information regarding frequency components of the sound signal, relying only on place coding. To ease perception and mitigate channel interference, less-significant spectral components are eliminated from the stimulation, while still providing good formant representation of the spoken language. (Nogueira, et al., 2005)

These strategies do not rely on the rate coding; therefore, a fixed stimulation rate is used. It provides good speech understanding of non-tonal language in adequate listening conditions, but highly impaired for intonation, melody and music perception.

Providing consistent results, variations of the N-of-M strategy were developed and adopted by most of the cochlear implant manufacturers: ACE, CIS, SPEAK, HDCIS, etc...

2.3.2.1 Advanced Combinational Encoder (ACE)

The ACE strategy is Cochlear's implementation of the N-of-M strategy. It uses 22 frequency bands, with the same number of intra-cochlear electrodes. Per each stimulation cycle, depending on the fitting parameters, between 6 and 10 electrodes are selected for stimulation (Choi & Lee, 2012) (Vondrasek, et al., 2008).

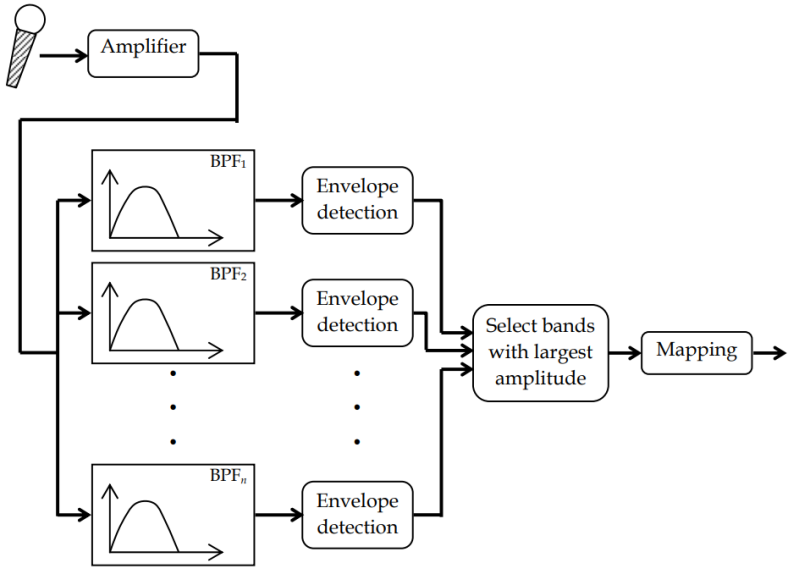


Figure 17: Block diagram of the ACE stimulation strategy (Choi & Lee, 2012)

Figure 17 shows the block diagram of the ACE stimulation strategy. The audio signal picked up by the microphone is fed into a bandpass filter corresponding to each of the 22 channels. Envelope detection blocks are used to detect the momentary amplitudes of the filtered signals and a band selector is used to identify the bands with the highest amplitude. After the bands with the highest amplitudes are selected, a mapping function is used to translate the amplitudes into stimulation intensities.

2.3.2.2 Continuous Interleaved Sampling (CIS)

The CIS strategy is a particular configuration of the N-of-M strategy, where N is equal M. In this case, in each stimulation cycle, all electrodes are stimulated. As it is shown in Figure 18, this strategy is implemented with a number of completely independent signal channels, consisting of a band-pass filter (BPF), a full-wave rectifier, a lowpass filter (LPF) to generate the signal envelope, and a nonlinear mapping function. The output of the mapping function is used to modulate a fixed rate impulse generator with typical stimulation rate varying between 250 Hz and 1500 Hz. The phases of the impulse generator are shifted in, so these impulses from different channels do not overlap, hence eliminating channel interactions and current surge between channels.

All major cochlear implant manufacturers are providing an implementation of the CIS stimulation strategy.

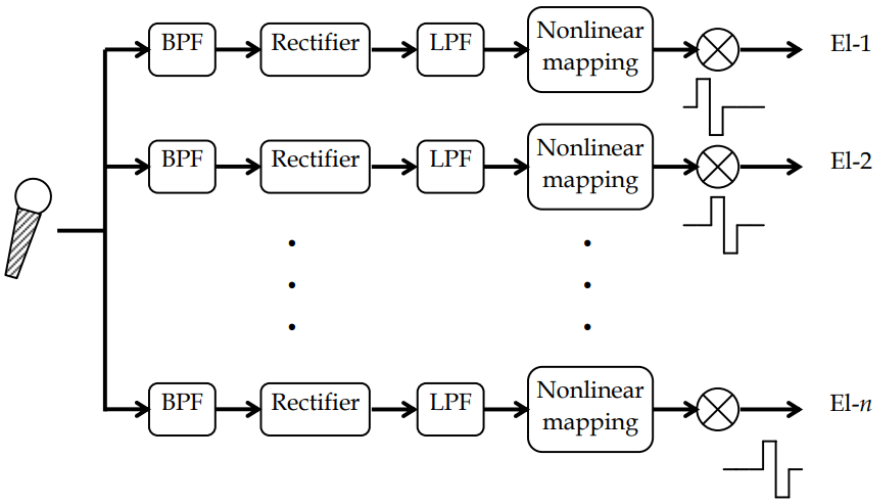


Figure 18: Block diagram of the CIS stimulation strategy. (Choi & Lee, 2012)

2.3.3 Fine Structure Strategies

The previously presented stimulation strategies are designed to transmit important, but rather coarse information regarding the sound structure. The aim is to transmit enough information through a low number of electrodes to allow useful speech recognition. According to experiments (Hochmair, et al., 2015) (Smith, et al., 2013) (Noble, et al., 2013), to easily understand non-tonal languages, it is enough to transmit the envelope of the speech signal and the approximate frequencies and envelopes of a few formants. The stimulation is delivered typically through 12 to 22 electrodes mapped to the same number of frequency bands, however, there are around 30.000 nerve fibers to be stimulated.

Given the low number of electrodes, advanced stimulation strategies were developed trying to deliver a better representation of the sound signal fine structure.

2.3.3.1 Stimulation using virtual channels (HiRes120)

To improve stimulation strategies based on place coding it is necessary to increase the number of individually addressable nerve groups. One way would be to increase the number of stimulation electrodes and reduce channel interferences. There are ongoing researches to develop high-density electrode arrays using thin-film technologies (Xu, et al., 2019), but as of now, there are no practical systems available. A technically feasible approach to allow better resolution in places coding is obtained using *current steering*. In case of current steering, targeted stimulation

impulses are delivered using two or more electrodes together to shape the form of the electric field inside the cochlea, providing electric field peaks at regions between the stimulating electrodes.

Using the current steering technique, it is possible to create so-called virtual channels and increase the spectral resolution delivered to the cochlea, while keeping the number of electrodes unchanged.

In Figure 19 the block diagram of the HiRes120 stimulation strategy is depicted, which was developed by Advanced Bionics and relies on current steering to provide up to 120 virtual channels (Choi & Lee, 2012). To achieve current steering, each stimulation impulse is delivered using two neighboring electrodes, balancing the intensity of the electric stimulation according to the nerve area targeted. Because the cochlear implants provided by Advances Bionic are using 16 electrodes, 15 neighboring electrode pairs can be used to deliver stimulation with current steering, therefore, the signal is divided into 15 frequency bands. At each frequency band, a Hilbert Envelope detector is used to determine the stimulation intensity, and a Spectral Peak Locator is used to identify the peak frequency of the filtered signal. Based on the envelope and the peak frequency, stimulation impulses are delivered concurrently to the two electrodes corresponding to the frequency band.

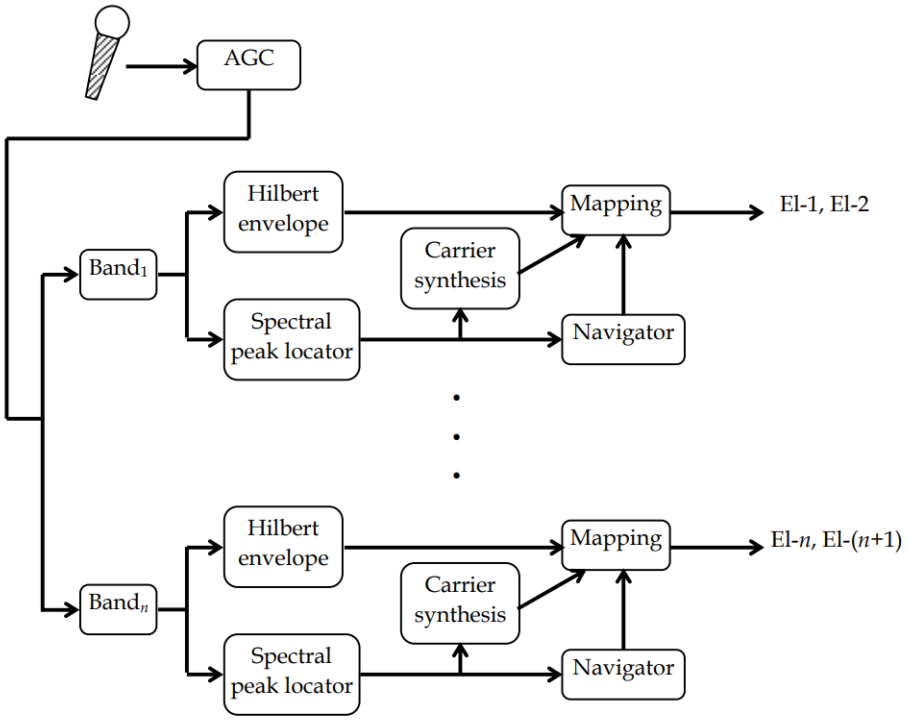


Figure 19: Block diagram of the HiRes120 stimulation strategy. (Choi & Lee, 2012)

2.3.3.2 MED-EL's Fine Structure Processing (FS4P)

Another approach to improve place coding-based stimulation strategies is provided by MED-EL, and it combines the CIS strategy with the rate coding approach for low frequencies. In case of the FS4P strategy the rate coding is available on the first four apical electrodes. The electrodes are still mapped to individual frequency channels, and the stimulation frequency is still modulated by the filtered signal amplitude, but instead of impulses with constant stimulation rate, as it is used in CIS, impulse trains are generated each time the filtered signal crosses 0. Because the stimulation impulses are synchronized with the phase of the sound signal, in case of bilateral implantation, sound direction identification using phase differences is possible too (Zirn, et al., 2015).

Figure 20 shows the theoretical mapping of the band filtered sound signals to stimulation impulses, while Figure 21 exemplifies an actual recording of signals generated by an Opus 2 cochlear implant processor.

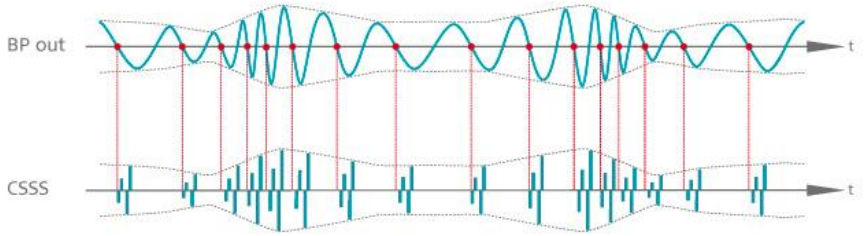


Figure 20: Temporal fine structure encoding using zero cross detection modulated by envelope. (MED-EL(2), n.d.)

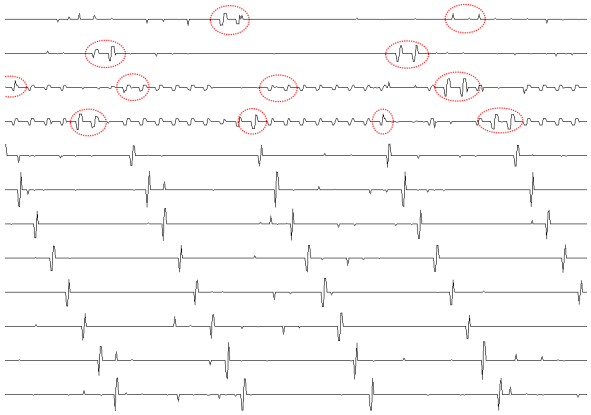


Figure 21: Simultaneous recordings of CI pulses for 12 channels (top to bottom: low to high frequencies) with 83 KSPS / Channel (Kuczapski & Andreescu, 2016)

It can be observed in Figure 21 that the impulse stimulations from the first 4 channels are generated using short impulse trains triggered with a rate corresponding to the frequency of the filtered signal of the channel, and with amplitudes modulated with

the envelope of the channel. Between the impulse trains, small oscillations are visible too. These oscillations are generated by the interaction with the high-frequency channels, but with amplitudes significantly below the activation thresholds of the nerves. The impulses recorded from the last 8 channels correspond to the CIS stimulation strategy and display fixed stimulation rates with an equal phase shift between consecutive channels

2.4 Current Challenges of the Cochlear Implant Technologies

In principle, the role of any cochlear implant is to restore hearing perception as close as possible to healthy natural hearing. Current approaches are limited by the physiology of the cochlea and by the available technologies, making it impossible to deliver stimulation signals accurately enough to replicate natural nerve firing patterns.

Various stimulation strategies were developed to deliver limited but useful information through a low number of electrodes, each approach having its strengths and weaknesses. Regardless of the used stimulation strategies, learning and adaptation of the patient is required, making it difficult to test and develop new strategies.

After assimilating a significant number of articles in this field, I found that nobody yet developed a *sufficiently accurate model to predict hearing quality*. This would be a key element in the feedback chain, in order to quantize and appreciate the quality of the artificial hearing supplied by the cochlear implant. Also, scientists fail to understand completely how electric impulses from the cochlear implant reach the neurological structure to be ultimately converted into sounds.

What all cochlear implant manufacturers today are doing, is to discharge some electrical current in the near vicinity of the modiolus, hoping that this will excite the acoustic nerve terminations in a manner that will allow enough discrimination between similar sounds. An unwanted outcome of this situation is that there are no models accurate enough to predict the interaction of the electrical fields generated by the inserted electrodes. This means that a predictive and reproducible fitting among several patients is not possible, this activity being more like a gamble than a deterministic approach. This explains, the rather large dispersion of the auditory performance of the patients wearing cochlear implants. To make this worse, many other factors contribute to such dispersion of auditory results like: age of implantation, auditory status (prelingual or postlingual), the duration and quality of speech therapy, or the environmental drive to verbal communication. All these reasons make it very hard to pinpoint and troubleshoot the real cause (or even causes) for a patient poor performance.

Stimulation strategies are not based on natural hearing models, because the processing power of a speech processor is very limited. Usually, it uses two or three Zn-Air type 675 batteries which must power both the speech processor itself and the implanted part of cochlear implant system. That’s why all these systems, regardless of the manufacturer, are very simplified models of the natural mechanical stimulation of the Organ of Corti.

The biggest setback is that there is no objective method to decide on fitting and stimulation strategies. That means that the fitter needs to be an exceptional professional with solid knowledge in many interconnected fields like: electronics (both digital and analog), acoustics, physiology of the hearing, and in case of prelingual little patients, a bit of psychology (in order to pick up the slightest reactions to a completely new type of stimulation). These requirements limit the number of people who could be trained as clinical engineers for cochlear implantation. The accessibility of fitting could be much higher if at least an important part of the process could be automated.

Unfortunately, the inserted electrode array is perceived by the cochlea, despite the biocompatible materials used, as a foreign object. Throughout its life, the body will try to build tissue structures in order to isolate the electrode array from the rest of the body. This affects significantly the paths of the electric currents determined by the charge applied to the electrodes, making the fitting itself a never-ending story. The lack of understanding of these biological processes prevents the *development of automated adjusting methods*.

3 SIMULATION MODELS OF THE EAR

The ear is the organ of hearing and balance responsible for translating mechanical vibrations and body accelerations into nerve signals interpreted by the brain. In this chapter, mathematical models are presented, describing the processes and mechanisms of the ear involved in the hearing process.

There are three delimited sections of the ear, named according to their position: outer, middle and inner ear. The sound wave travels through each section entering at the outer ear and exiting through the round window of the cochlea in the inner ear, each section having its own function in the hearing process.

From a mathematical point of view, the sections can be viewed as some complex transfer functions, characterized by their input signals, output signals, and transfer functions.

Ear Section	Input Signals	Output Signals
Outer Ear Conveys, shapes and amplifies the sound vibrations of the air.	**Primary** Air pressure function **Secondary** -	**Primary** Tympanic membrane Displacement **Secondary** -
Middle Ear Transduces air vibrations into liquid vibrations. Limits the amplitude of the vibrations, protecting the cochlea	**Primary** Tympanic membrane Displacement **Secondary** Stapedius Muscle Tension	**Primary** Oval-window membrane Displacement **Secondary** -
Inner Ear (Cochlea) Spectral decomposition of the liquid vibrations. Translates mechanical vibrations into nerve impulses	**Primary** Oval-window membrane Displacement **Secondary** -	**Primary** Auditory Nerve Impulses **Secondary** Basilar membrane displacement Inner/Outer hair cell polarization Inner hair cell gain factor

Figure 22: Input and output signals of the ear sections.

3.1 OUTER EAR

The outer ear consists of the auricle and the ear canal. The function of the auricle is to collect, amplify and direct soundwaves into the ear canal, creating pressure waves perpendicular to the tympanic membrane.

In free field stimulation, the frequency spectrum of the sound wave arriving to the auricle, is heavily affected by the filtering effect of the body which might or might not be positioned between the source of the sound waves and the auricle. (Meddis & Lopez-Poveda, 2010)

Besides the filtering effect of the body, due to the shape and size of the auricle, a small segment of the frequency spectrum is eliminated while the sound waves are conveyed to the ear canal. This effect is caused by the interference between the sound wave entering directly the ear canal and those entering after being reflected by the different surfaces of the auricle causing phase cancellation. This notch effect is observed between 6 kHz and 10 kHz, and its center frequency and attenuation depend on the direction of the sound wave (Zhang, 2010).

Combining the filtering effect of the body and the notch effect created by the auricle, the spectrum of the sound waves conveyed to the tympanic membrane caries important information regarding the source of the sound.

Because the filtering effect of the body and the auricle depends on the position of the sound source relative to the position of the head and auricle, the model of the outer ear differs from one listening setup to the other.

A *Head-related transfer function* (HRTF) is a linear transfer function describing accurately the filtering effect of the body and the auricle, in a certain listening setup (e.g. sound direction). Because its linear nature, it can be determined by measuring the *Head-related impulse response* (HRIR) by placing a miniature microphone into the ear canal, close to the tympanic membrane.

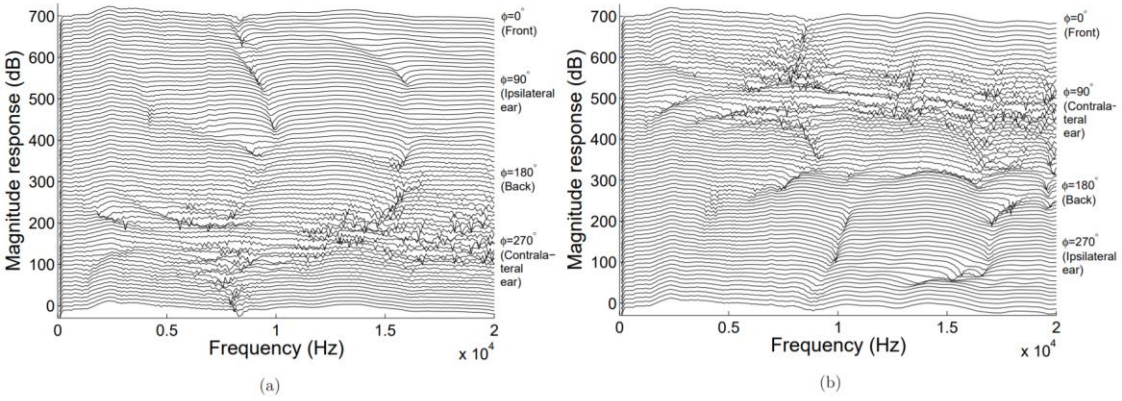


Figure 23: Frequency domain representation of recorded HRTF as a function of azimuth in the horizontal plane. (a) Left ear; (b) Right ear (Zhang, 2010)

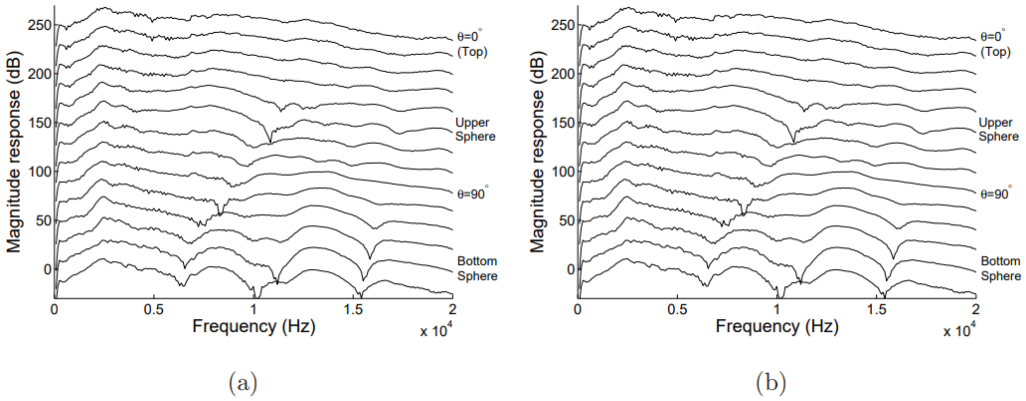


Figure 24: Frequency domain representation of recorded HRTF, as a function of elevation in the median plane. (a) Left ear; (b) Right ear (Zhang, 2010)

Figure 23 and Figure 24 exemplify a set of binaural HRTF recordings, with varying azimuth and elevation over the horizontal respectively median plane. It can be observed that the recordings from the left and right ear are in mirror when the *azimuth* is changed and similar when the elevation is changed. In the case of *elevation* changes, the characteristics are similar because keeping the sound source on the median plane guarantees that the sound waves will arrive at the ears with the same relative direction.

There are several studies (Miller & Matin, 2011) (Zhang, 2010) (Meddis & Lopez-Poveda, 2010) on the mathematical deduction of the HRTF based on the sound source positioning, and numerical model of the body and ear. These studies aim to develop models and technologies to produce virtual surround sounds mainly for entertainment or virtual reality setups, however, to study the hearing perception of recorded HRTFs is more adequate.

Because, except some particular models, the microphone of the cochlear implant is not situated in the ear canal, the *processed sound waves are not subject to the HRTF*, but rather follow the directional characteristics of the microphone. And, because, such fine details like frequency notches cannot be represented by current cochlear implant systems, *we will not include HRTFs in our models when studying the hearing perception of cochlear implant users*.

3.2 MIDDLE EAR

The function of the middle ear is two-fold. Because the sounds to be perceived are travelling through air, and the physical receptors (hair cells) are floating into perilymph, there was a need of an impedance adaptor between more fluid air into the less fluid perilymph. In millions of years of evolution, nature's solution to the problem

was the middle ear which connects the outer ear (tympanic membrane) to the inner ear (cochlea) through an ossicular chain composed of the smallest three ossicles in the human body: hammer (*malleus*), anvil (*incus*), and stapes. It also limits the amplitude of the vibrations by tensioning the stapedius muscle when high intensity sound is perceived.

In terms of signal processing, if the sound levels are moderate to low, the middle ear can be modelled as a linear system. In case of high intensity sound signals, the stapedius muscle tensions the ossicles attenuating the vibrations, making the middle ear transfer function non-linear.

Similar to the outer ear, the function of the middle ear is completely replaced by the cochlear implant processor. In order *to replicate the function of the middle ear, most cochlear implant systems implement an automatic gain control (AGC) strategy.*

3.3 INNER EAR

The inner ear (cochlea) is the "transducer" part of the ear. Its main function is to translate mechanical vibration, conveyed by the outer and middle ear, into electric impulses carried by the specialized nerve structures to the brain, to be interpreted as meaningful sound perceptions. In terms of signal processing the structures of the cochlea are highly nonlinear.

Figure 5 shows the schematic representation of the cochlea. The basilar membrane vibrates in response to the sound waves traveling through the intracochlear fluid, each longitudinal position resonating with a certain frequency band, and thus decomposing the sound vibration into frequency components, stimulating the inner and outer hair cells. The inner hair cells are responsible for generating nerve impulses as a response to the mechanical stimulation, while the outer hair cells are answering to the stimulation with contraction and elongation in resonance with the basilar membrane, amplifying its vibration. This positive feedback loop between the basilar membrane and the outer hair cells gave rise to a highly non-linear response function.

According to the physiology of the cochlea, the following signals are of interest:

- Displacement or Velocity of the Oval Window;
- Displacement or Velocity of points on the basilar membrane;
- The polarization of the hair cells, as a response to the mechanical stimulation;
- Nerve impulses.

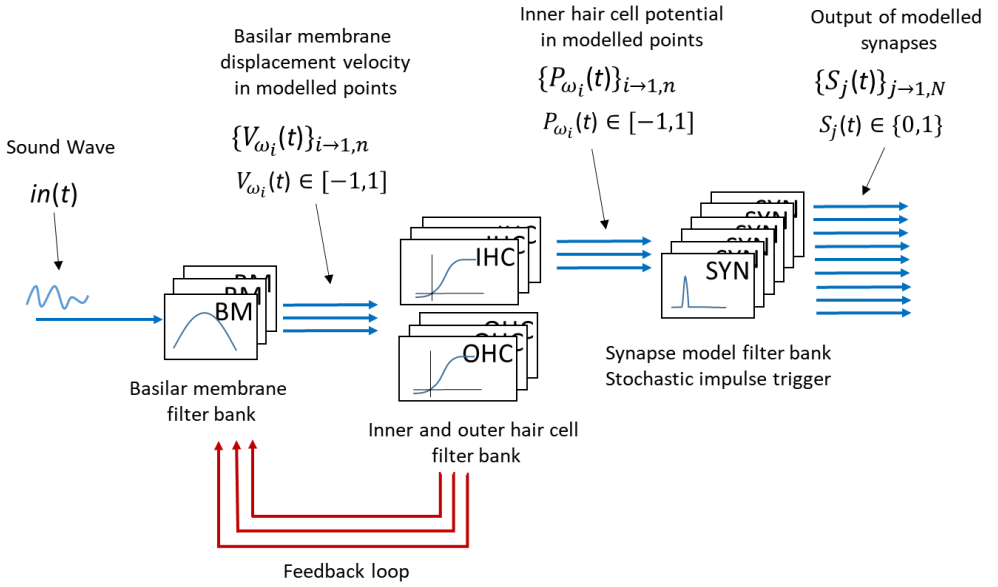


Figure 25: General physiological model of the cochlea, where n is the number of basilar membrane segments modelled and N is the total number of synapses modelled.

In Figure 25 a general simulation model is synthesized according to the physiology of the cochlea. In order to model all important aspects, we have defined the following signals, representing physiological values:

- $in(t)$ – Displacement velocity of the membrane of the oval window (equal to the displacement velocity of the stapes ossicle)
- $V_{\omega_i}(t)$ – A series of signals representing the normalized displacement velocity of the basilar membrane in modelled points
- $P_{\omega_i}(t)$ – A series of signals, representing the normalized average potential of the inner hair cells (IHCs) in the vicinity of the modelled points of the basilar membrane
- $S_j(t)$ – A series of signals, representing the *firing state of each individual auditory nerve*. Each $P_{\omega_i}(t)$ signal to an individual set synapse (SYN) elements resulting that $N=n*k$, where k is the number of SYN elements connected to the output of a single IHC element

In this setup, the basilar membrane is divided into a finite number *(n)* of segments. Each segment, having the same input *in(t)*, is individually simulated by a dedicated basilar membrane filter (BM), inner and outer hair cell filter (IHC, OHC) connected in a feedback loop, *resulting a series of signals* $P_{\omega_i}(t)$ *representing the average potential of the inner hair cells*. The average cell potential $P_{\omega_i}(t)$ of each segment is conveyed

to a set of synapse models (typically 20 synapse models per segment) in order to produce a final series of *N* signals $S_j(t)$ representing the resulting nerve impulses.

The nature of the feedback loop signal is not defined, as it can vary from implementation to implementation.

3.3.1 Gammatone Filter Bank and the Basilar Membrane

The Gammatone filter is a linear band pass filter designed to describe the displacement of an individual position along the basilar membrane (Patterson, 1992) (Meddis & Lopez-Poveda, 2010). The impulse response of the Gammatone filter is the product of a sinusoidal tone and a gamma distribution function shown by the equation below (Patterson, 1992):

$$g(t) = at^{n-1}e^{-2\pi bt}\cos(2\pi f_c t + \Phi) \qquad (1)$$

where *fc* is the center frequency of the filter while Φ is the phase shift, and *a* is the amplitude. There are two primary parameters of the filter: *n* – the order of the filter; *b* – the bandwidth of the filter. These two parameters can be used to calibrate the filter to match the characteristics of the basilar membrane of different species. For humans, *n* should be in the range of 3-5 and *b* should be selected depending on the *fc* in order to match the Equivalent Rectangular Bandwidth (ERB) of the auditory filter determined experimentally (Patterson, 1992) as:

$$ERB_{f_c} = 24.7\left(\frac{4.37\,f_c}{1000} + 1\right) \qquad (2)$$

Figure 26 depicts the impulse response of a 4th order gammatone filter, with the center frequency set to 1 kHz and the bandwidth to 125 Hz.

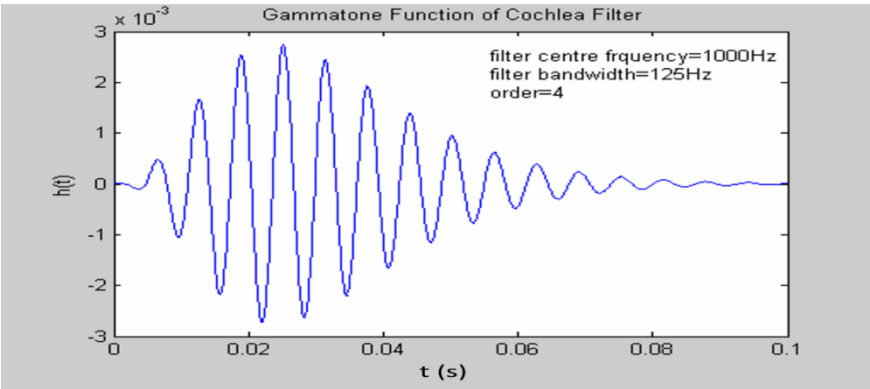


Figure 26: The Gammatone function (Yushi Zhang, 2006)

As it was described in the previous section, different positions of the basilar membrane resonate with different sound frequencies, basically decomposing the sound frequencies into individual spectral components. In Figure 27, it is demonstrated with an FM sweep signal, how a section of a certain frequency is detected by a gamma tone filter with the corresponding frequency.

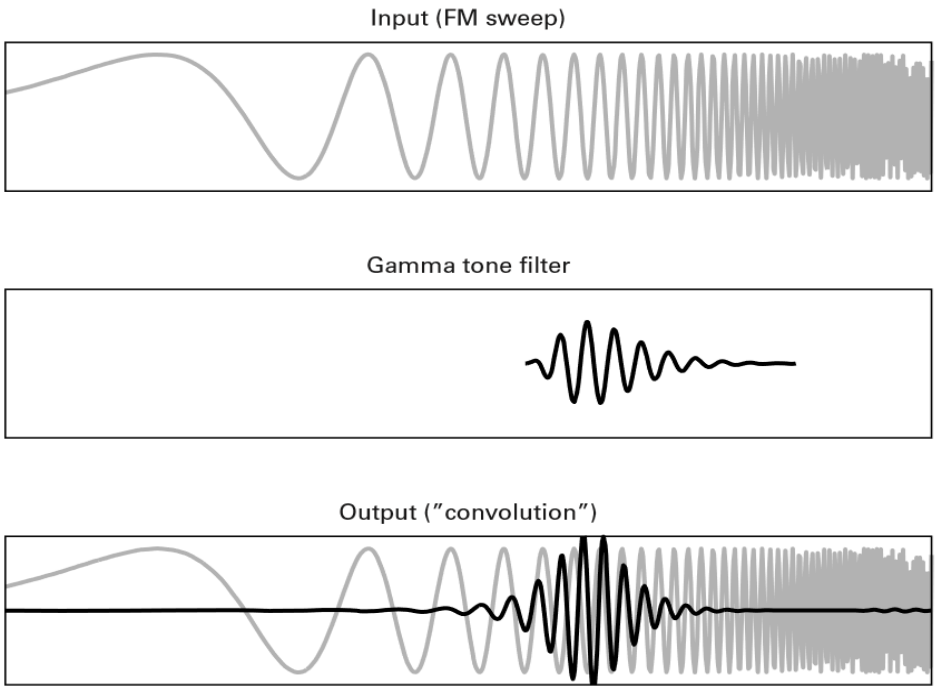


Figure 27: An FM sweep filtered by a gamma-tone filter (Schnupp, et al., 2011)

The complete basilar membrane can be modelled by a series of gammatone filters with exponentially increasing f_c, each individual filter estimating the displacement of the corresponding position along the basilar membrane. Figure 28 shows the frequency response of 16 gammatone filters with central frequencies below 7 kHz, whereas Figure 29 depicts the simulation results of a filter bank comprised of 49 gammatone filters with center frequencies between 220 Hz and 4400 Hz receiving as input four cycles of a pulse train with a 8 ms period.

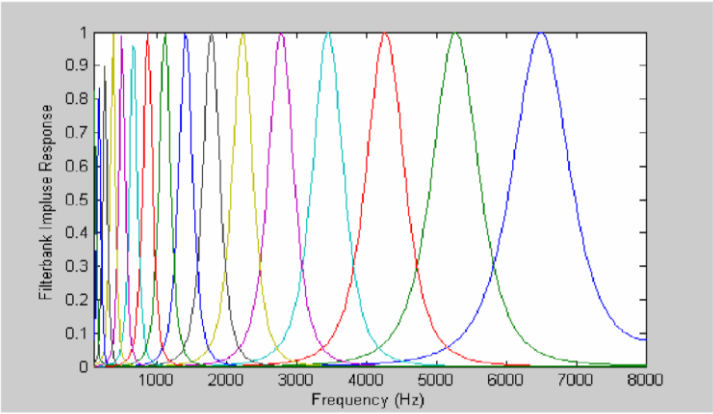


Figure 28: Frequency Response of a Gammatone Filter bank (Yushi Zhang, 2006)

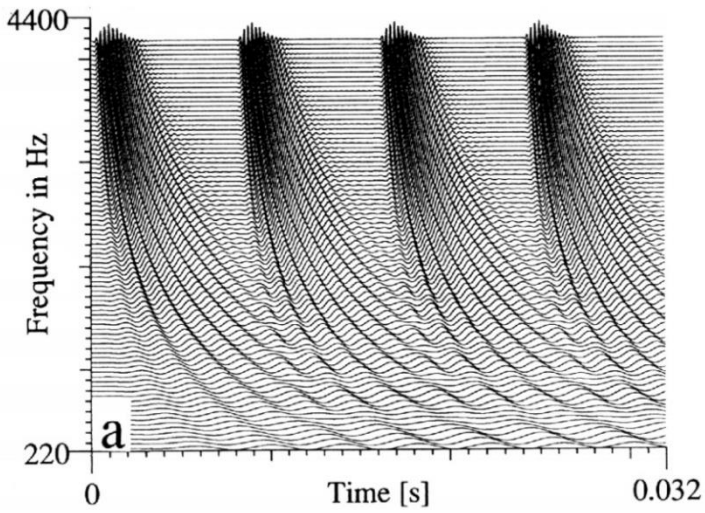


Figure 29: Output of a 49-channel filter bank as response to four pulse trains of 8ms (Patterson, 1992)

Multiple, more advanced basilar membrane models were developed around the gammatone filter model, aiming to better replicate the nonlinearity of the biological ear (for a healthy person). The diagrams of three such models can be seen in Figure 30. These models are more elaborated compared to the simple gammatone filter bank, yet we preferred to use a simpler solution because it meets the experimental needs within the scope of this work.

To say it in a clearer way, we couldn't find a reason to better simulate the natural hearing (because this is done every day by a healthy person), *but to simulate the hearing of patients with cochlear implants*. In this case, the gammatone filter bank is more than enough and therefore all the speech processors from all manufacturers use similar approaches.

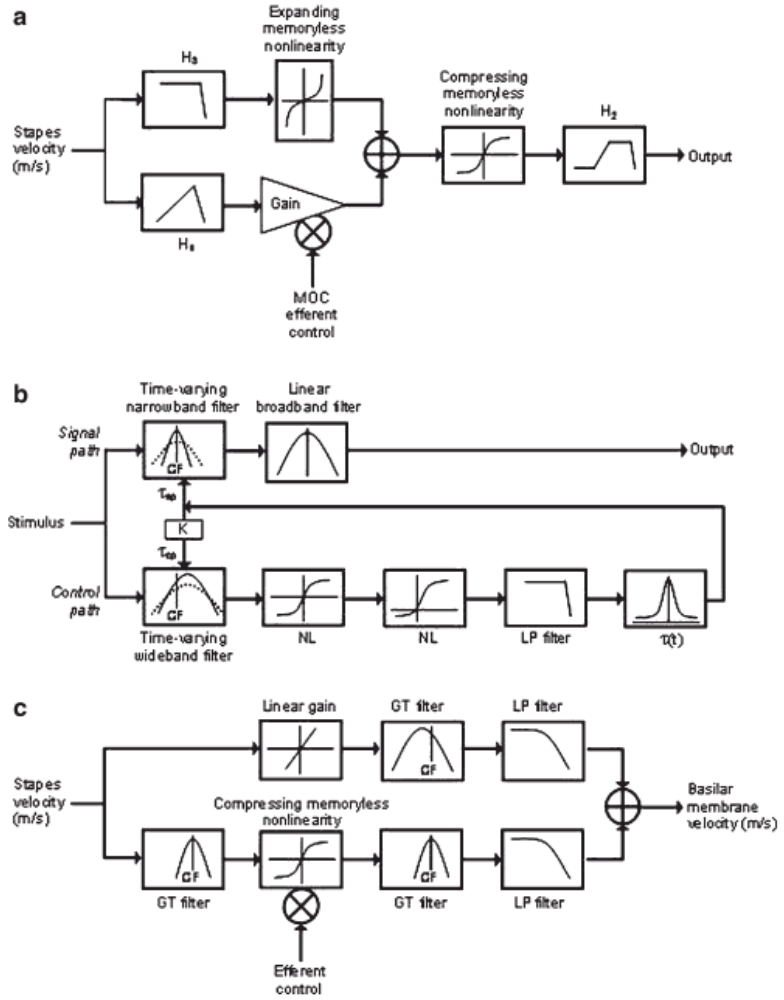


Figure 30: Comparative architecture of three phenomenological nonlinear BM models (Meddis & Lopez-Poveda, 2010)

3.3.2 Inner Hair Cells and Auditory Nerve Model

Inner hair cells (IHC) are sensory receptors responsible for transducing mechanical vibrations into electric potential variations in order to excite the synapses of the auditory nerves generating nerve impulses. In effect, an inner hair cell model shall describe the relation between the basilar membrane vibrations and the evoked changes in the intracellular potential of the hair cell. Consequently, the activity of the auditory nerve fibers can be modelled as a function between the inner hair cell potential and the firing patterns of the auditory nerve fibers.

In nature, the polarization of the inner hair cells is generated by the flow on positive ions between the interior and exterior of the hair cell. This ion in-flow is controlled by the deflection of the stereocilia situated on the top of the inner hair cell floating in the endolymph. When the stereocilia is deflected toward the tallest cilium, the non-specific ion channels are open, facilitating the inward flow of positive ions thus depolarizing the cell. In turn, when deflected in opposite direction, the ion in-flow is inhibited allowing the repolarization of the hair cell through the bottom of it surrounded by the perilymph. (Meddis & Lopez-Poveda, 2010)

The response characteristics of the inner hair cells were extensively studied and measured both in-vivo and in-vitro experiments (Meddis & Lopez-Poveda, 2010). It was shown that inner hair cells respond nonlinearly both to the DC and the AC component of the basilar membrane movement. The DC component of the input signal is transmitted similarly to a compression function having 2 dB/dB gain at threshold sound levels and <1 dB/dB at moderate to high sound levels. The transition of the AC component varies based on the stimulation frequency. For low frequencies, the cell potential reflects the input frequency too. However, rising of the stimulation frequency reduces the AC/DC ratio to a point where at higher frequencies the AC component is not transmitted anymore (Meddis & Lopez-Poveda, 2010).

Most commonly, the inner hair cells are modelled using either biophysical analogs (Lopez-Poveda & Eustaquio-Martin, 2006) or signal processing analogs (Zhang, et al., 2001) (Eriksson & Robert, 1999). The latter approach relies on a well-known signal processing element, therefore it is much easier to implement and evaluate. They also require very few parameters easing the calibration of such models. Although, the signal processing analogs does not reflect completely all characteristics of the inner hear cells, and does not use biophysical parameters explicitly, they are still extensively applied providing reasonably good approximation of the inner hair cells function.

A *simple yet functional model of the inner hair cell* is provided by (Zhang, et al., 2001) and depicted in Figure 31. In this approach the output of the basilar membrane model $P_{sp}(t)$ is connected to the input of a non-linear element of which output is conveyed into a low-pass filter. The first element emulates the nonlinear response of the inner hair cells to the DC component of the stimulus, while the low-pass filter attenuates the AC component when its frequency rises. The parameters of these filters are fine-tuned until the output of the model resembles the transfer characteristics of a real inner hear cell.

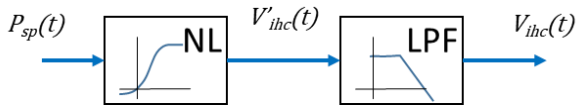


Figure 31: Inner hair cell model using signal processing elements (Zhang, et al., 2001)

The following equations describe the nonlinear transfer function:

$$\boldsymbol{V'_{ihc}(t) = A_{ihc}\left[P_{sp}(t)\right] * log\ (1 + B_{ihc} * \left|P_{sp}(t)\right|)} \quad (3)$$

$$A_{ihc}\left[P_{sp}(t)\right] = f(x) = \begin{cases} A_{ihc0} & , for\ P_{sp}(t) \geq 0 \\ -\frac{|P_{sp}(t)|^{C_{ihc}+D_{ihc}}}{3*|P_{sp}(t)|^{C_{ihc}+D_{ihc}}} * A_{ihc0} & , for\ P_{sp}(t) < 0 \end{cases} \quad (4)$$

where A_{ihc0}, B_{ihc}, C_{ihc}, and D_{ihc} are scalar parameters of the nonlinear function used to calibrate the model response. The values proposed for these parameters (Zhang, et al., 2001) are:

$$A_{ihc0} = 0.1,\ B_{ihc} = 2000,\ C_{ihc} = 1.74,\ D_{ihc} = 6.87 * 10^{-9} \quad (5)$$

The low-pass filter used in the inner hair cell model is a 7th order filter with a cutoff frequency of 3800 Hz.

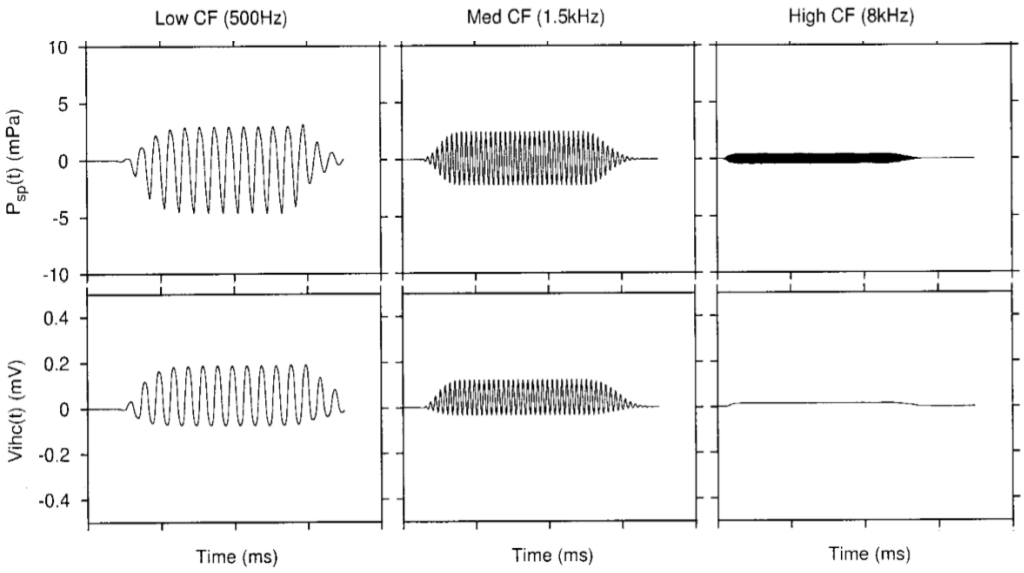


Figure 32: The responses of the inner hair cell model at different frequencies (Zhang, et al., 2001)

The above presented inner hair cell model reflects the most important transfer characteristics of the inner hair cell. Figure 32 shows the output signal of the inner

hair cell model at three different stimulation frequencies. It can be observed that the AC component of the input signal does fade away at high frequencies, but it is present at low and medium frequencies, these signals resemble quite well the in vivo measurements displayed in Figure 6.

The *firing of the auditory nerve* is triggered by the release of glutamate (a neurotransmitter) into the synaptic cleft between the auditory nerve fiber and the inner hair cell. Two factors are regulating for the transmitter release: the electric potential of the inner hair cell and the availability of the receptors in the presynaptic area. The rise in the electric potential of the inner cells increases the likelihood of transmitter release, in the same time any release of the transmitters reduces the availability of it making a consequent release more unlikely. As a consequence, when a stimulation appears after an *idle period*, a higher amount of transmitters will be released, causing a high density of nerve firings, which is quickly reduced to much lower level even if the stimulation intensity is maintained (Meddis & Lopez-Poveda, 2010).

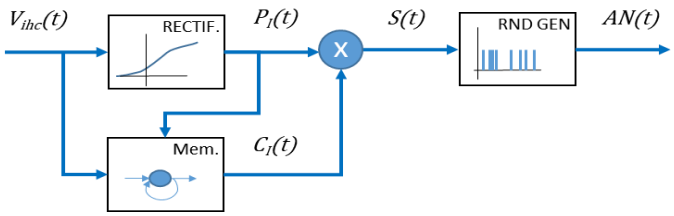


Figure 33: Auditory nerve firing model in (Zhang, et al., 2001)

This process can be modelled using a two-step approach as described in (Zhang, et al., 2001) and depicted in Figure 33. First, the probability of the nerve firing *S(t)* is calculated as a product of instantaneous probability $P_i(t)$ and the concentration of the available transmitters $C_i(t)$. $P_i(t)$ is estimated using a soft rectifier component, while $C_i(t)$ is estimated with more complex recurrent function reflecting the reduction of transmitter availability due to previous firings. In the second step, the nerve firings *AN(t)* are reproduced by a random impulse generator controlled by the probabilities given by *S(t).*

The probability of nerve firing without considering previous nerve firings is given by:

$$P_I(t) = p_1 * \log\left(1 + e^{p_2 V_{ihc}(t)}\right) \tag{6}$$

where, *p1* determines the immediate permeability at rest and p_2 is determined according to the central frequency (CF) corresponding to the nerve fiber location in the cochlea:

$$p_2 = \begin{cases} 1165, & for\ CF < 685\ Hz \\ -5430 + 1010 * \log(CF), & for\ CF \geq 685\ Hz \end{cases} \tag{7}$$

The detailed description of the transfer functions is provided in (Zhang, et al., 2001) and the C++ source code on the discrete-time implementation is provided online.

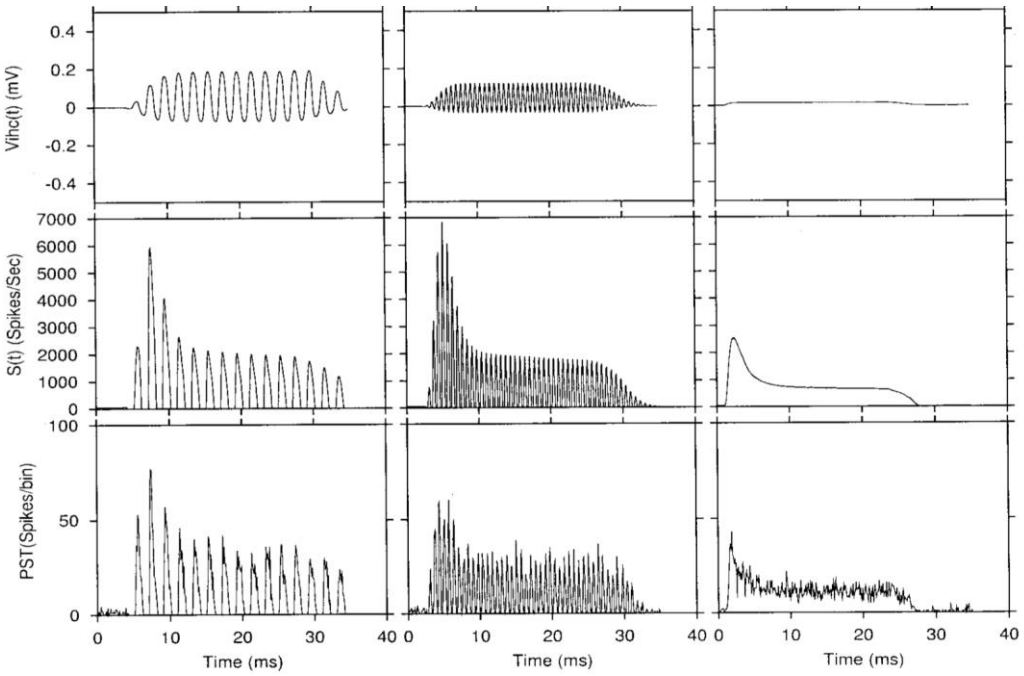


Figure 34: The responses of the auditory nerve firing model at different frequencies (Zhang, et al., 2001)

Figure 34 shows the output signal of the inner hair cell model at three different stimulation frequencies.

4 SIMULATION OF HEARING PERCEPTION

Based on clinical experiments, simulation software has been developed to *synthesize sounds as they are perceived by cochlear implant users (auralization),* making possible to assess expected hearing quality by technicians and fitting specialists (Mahalakshmi & Reddy, 2012), (Chilian, et al., 2011), (Loebach, 2007).

Although these simulation algorithms are based on observations from real-life experiments, the resulted synthesized sounds seem to underestimate the hearing quality of the patients. This is underlined by the opinion of *cochlear implant users who declare that they perceive the synthesized sounds unpleasant and lower in quality than the usual sounds*, and also by the fact that many implant users perform very well with musical instruments despite the result of the simulations where no music or melody appreciation is predicted (Drennan & Rubinstein, 2008), (Wang, et al., 2011).

In our option, the *discrepancy between predicted and experienced hearing quality is given by the fact that hearing simulations are done considering strictly sound perception patterns identified in natural hearing, and ignores the capability of the brain to adapt to new stimulation patterns*. This idea is reinforced also by the observation that most electric stimulation experiments were done on adult implant users who lost their hearing at a later stage in their life, already learning to interpret natural stimulation patterns (postlingual deafness). Therefore, the results of the experiments rarely reflect newly learned pattern recognition capability. In experiments with prelingually deafened patients, it is possible only to assess the capability of differentiation between different stimuli (Wang, et al., 2011), but not to compare the hearing sensations to the sound they were perceived before the hearing loss.

4.1 BASE MODEL OF THE NATURAL HEARING

As a first step in developing the auralization method, the *nerve impulse patterns created by the natural hearing are necessary to be reproduced*. *Human and animal auditory models* are well studied and described in the current literature (Schnupp, et al., 2011) (Schnupp, et al., 2011) (Meddis & Lopez-Poveda, 2010) (Lopez-Poveda & Eustaquio-Martin, 2006) (Tan & Carney, 2003) (Zhang, et al., 2001). These models focus mainly on the following aspects:

- Outer ear transfer function: how the sound waves are shaped by the structure of the ear and the ear canal.
- Middle ear transfer function: Transferring the vibrations form air to liquid and amplitude limitation of the stapedius reflex.
- Basilar membrane model: Estimates the displacement patterns of the basilar membrane.
- Hair-cell polarization models: Estimates the hair-cell polarization by the basilar membrane displacement.
- Auditory nerve firing models: Estimates the nerve firing patterns evoked by the hair-cell polarization

- Holistic cochlea models: Describe multidirectional interactions between stapes, basilar membrane, hair-cells and nerves, modelling complex feedback loops responsible for better sound decomposition and discrimination like active frequency filtering.

In the context of this thesis, a *simplified auditory model was studied and implemented* (Kuczapski & Andreescu, 2016). Its main purpose is to *reproduce the nerve impulses generated by the natural hearing*

process in *normalized* listening conditions. It was considered that complex feedback loops and active filtering capabilities, mostly implied in difficult listening conditions, are not important in the performed experiments, therefore the implemented simplified auditory model does not include such elements. *For CI patients, the electric pulses directly excite the auditory nerves and therefore, the sound way through the outer and middle ear is not present*. Considering the limitations of the implemented auditory model, all sound signals were normalized. The sampling rate of the signal used in experiments being 44.1 KHz. The graphical representation of the cochlea modeled by the simplified auditory model is depicted in Figure 35.

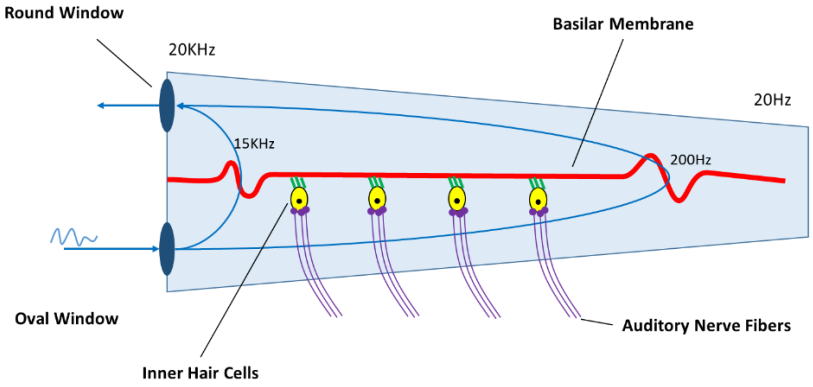


Figure 35: Representation of the cochlea structure according to the simplified auditory model, basilar membrane, hair cells and auditory nerve fibers

The simplified auditory simulation model consists of a series of filter banks (

Figure 36), each filter bank is responsible for modeling a different part of the cochlea. The first filter bank computes the displacement of the basilar membrane using gammatone filters (GT). The transfer function of an individual filter is described in the previous chapter.

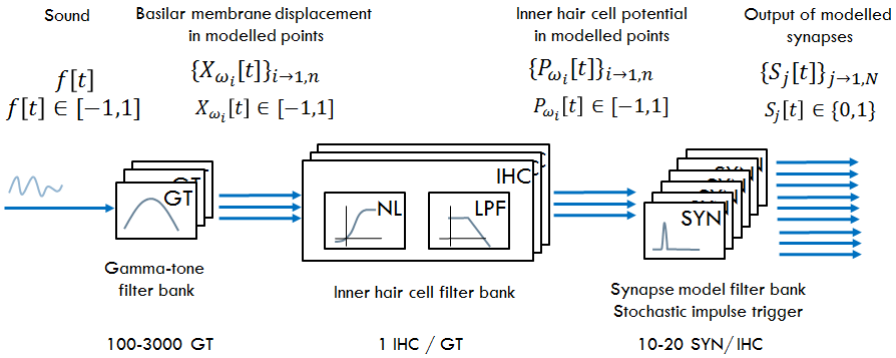


Figure 36: Simplified auditory model used for auditory nerve impulse study: gammatone (GT) filter bank for basilar membrane modelling; inner hair cell model (IHC) to estimate cell potential; stochastic synapse model (SYN) to generate nerve firing patterns

The output of the basilar membrane filter bank is fed to the inner hair cell (IHC) filter bank. Each modelled basilar membrane segment (e.g., gammatone filter) is connected to a single inner hair cell modelled by an asymmetric compression filter (NL) and a low pass filter (LPF). The first nonlinear filter reflects the tendency of the inner hair cells to polarize positively, while the second filter limits the frequency at which the potential of the cell can oscillate. Figure 37 depicts the polarization patterns of the simulated inner hair cells resulting from different input sounds. The X axis of the pictures represents the time of the simulation while the Y axis of the images represents the location along the basilar membrane. The color intensity shows the polarization potential: red color indicating positive charge, blue color indicating negative charge and black color indicating zero potential.

It can be easily observed that periodic polarization patterns are created at low frequencies (Figure 37a and b), and continuous polarization patterns at high frequencies (Figure 37c). These findings are in consensus with the in vivo / in vitro experiments found in (Schnupp, et al., 2011) (Meddis & Lopez-Poveda, 2010). When composite sounds are introduced in the model, the same polarization patterns are combined (Figure 37d).

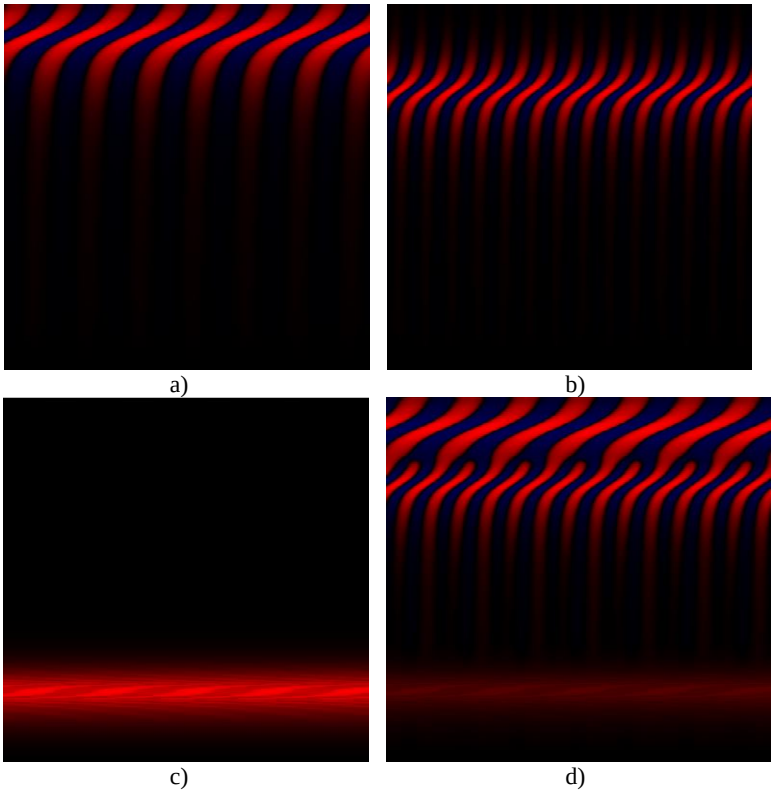


Figure 37: Polarization patterns of inner hair cells for input sounds at: a) 300 Hz; b) 600 Hz; c) 8000 Hz; d) composite sound of a), b) and c). X-axis represents the time, while Y-axis represents the distance from the oval window along the basilar membrane. Red color indicates positive charge and blue color indicates negative charge of the hair cells

The output of each inner hair cell model is connected to a sequence of synapse models (SYN), typically 10 to 20 synapses per inner hair cell. The synapse model consists of a stochastic binary impulse trigger and a recovery cycle to mimic real-life behavior of nerve fibers. All parameters and coefficients of the model were fine-tuned according to experimental results and values found in references (Zhang, et al., 2001), (Lopez-Poveda & Eustaquio-Martin, 2006), (Meddis & Lopez-Poveda, 2010). The resulting nerve impulse patterns of the 300 Hz + 600 Hz + 8000 Hz composite input signal are depicted in Figure 38. It is easy to observe that clear impulse patterns are generated according to the input signal frequency components, thus reconstruction of the original sound using pattern recognition and machine learning methods is more than plausible.

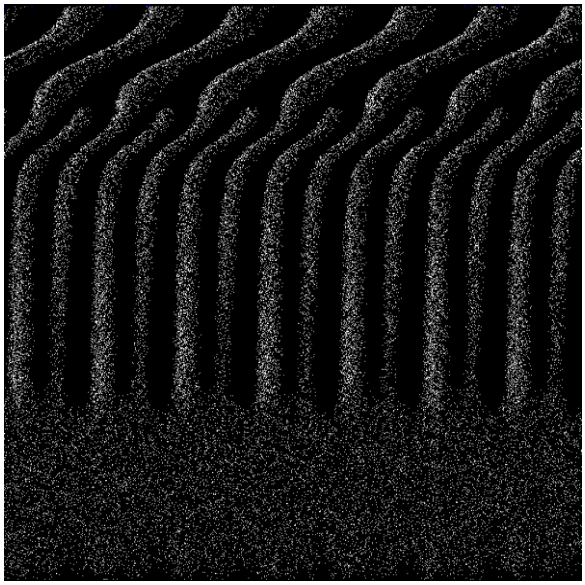


Figure 38: High resolution auditory nerve impulse map for a composite sound of 300 Hz +600 Hz +8000 Hz. X-axis represents the time, while Y-axis represents the distance from the oval window along the basilar membrane; white plot represents a nerve binary impulse

4.2 EXISTING AURALIZATION METHODS

The auralization of the hearing perception of cochlear implant users is a process where the output signals generated by a specific cochlear implant coding strategy are transformed back to audible sounds to be listened by people with normal hearing, demonstrating the sounds as they are supposed to be perceived by the implant users. The synthesized sounds present distortions and artifacts caused by the reduced number of implant stimulation electrodes which implies the transition of only a few spectral components (Loebach, 2007), (Chilian, et al., 2011), (Mahalakshmi & Reddy, 2012). We have reviewed existing auralization methods (Kuczapski & Andreescu, 2016) (Chilian, et al., 2011) (Loebach, 2007) and found that in principle only two methods were previously used. These two methods are briefly presented in the following sections.

i) Simple auralization method based on channel envelope detection only

This first auralization method, depicted in Figure 39*,* employs a *simplified model of the cochlear implant sound processing* (CI*). The input sound signals are filtered by a number of bandpass filters equal with the number of electrodes used by the cochlear

implant. For each channel, the filter output signal is conveyed into an envelope detection filter and a low pass filter that limits the dynamic of the envelope.

For each filtered channel, there is a carrier signal with the central frequency of the corresponding bandpass filter. The carrier signal can be as simple as a sinus, or can be obtained by filtering white noise (WNG) through a similar bandpass filter (BPF) as used for the corresponding channel filtering. Each carrier signal is modulated with the corresponding channel envelope and then summed up all together to form the output sound of the auralization model (Loebach, 2007), (Mahalakshmi & Reddy, 2012).

Although the synthesized sounds obtained with the noise carriers, respectively with sinus carriers are quite different, the speech understandability is similar in both cases. This is due to the fact that the *actual information extracted by the listener brain is mostly limited to the encoded envelopes, giving very limited information about the sound frequency components*. Lacking the frequency information eliminates also, almost totally, the possibility of music perception.

Such auralization method appreciates correctly the information coded by spatial cues, and predicts accurately the number of electrodes required for useful speech perception in quiet. On the other hand, it fails to reproduce the effect of temporal coding in low frequencies, and does not demonstrate tone discrimination capabilities observed in some well-trained cochlear implant users.

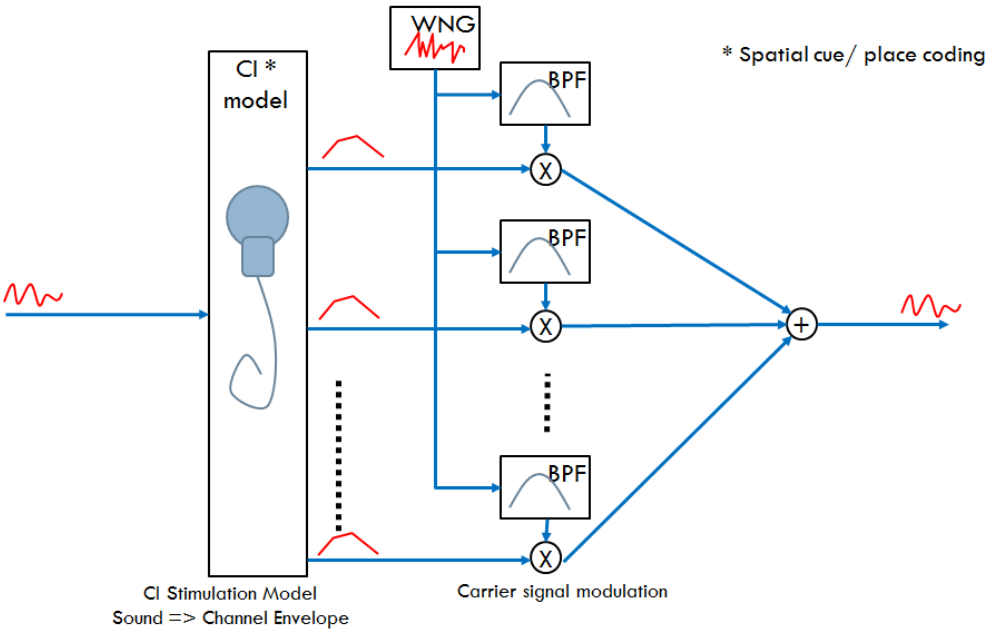


Figure 39: Simple auralization method based on channel envelope detection only: cochlear implant (CI*) model generates channel envelopes only; white noise

generator (WNG) and bandpass filters (BPF) used for carrier signal generation; carrier signal modulation and summation (Loebach, 2007)

ii) Auralization method based on channel envelope & frequency detection

In order to give more natural and richer hearing sensations, temporal coding of sound information is employed in several new coding strategies (Chen & Zhang, 2006), (Harczos, et al., 2013), (Nogueira & Buechner, 2012). Improvements to the auralization method have been developed (Chilian, et al., 2011) in order to estimate and also to demonstrate the increasing hearing quality of newly developed coding strategies. Instead of using channel envelopes to modulate carrier signals, the *cochlear implant (CI*) is completely modelled down to the electric impulses generated by electrodes*. The output signals, representing the electric impulses of each channel, are conveyed into a *current spread model to estimate interference between electrodes*. Then, the resulting signals are fed to a *perception model which estimates the loudness based on the current intensity, and the frequency (F_{center}) based both on the stimulation place and stimulation rate*. The frequency and loudness information are used to synthesize audible sound similarly to the previous auralization method. The schematics of this auralization model is shown in Figure 40.

This second auralization method considers both spatial and temporal cues and therefore provides better approximation of the perceived sounds. However, the learning and adaptation capabilities are not still modelled.

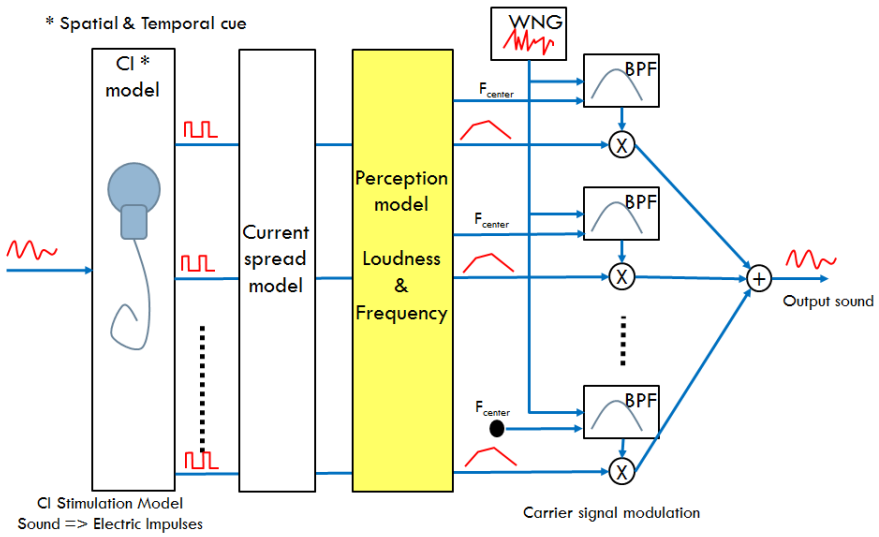


Figure 40: Auralization method based on channel envelope and frequency detection: cochlear implant (CI*) model generates electric stimulation patterns; current spread model to estimate nerve stimulation; perception model to detect envelopes and

frequencies; white noise carriers (WNG) and bandpass filters (BPF) with varying central frequency; carrier signal modulation and summation (Chilian, et al., 2011)

From a practical point of view, *auralization methods* can be used to *answer the following questions*:

- What amount of information is coded in the electric stimulation patterns?
- What will be extracted by the auditory nerves and the brain, considering natural hearing patterns?
- What can be extracted considering the limitation of the auditory nerves and the learning ability of the brain?

To be able to answer all these questions, the concept of a *novel auralization method* is proposed (Figure 41). It takes from a CI* stimulation model a series of electric stimulation impulses, calculates the current spread in the cochlea, and estimates the nerve impulses generated by the current with an auditory nerve firing model. The estimated neural impulses are conveyed into a pattern recognition algorithm based on machine learning, which turns the signals into a frequency domain representation of the perceived sound that is then converted back to sound.

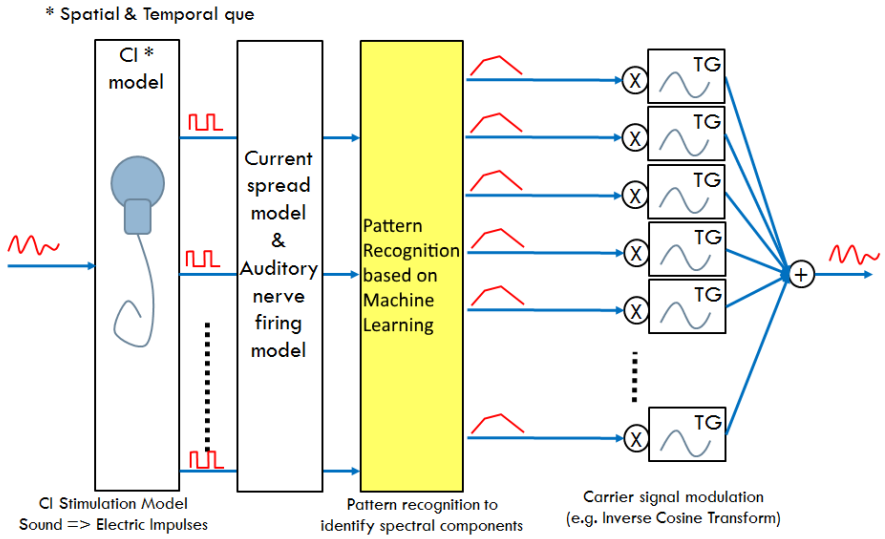


Figure 41: New auralization concept based on neural firing pattern recognition: cochlear implant (CI*) model generates electric stimulation patterns; current spread model to estimate nerve stimulation currents; auditory nerve firing model to generate nerve impulse patterns; Pattern recognition to estimate envelope magnitude of spectral components; tone generation (TG) and summation

Such an auralization approach can be used to simulate different hearing conditions.

Example 1: The pattern recognition algorithm could be trained to recognize neural impulse patterns of healthy hearing, resulting in a model that would estimate the hearing experience of newly implanted postlingually deafened patients.

Example 2: The pattern recognition algorithm could be retrained to recognize neural impulse patterns generated by the artificial stimulation, therefore the auralization of adapted hearing could be realized.

Following the principles of above depicted concept, suitable approaches for implementing the *pattern recognition* algorithm have been investigated. Initially, the utilization of artificial networks was considered (Kuczapski & Andreescu, 2016), but later on the initial implementation trials, it lead to a novel and more robust pattern learning an matching algorithm based on simple auto- and cross correlation methods (Kuczapski & Andreescu, 2017). This later approach not just allowed to identify the transmitted frequency components in the nerve firing patterns, but also it resulted in easily interpretable numerical characteristics to describe the expected hearing characteristics.

4.3 New Auralization Method using Autocorrelation based Pattern Recognition

Based on the conclusions shown in the previous section, a pattern learning and recognition algorithm was developed and presented in (Kuczapski & Andreescu, 2017) and in the patent (Kuczapski, 2015). This chapter presents the results of these developments and demonstrates how a simple *row-wise image autocorrelation of cochlear nerve firing patterns* can be used to develop a self-learning system, which is *completely agnostic of the CI coding strategy*, capable of *accurately predicting pitch perception characteristics of CI users*. Simulation results are shown using comparisons between: i) natural hearing, ii) hearing with CI using Advanced Combination Encoder (ACE) strategy, iii) hearing with real MED-EL Opus 2 CI processor.

4.3.1 Cochlear Nerve Firing Pattern Identification Based on Autocorrelation in Natural Hearing

In the case of natural hearing, the cochlea translates the sounds coming in form of pressure changes in the air, into intricate patterns of cochlear nerve firings (Schnupp, et al., 2011). By nature, the auditory system relies on two important cues: i) the position of stimulated nerves on the basilar membrane – *spatial cue*, and ii) the rate of stimulation – *temporal cue*. The temporal cues are only present for the spectral components below 4 kHz (Schnupp, et al., 2011).

Figure 42 and Figure 43 depict the cochlear nerve firing patterns in the case of natural hearing, where the X axis represents the time and the Y axis - the stimulation location along the basilar membrane. For each nerve firing, a black dot is shown. Following the cochlea tonotopic structure, each position along the basilar membrane corresponds to a sound frequency, and the frequencies are ordered descending from the lower part of the image (high frequencies - at the cochlea base) to the upper part (low frequencies).

The cochlear nerve firing patterns were obtained using a *simplified auditory model* (Figure 36) (Kuczapski & Andreescu, 2016), where: i) the movement of the basilar membrane was approximated using gammatone filters, ii) the inner hair cells charges were modeled with asymmetric compression filters and low pass filters, and iii) the cochlear nerves were simulated using a stochastic synapse model. All simulation parameters were set to match the human auditory system (Meddis & Lopez-Poveda, 2010) (Lopez-Poveda & Eustaquio-Martin, 2006) (Tan & Carney, 2003) (Zhang, et al., 2001).

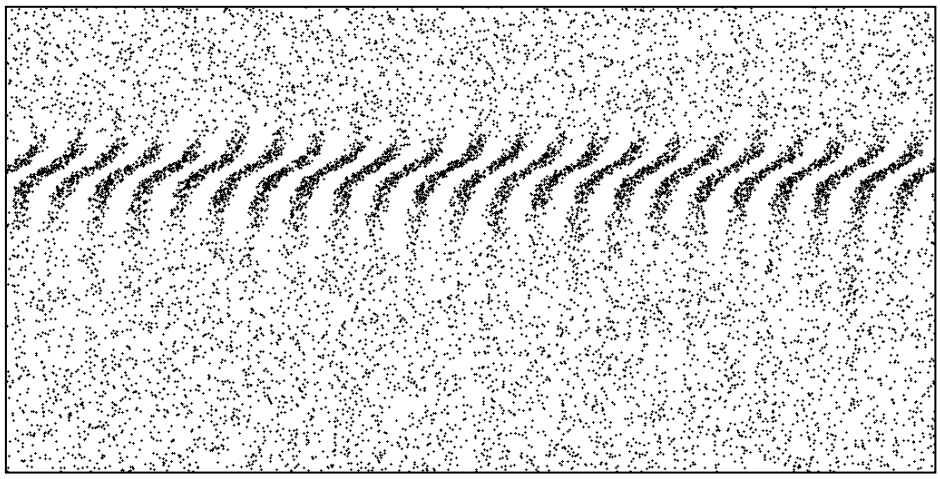


Figure 42: Natural cochlear nerve impulses firing pattern for 1 kHz tone input. Note1: X axis - time, Y axis - stimulation location along the basilar membrane corresponding to a specific frequency (high-frequency to origin). (Kuczapski & Andreescu, 2017)

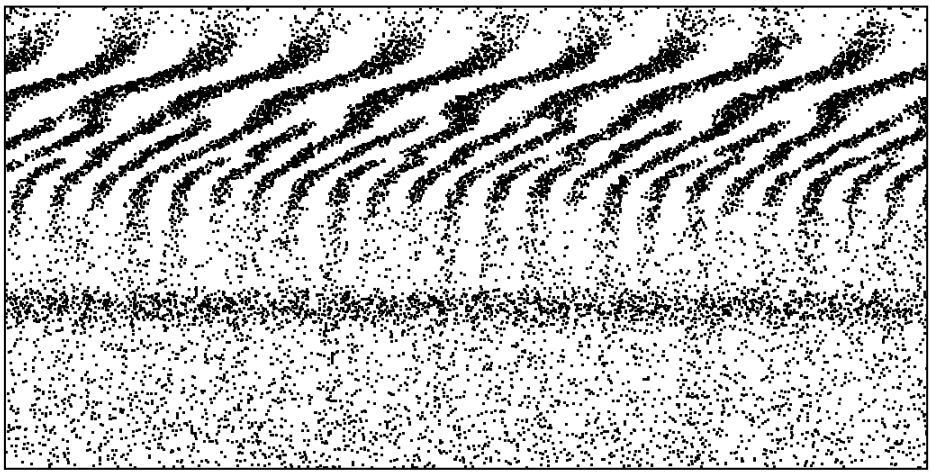


Figure 43: Natural cochlear nerve impulses firing pattern for mixed tone inputs at 427 Hz, 758 Hz, 1 kHz, 3.9 kHz; see Note1 as in Figure 42. (Kuczapski & Andreescu, 2017)

In Figure 42, a 1 kHz pure tone is used as input. The cochlear nerve firings are concentrated to rather broad band along the basilar membrane showing the sensitive area to the 1 kHz frequency. Also, the nerve firings are modulated with 1 kHz providing important temporal cues (Schnupp, et al., 2011) to the auditory system.

The Figure 43 shows the simulation result for a mix of pure tones at 427 Hz, 758 Hz, 1 kHz and 3.9 kHz. The first three tones are translated into a complex oscillating pattern of cochlear nerve firings, while the fourth tone at 3.9 kHz is mapped to a narrow continuous band of higher density firing.

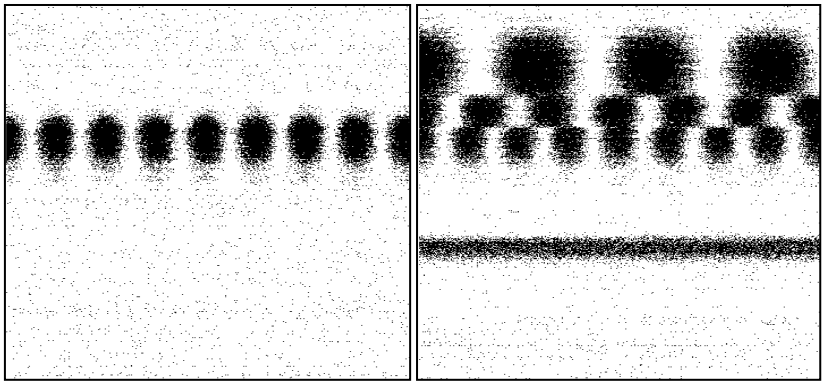
In the case of complex sounds, the analysis of the cochlear nerve impulses can be difficult in the raw form. To overcome the complexity of extracting frequency specific patterns, the *cochlear nerve firing patterns are transformed* by using the *row-wise autocorrelation* (8)

$$\mathrm{AC}_{\mathrm{row_i}}^{\mathrm{to}}(\mathrm{k}) = \sum_{\mathrm{t=to}}^{\mathrm{to+n}} \mathrm{SYN}_{\mathrm{row_i}}(\mathrm{t}) * \mathrm{SYN}_{\mathrm{row_i}}(\mathrm{t}+\mathrm{k}) \qquad (8)$$

where $AC_{row_i}^{to}(k)$ is the autocorrelation function of the k[th] row computed for the time interval [t_o, t_o+n], $SYN_{row_i}(t)$ is one row of the nerve firing pattern and n is the length in samples of the autocorrelation window.

The results are shown in Figure 44. *Both spatial and temporal cues can be easily observed* in both pure tone and mixed tone experiments. Moreover, the background noise, caused by the random nerve firings is considerably reduced. These patterns resulting from the row-wise autocorrelation are more suitable for machine learning approaches.

A process similar to this autocorrelation process is supposedly happening also in the human auditory cortex (Schnupp, et al., 2011) (Meddis & Lopez-Poveda, 2010), therefore this transformation is not an unnatural one.



a) *b)*

Figure 44: Row-wise autocorrelation pattern of cochlear nerve impulses for: a) 1 kHz signal; b) 427 Hz, 758 Hz, 1 kHz and 3.9 kHz mixed signal. Note2: X axis - time shift of the autocorrelation, Y axis - frequency related position. (Kuczapski & Andreescu, 2017)

4.3.2 Frequency Detectors from Cochlear Nerve Firing Patterns – Frequency Autocorrelation Masks and Amplitude Cross-Correlation Coefficients

Based on these experiments, a new method, to create reliable *audio frequency detectors for extracting the spectral components* of *perceived sound encoded in the cochlear nerve firing patterns,* is developed.

The detector for each frequency component f_i uses:

I) the *frequency specific autocorrelation mask*, and

II) the frequency specific *amplitude response* characteristic (i.e., the mapping curve) obtained from amplitude cross-correlation coefficients.

I) **Frequency specific autocorrelation masks** are extracted by a developed *supervised machine learning algorithm*. When a specific cochlear nerve firing pattern approximation model is used connected to the developed learning algorithm, the resulting patterns will be characteristic to this.

The following steps summaries the proposed learning algorithm:

1. The analyzed audio spectrum (220 Hz - 19 kHz) is divided into a logarithmic scale with ¼ of a semitone step, creating a *frequency list FL* of 312 f_i frequencies. This resolution is good enough to match most people's hearing resolution.

2. For each frequency f_i, the *learning algorithm* is computed as:

   2.1. Estimating the *cochlear nerve firing patterns* resulted from a random mix of pure tones with 50% chance that the frequency f_i is present in the mix, by using the analyzed stimulation strategy of CI, or the simplified auditory model of natural hearing (Figure 36).

   2.2. Subtracting the average of the row-wise autocorrelation patterns (8) where f_i is not present from the average of the row-wise autocorrelation patterns where f_i is present, resulting the f_i *frequency specific autocorrelation mask*.

   2.3. The length of the autocorrelation patterns is limited to 4 periods of f_i.

   2.4. The f_i frequency specific autocorrelation mask is *normalized* with a cut-off of 65%, so the intensities above 65% are mapped to 1, and the others are mapped to 0.

Figure 45 depicts the flowchart of the frequency specific pattern learning algorithm and the results are exemplified in Figure 46 showing the nerve firing autocorrelation masks obtained for 1 kHz and 6 kHz detection in case of a cochlear nerve firing pattern approximation model simulating natural hearing.

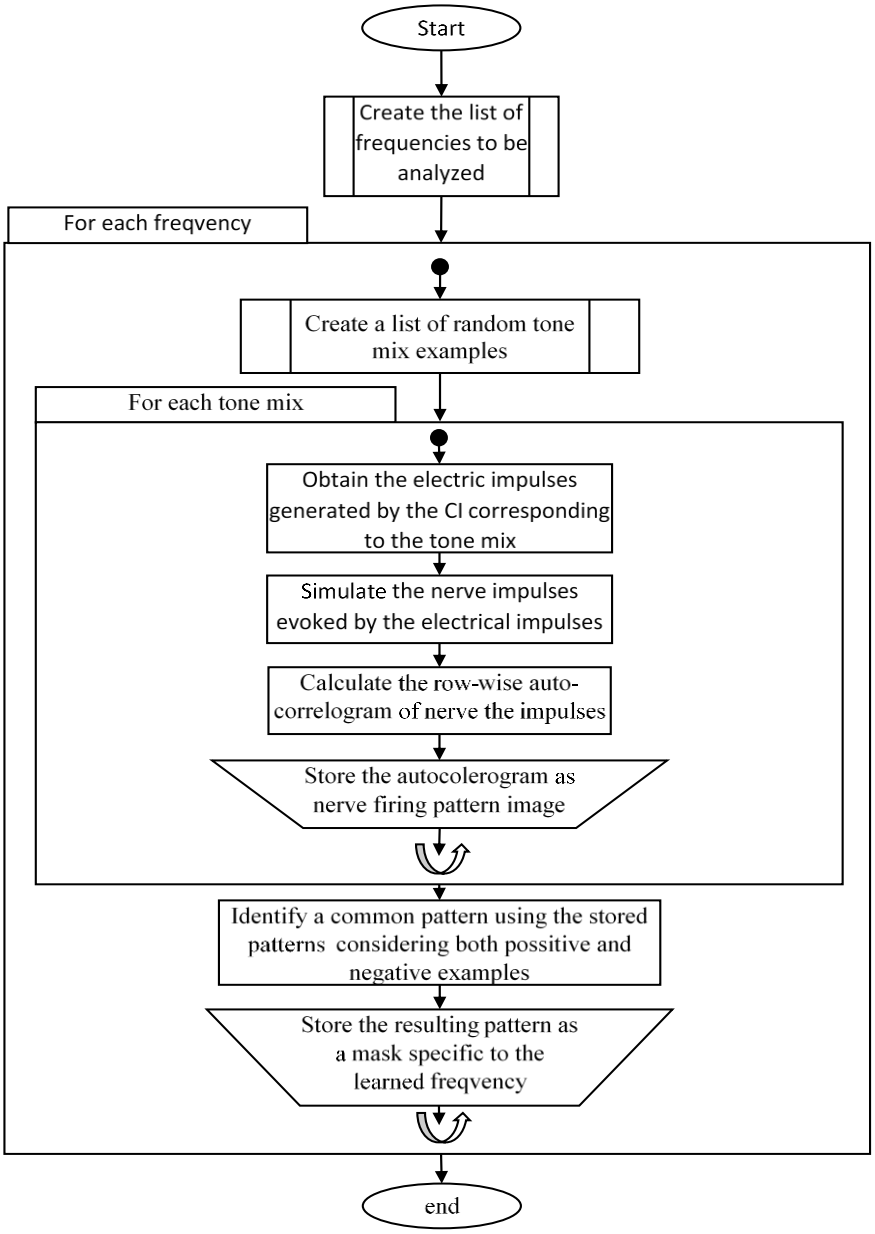
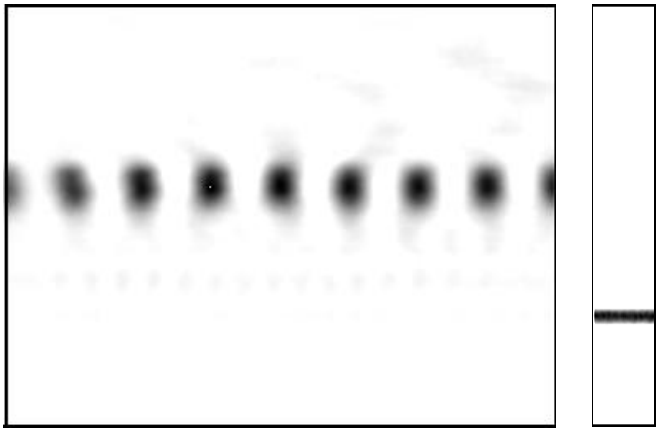


Figure 45: Flowchart - Learning algorithm for frequency specific autocorrelation masks calculated from simulated nerve impulses.



a) 1 kHz *b) 6 kHz*

Figure 46: Frequency specific autocorrelation masks obtained by the learning algorithm: a) Pattern for 1 kHz detection; b) Pattern for 6 kHz detection; see Note2 as in Figure 42

II) **Amplitude cross-correlation coefficients.** Once the frequency specific autocorrelation masks are determined, the intensity of the frequency components perceived for any arbitrary sound sample, according to the cochlear nerve firing pattern approximation model, are estimated by the following algorithm:

1. A1. The *cochlear nerve firing pattern* for the sound sample is estimated by using the analyzed firing pattern approximation model obtained from the simplified auditory model of natural hearing (Figure 36), or from the cochlear implant stimulation strategy.

2. The *row-wise autocorrelation pattern* of the cochlear nerve firing pattern for an unknown sound is computed (8) .

3. For each frequency f_i, the *amplitude cross-correlation coefficient approximating the perceived intensity (amplitude)* is calculated:
   3.1. The f_i frequency specific autocorrelation mask is cross-correlated with the row-wise autocorrelation pattern for the unknown sound sample, resulting in the *amplitude cross-correlation coefficient (9),* proportional to the *perceived amplitude of f_i.*

Cross-correlation formula to calculate the *amplitude coefficient* is:

$$\propto_{f_i}^{to} = \sum_{row_i=1}^{312} \sum_{k=1}^{m} AC_{row_i}^{to}(k) * MASK_{row_i}(k) \qquad (9)$$

where $\propto_{f_i}^{to}$ is the cross-correlation coefficient of frequency f_i at time *to*, $AC_{row_i}^{to}(k)$ is one row of the autocorrelation image of the analyzed signal corresponding to time *to* and

$\mathrm{MASK}_{\mathrm{row}_i}(k)$ is one row of the autocorrelation mask of the frequency f_i., and m is the length of the sampling window.

The determined amplitude cross-correlation coefficient, indicating the perceived intensity, usually is not linearly proportional to the equivalent amplitude of the frequency component in the analyzed sound sample.

To mitigate this problem, the following algorithm was developed to determine the amplitude response characteristics corresponding to each frequency mask:

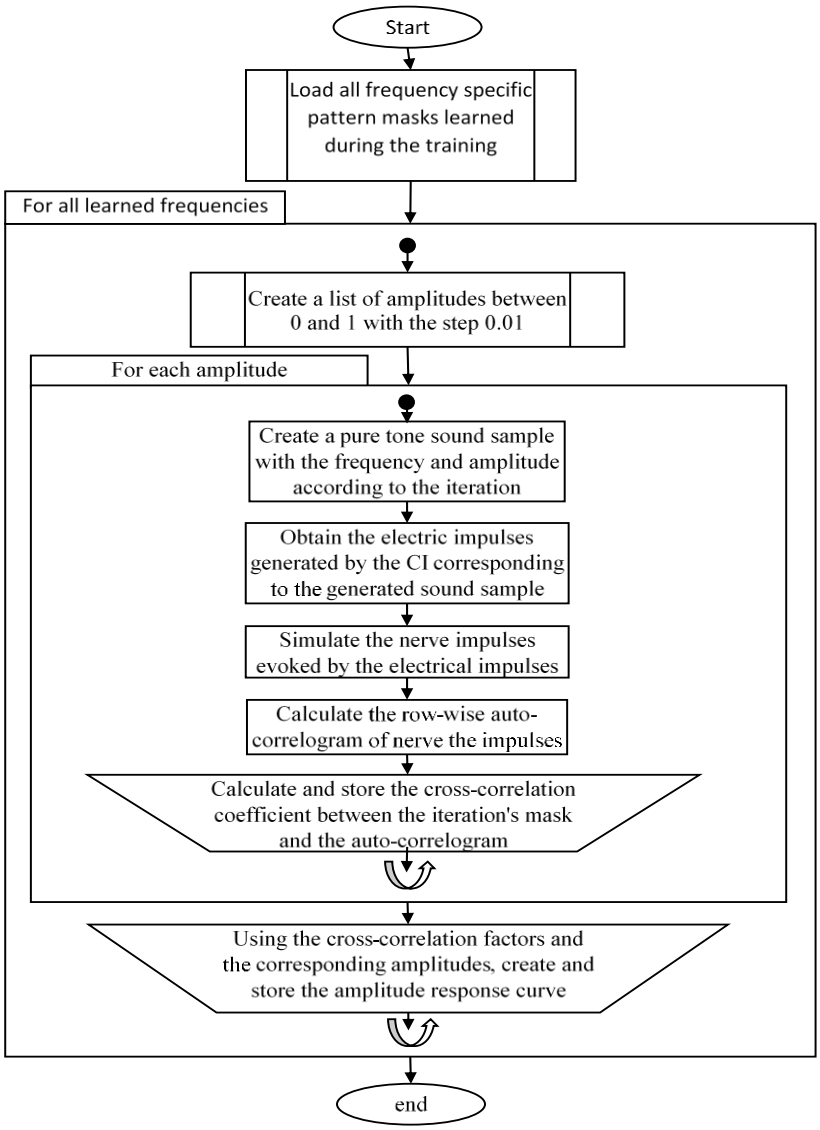


Figure 47: Flowchart – Algorithm to determine the amplitude response characteristics corresponding to each frequency mask

Using the above algorithm, a *mapping curve is determined for each frequency specific autocorrelation mask showing the relation between the amplitude cross-correlation coefficient and the equivalent amplitude of frequency component*.

Using these mapping curves and the set of amplitude cross-correlation coefficients calculated for an arbitrary sound sample, it is possible to compose a curve showing the *spectral components of the perceived sound sample similar to a Fourier Transform.*

4.3.3 Frequency Response of the Frequency Detectors in the Case of Natural Hearing

Frequency detectors based on nerve firing autocorrelation masks are not only sensitive to the specific frequencies but to the frequency vicinity too. How broad this vicinity is, depends on the cochlear nerve firing pattern model, allowing a quantitative estimation of the frequency discrimination capability of the modeled hearing.

An algorithm was developed (Figure 51) to estimate the frequency response of the frequency detectors. Due to the nonlinear nature of amplitude response of the frequency detectors, the frequency response curves are specific to the amplitude of the input sound. Therefore, instead of one frequency response curve per detector, a collection of frequency response curves is calculated, for each detector, utilizing different input sound levels.

By simulating the cochlear nerve firing patterns in the case of natural hearing (Figure 36) for all frequencies in the frequency list *FL*, and calculating the corresponding cross-correlation coefficients, the *frequency response characteristic for each frequency detector* f_i is obtained.

Figure 48 displays the frequency response curve for the 440 Hz detector, using maximum sound level in the case of natural hearing model, with the following characteristics: i) The frequency response is quite narrow at this frequency which is expected in the case of natural hearing. ii) There are smaller peaks on both sides of the response peak. After researching on the meaning of these peaks, it was observed that the *corresponding frequencies of these peaks* are equal to the frequencies of the notes in *the musical harmony* starting or ending with 440 Hz. This finding might show that the human ear is hard wired to easily detect harmonies.

In case of the 6 kHz detector (Figure 49), the frequency response curve has the following characteristics: i) it is less narrow, and ii) does not show artifacts related to the musical theory. This conclusion is also consistent with the literature (Schnupp, et al., 2011) (Meddis & Lopez-Poveda, 2010) as the pitch perception in humans gets worse at high frequencies and also, the fundamentals of most tonal musical instruments are lower than 4 kHz.

Finally, the *frequency response of all detectors* is represented as a chart in one square plot (Figure 50), where the X-axis represents the frequency of the detector and the Y-axis represents the frequency of the tested tone. The intensity of each point in the plot shows the magnitude of the cross-correlation coefficient among the f_i mask and the cochlear nerve firing pattern. In this case*, the values are normalized per column* to equalize the differences between the maximum responses of the detectors.

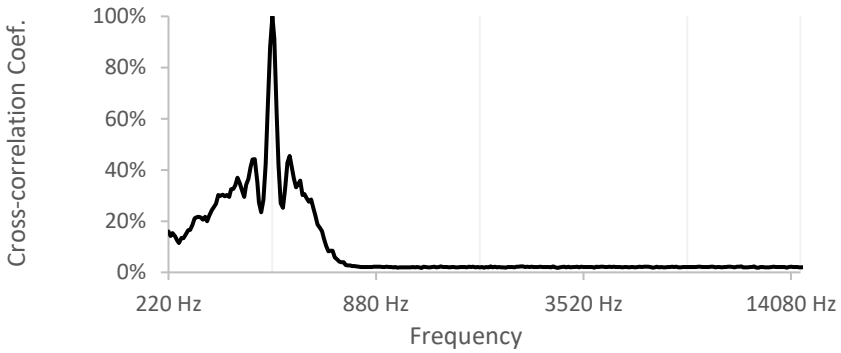


Figure 48: Frequency response of the 440 Hz detector in case of natural hearing.

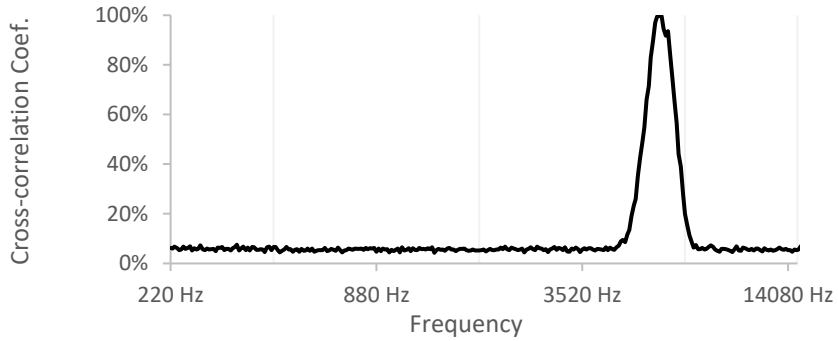


Figure 49: Frequency response of the 6 kHz detector in case of natural hearing.

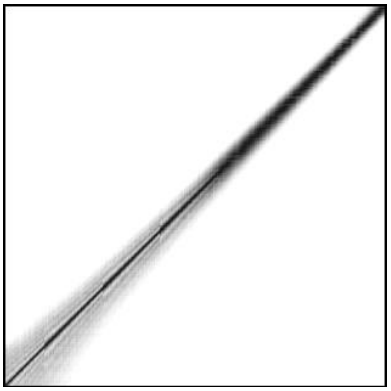


Figure 50: Frequency response chart for all detectors in case of natural hearing. X-axis - detector frequency, Y-axis - all input frequency range, Intensity – The cross-correlation coefficient between the detector mask at frequency X and the nerve impulse pattern generated as response to a sound at frequency Y

The plot in Figure 50: i) clearly predicts the better pitch perception at low frequencies compared to high frequencies, and also ii) predicts some form of channel interactions or resonances on the lower half of frequency spectrum, in the case of natural hearing.

These resulted frequency responses from Figure 48-50, with specific characteristics, validate the proposed frequency detectors in the case of natural hearing.

In *conclusion*, *although the learning algorithm of the frequency specific autocorrelation mask of the cochlear nerve firing pattern is completely agnostic of the spatial and temporal cues*, *it seems to correctly estimate the pitch perception quality of natural hearing.*

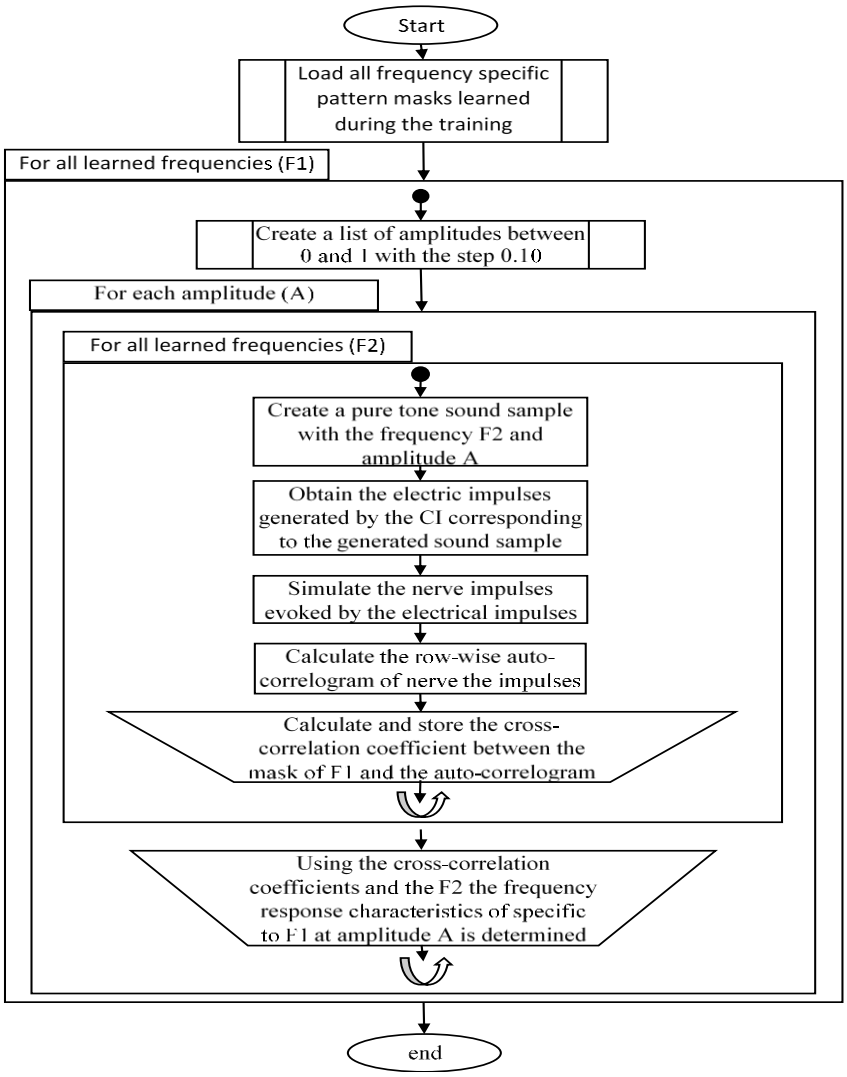


Figure 51: Flowchart –Algorithm to determine the frequency response corresponding to each frequency mask (calculated at different amplitudes)

4.3.4 Frequency Response of the Frequency Detectors in the Case of Cochlear Implants

The main purpose of the proposed auralization method is to provide an *objective method to predict hearing quality* with existing or under development cochlear implant devices.

In this section, simulation results regarding the frequency response of the proposed frequency detectors from cochlear nerve firing patterns are presented by using: i) a simulated cochlear implant device with ACE strategy, and ii) an integrating real MED-EL Opus 2 cochlear implant processor.

For cochlear implant simulation, only frequency detectors above 220 Hz are used, as this frequency is usually used as lower cutoff frequency for cochlear implants.

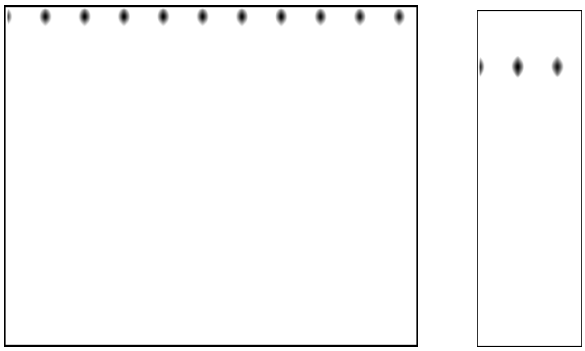
4.3.4.1 Frequency Response of the Frequency Detectors in the Case of ACE Stimulation Strategy

The *Advanced Combination Encoder - ACE* stimulation strategy for CI is based on a so-called N of M stimulation strategy (Harczos, et al., 2013) using 24 electrodes to stimulate the cochlear nerves. The electrodes are activated cyclically with a constant rate, but with modulated intensity. *A bandpass filter and an envelope detector correspond to each electrode* and determine the amount of electrical current to be sent to the nerves. To reduce the channel noise and channel interaction, only the most dominant N electrodes from all M electrodes are used at one moment stimulating with a specific firing rate.

In this setup, the *ACE strategy was implemented in a MATLAB model* (Harczos, et al., 2013), and the *nerve firing intensity* was approximated proportional with the electric impulse intensity with a fixed decay. The *current spread* was approximated with a simple linear model of the cochlea where the nerve endings and the electrodes are placed on two parallel lines. Based on this model, the cochlear nerve firing patterns are obtained.

Figure 52 shows the *autocorrelation mask* obtained for 320 Hz and 1 kHz. In contrast to natural hearing, the firing rate is not influenced by the frequency of the sound (the distances between the spots are equal in both cases), therefore the temporal cue is completely missing.

The frequency responses in the case of CI with ACE strategy at 440 Hz (Figure 53) and at 6 kHz (Figure 54) are significantly worse when comparing with the responses of the natural hearing model. Artifacts and peaks on the curves seem to be related to the neighboring electrodes and not to the musical harmonies.



a) 320 Hz *b) 1 kHz*

Figure 52: Autocorrelation masks obtained by the learning process for ACE strategy: a) Pattern for 320 Hz detection; b) Pattern for 1 kHz detection. Note2: X - time shift of the autocorrelation, Y - frequency related position

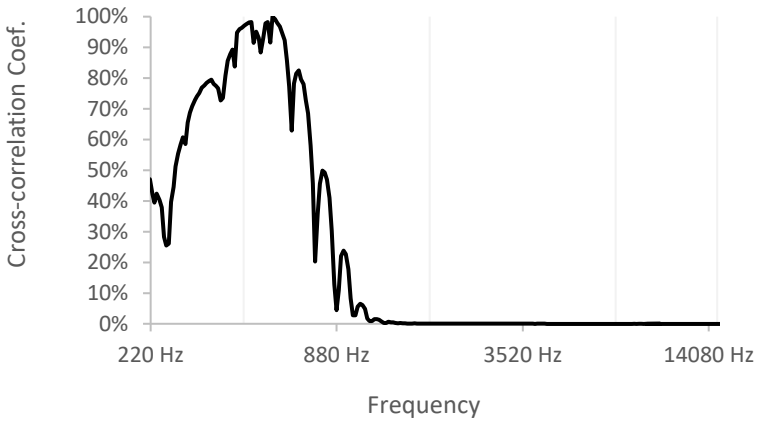


Figure 53: Frequency response of the 440 Hz detector in case of CI with ACE strategy

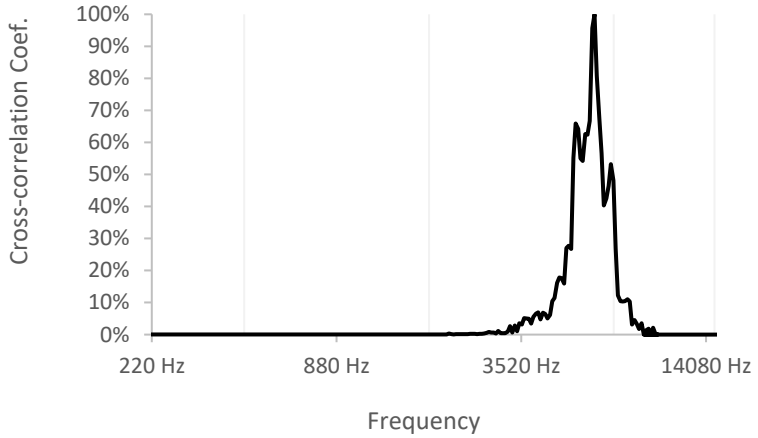


Figure 54: Frequency response of the 6 kHz detector in case of CI with ACE strategy

Figure 55 shows the *overall frequency response plot* in the case of CI with ACE strategy with the following characteristics: i) The pitch *perception gets better with high frequencies*. This is again consistent with previous findings (Hochmair, et al., 2007), which reports that *CI users typically understand female voice easier* as opposed to male voice. ii) The left part of the image is empty, that means that very low frequency spectrum is missing, is not transmitted at all by this coding strategy. iii) The fact that the pitch perception at low frequencies is significantly worse than natural hearing, prevents CI users with ACE or similar strategy to enjoy musical

melodies (Drennan & Rubinstein, 2008), although they are still capable of enjoying rhythm, and have a good speech perception.

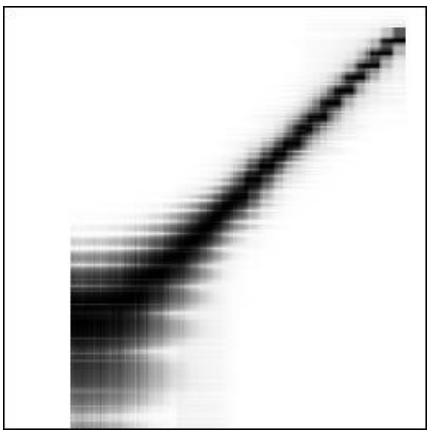


Figure 55: Frequency response chart for all detectors in case of CI with ACE strategy. X-axis - detector frequency, Y-axis - all input frequency range

4.3.4.2 Frequency Response of the Frequency Detectors in the Case of MED-EL Opus 2 CI Processor with FS4P Strategy – Experimental Results

In the second experiment, a real MED-EL Opus 2 CI processor was interfaced with the learning system. The sound signal was sent to the processor using a dedicated audio cable, but the processor microphone was not switched off, thus leading to some degree of noise in the experiment. The generated impulses were collected using the I100+ interfacing box and an own developed 12 channel analog data acquisition box (Kuczapski & Andreescu, 2016) that will be presented in the section 5.5.

For current spread and the nerve excitation, the same model was used as in the case of the ACE stimulation strategy. Based on this chain, the cochlear nerve firing patterns are obtained.

The *FS4P (Fine-Structure 4-Channels Parallel*) stimulation strategy of the Opus 2 CI processor (see section 2.3.3.2) was used during the experiment, which is designed to transmit both spatial and temporal cues to the cochlea, but using only 12 electrodes. An example of cochlear nerve impulses firing pattern after the current spread and decay model are displayed in Figure 56.

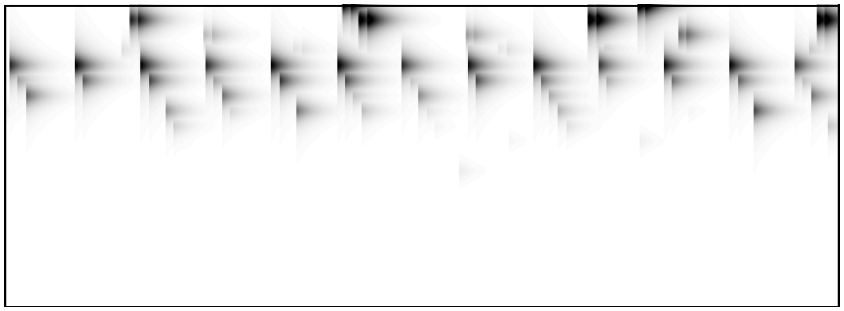
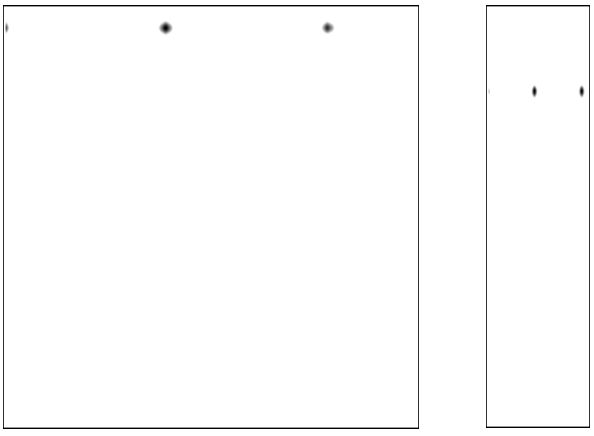


Figure 56: Experimental recorded cochlear nerve impulses firing pattern from an Opus 2 processor and simulated decay for a mixed signal with two tones; Note1: X axis - time, Y axis - stimulation location along the basilar membrane corresponding to a specific frequency (high-frequency to origin)

The presence of the predicted temporal cues is observed in Figure 57, in the autocorrelation masks for these frequencies, as the distance between the spots differs from 320 Hz to 1 kHz. Besides the stimulation rate differences the location of stimulation is frequency dependent, similarly to the ACE strategy.

The pitch perception on low frequencies is predicted to be closer to the natural hearing as the Figure 58 shows. However, the exact shape of the frequency response is hidden by the noise caused by the processor microphone. At high frequencies (Figure 59), the frequency response is significantly worse compared to both natural hearing and CI with ACE strategy.



a) 320 Hz *b) 1 kHz*

Figure 57: Autocorrelation mask obtained by the learning process for FS4P strategy: a) Pattern for 320 Hz detection; b) Pattern for 1 kHz detection. Note2: X - time shift of the autocorrelation, Y - frequency related position

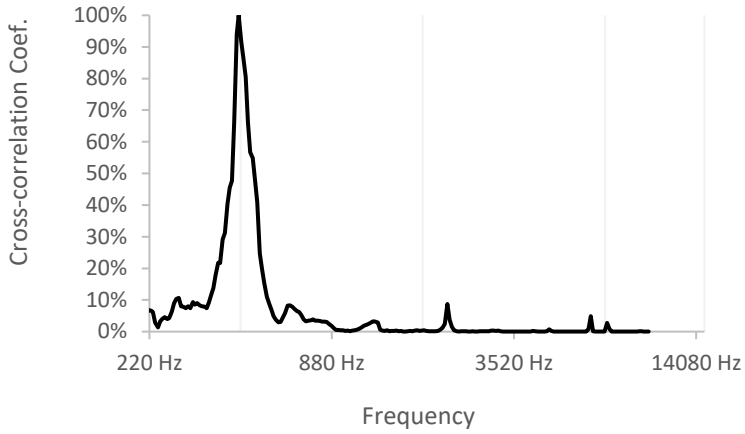


Figure 58: Frequency response of the 440 Hz detector in case of MED-EL Opus2 CI with FS4P.

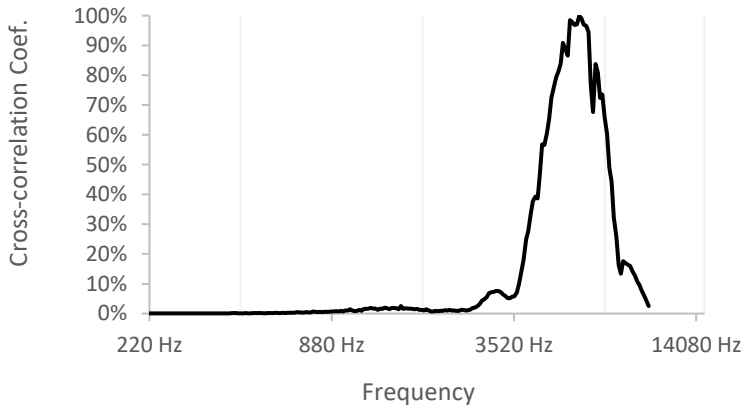


Figure 59: Frequency response of the 6 kHz detector in case of MED-EL Opus 2 CI with FS4P strategy.

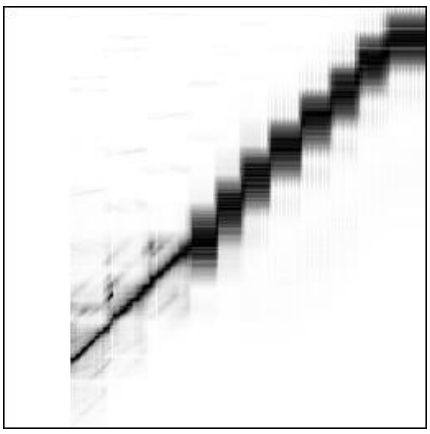


Figure 60: Frequency response chart for all detectors in case of MED-EL Opus 2 CI with FS4P strategy. X - detector frequency, Y - all input frequency range.

Figure 60 displays the *overall frequency response plot* of the Opus 2 CI processor with FS4P strategy. It predicts better pitch perception at low frequencies than at high frequencies, which is exactly the opposite of the ACE strategy characteristics.

4.3.5 Auralization Vocoder

Using the previously presented setup of frequency detectors, the auditory nerve stimulation is translated into 312 continuous signals representing together the spectral image of the perceived sound. Even though this number of channels provides a very high spectral resolution, the effective resolution is highly affected by the frequency response of the individual detectors. The broader the frequency bands indicated by the frequency response of the detectors, the more faded is the spectral image extracted.

In order to synthesize the perceived sound (auralization), the effective resolution of the spectral image must be taken into consideration. In other words, the synthesized sound must only contain the minimum number of spectral components which could potentially generate the same spectral image.

To resolve the above problem, a dedicated channel vocoder algorithm was developed. During the sound synthesis, for each sample, the *algorithm identifies the list of those dominant frequencies* of which corresponding detector's frequency response summed up approximates the curve of the momentary spectral image. The outputs of the selected frequency detectors and their frequencies are used then to modulate a collection of tone generators.

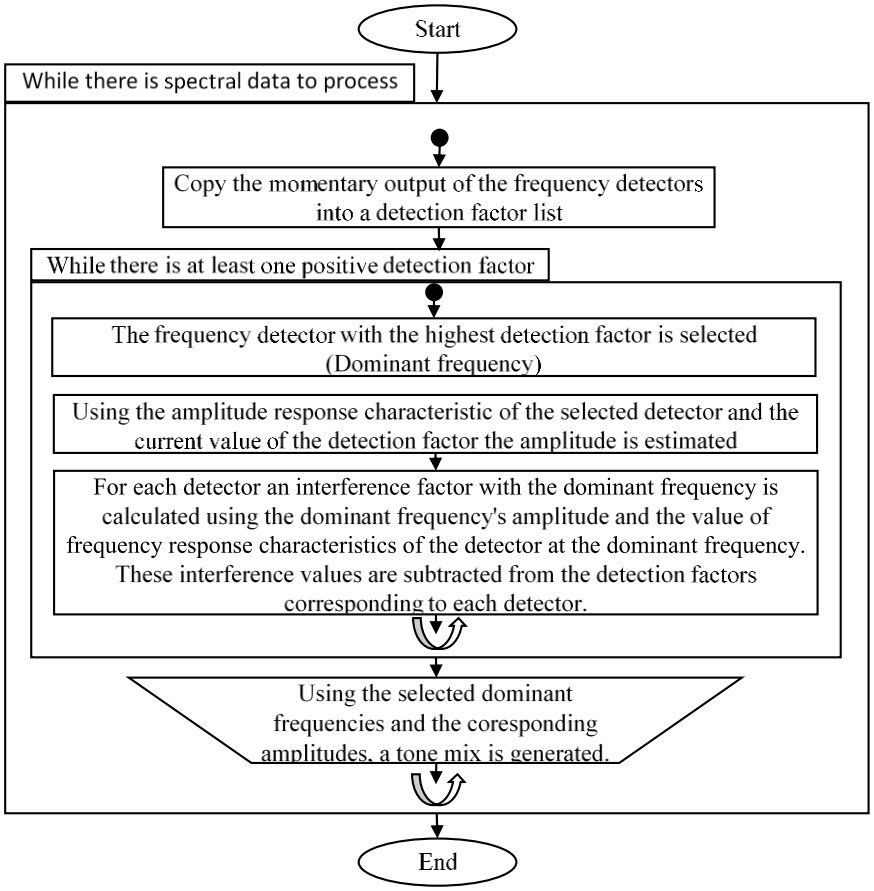


Figure 61: Flowchart - Dominant frequency selection and sound synthesis

The illustrations in Figure 62 and Figure 63 demonstrate how the data is processed through the auralization algorithm.

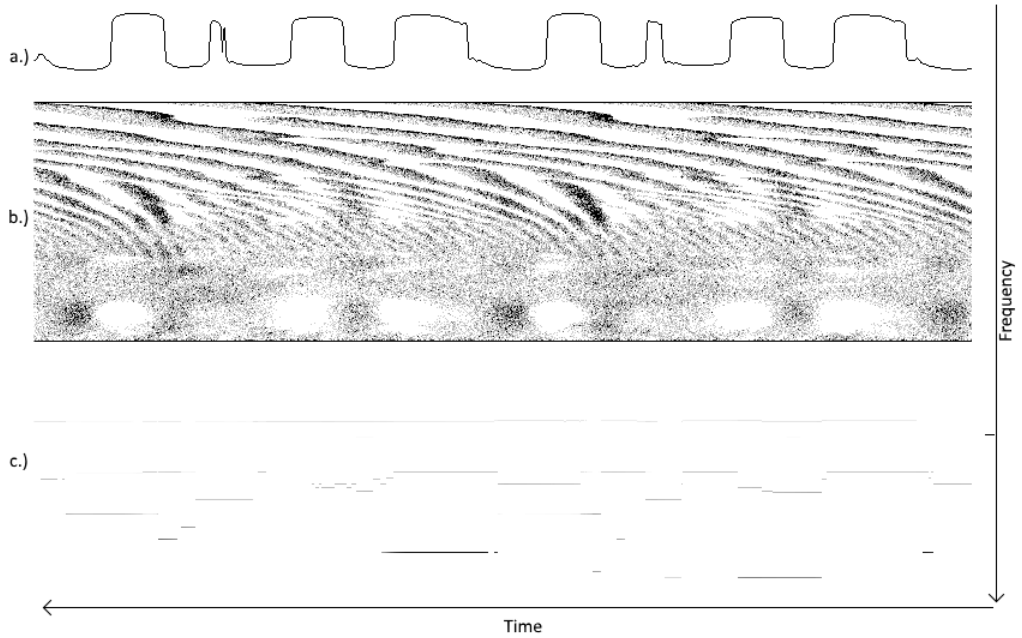


Figure 62: a.) Input sound signal using logarithmic amplitude scale; b.) Nerve firing pattern; c.) output signals of the 312 frequency detectors.

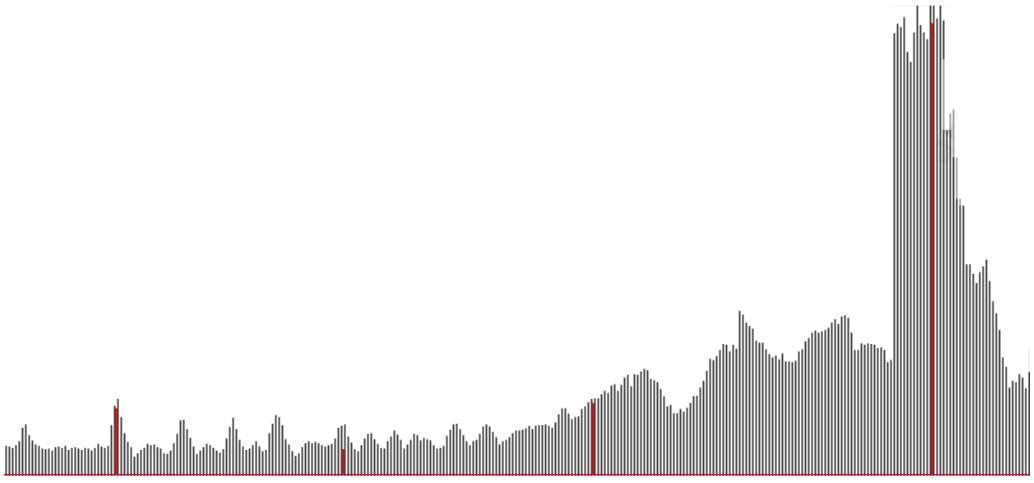


Figure 63: Momentary view of the detected spectral components overlapped with the selected dominant frequencies (Vertical cross section of the previous figure). X – detector frequencies; Y – detected amplitudes.

4.3.6 Implementation of the New Autocorrelation based Auralization Method

All simulation models, filters, algorithms and modules were developed and executed using Java programming language and Eclipse IDE development framework. In order to allow a continuous development implementation of new features and various experimental models, a signal processing library was created as a foundation. This own developed library consists of the following main packages:

- **edu.kuczapski.math** – Implementation or adaptation of generic mathematical concepts, data models and functions:
  - Fourier transformation
  - Matrix Operations
  - Windowing functions and resampling
  - Cyclic memory buffers
  - Autocorrelation

- **edu.kuczapski.signal.discrete** – Definitions of interfaces and generic model classes used for discrete signal processing:
  - ISignalSink, ISignalSource and IFilter
  - DiscreteCircuitElement, ComposedDiscreteCircuitElement and DiscreteSignalModel
  - AbstractFilterBank

- **edu.kuczapski.signal.discrete.fir** - Filter implementations using finite inpulse response models
  - Custom FIR Filter
  - Gammatone Filter

- **edu.kuczapski.signal.discrete.iir** – Filter implementations using infinite response models
  - 1st order lowpass filter
  - 2nd order bandpass filter
  - 4th order Gammatone filter

- **edu.kuczapski.signal.discrete.ops** – Implementation of simple signal operation elements
  - Add or multiply signals
  - Signal Integrator
  - Zero cross detection element
  - Signal compression and decompression module

- **edu.kuczapski.signal.discrete.signals** – Implementation of various signal sources:
  - Impulse and step functions
  - Pure tone signal
  - White noise signal

- **edu.kuczapski.utils** – Utility function used for audio recording, signal plotting and visualization

- **edu.kuczapski.cochlear** – Filters and signal processing models used to simulate the inner ear:
    - Basilar membrane filter bank
    - Inner hair cell model and filter bank (Adapted from Zang et al.) (Zhang, et al., 2001)
    - Synapse model and filter bank (Adapted from Zang et al.) (Zhang, et al., 2001)

The proposed auralization system, patented by the thesis author (Kuczapski, 2015), relies on the above signal processing framework, and it is designed in a highly modular way allowing different experimental setups. Figure 64 depicts the module structure of the developed auralization system. M1, M2, M3, M4 and M5 are modules responsible to implement the signal processing model of various stages of the auralization or the training phase, G1 is used as audio input during the training, I1 is used as audio input for auralization. R1 is the result of the auralization, while R2 are numerical reports resulting from the training process.

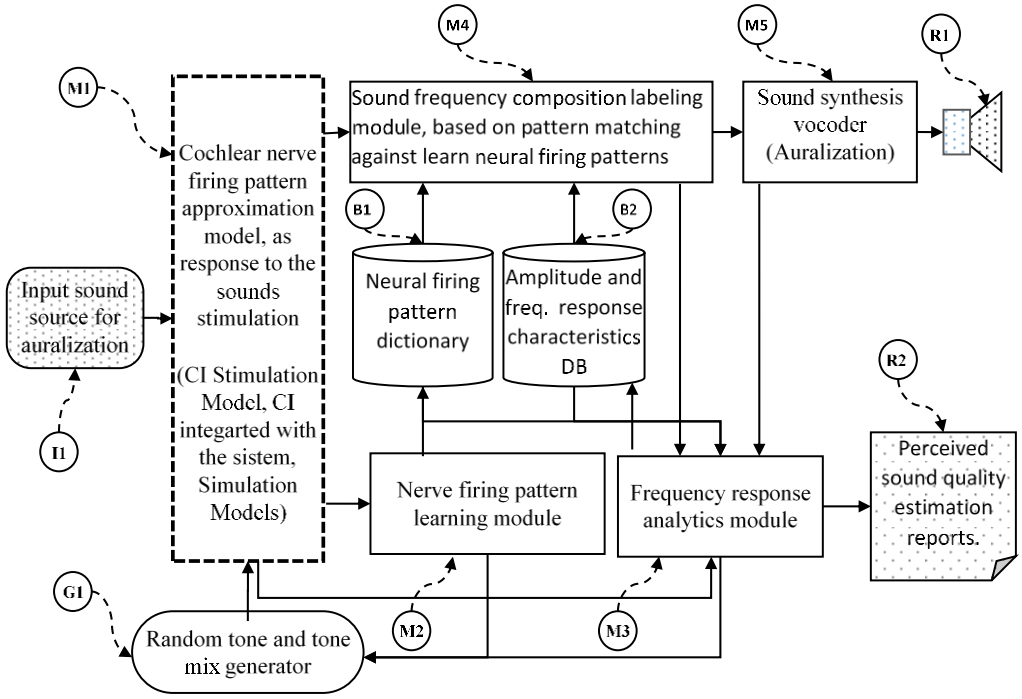


Figure 64: Block diagram - proposed complete auralization system (Kuczapski, 2015)

While the modules from M2 to M5 are the implementation of the new auralization method and therefore they are fixed, the implementation of the module M1 is highly dependent on the experiment setup and it is easily exchangeable. The following implementations were developed for M1:

- i) Cochlear nerve firing pattern approximation using natural auditory model (Figure 65)
- ii) Cochlear nerve firing pattern approximation using emulated cochlear implant (Figure 67)
- iii) Cochlear nerve firing pattern approximation using a real cochlear implant Processor (Figure 68)

i) The 1st implementation of M1 serves as a reference implementation used to develop and fine-tune the auralization method. A simplified natural hearing auditory model (S3.1), implemented based on the model introduced in chapter 4.1, supposedly preserves all information carried by the audio signal, therefore the auralization method should perform best with this model.

Furthermore, the development of the sound synthesis module (M5) was greatly helped by the experiments performed with the natural hearing auditory model, resulting in the method described above, and implemented in the modules S4.3 and S4.4 depicted in Figure 66 - Frequency detections and sound synthesis blocks.

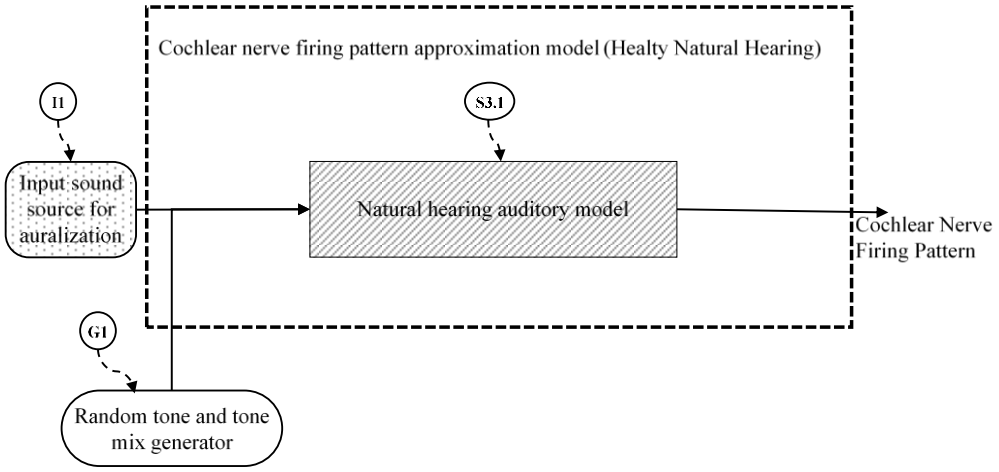


Figure 65: Block diagram – Cochlear nerve firing pattern approximation using the simplified natural auditory model presented in Figure 36 (Kuczapski, 2015).

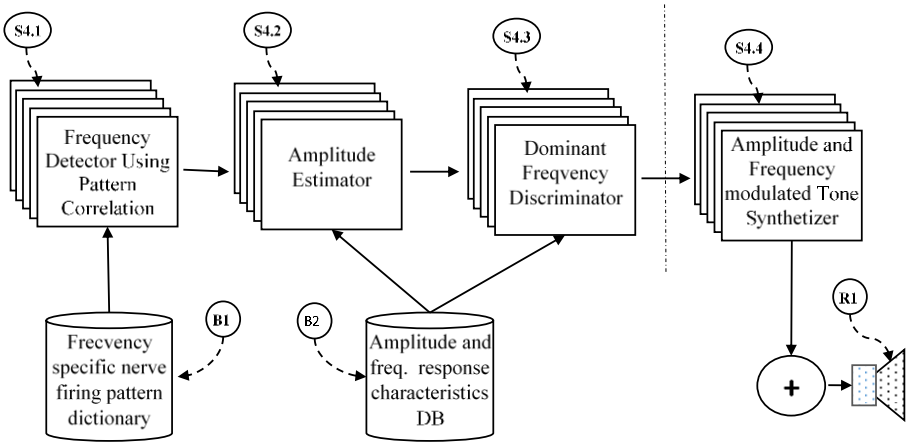


Figure 66: Block diagram – Frequency detections and sound synthesis (Kuczapski, 2015)

ii) Figure 67 depicts the 2nd implementation of the module M1. This module was developed to fulfill the main scope of the work, namely, to be used to create auralizations of the sounds perceived by cochlear implant users. This approach uses a *software CI simulation model of the stimulation strategy* used by the targeted cochlear implant type (S1.1). The output of this simulation model, consisting of discrete signals representing the electric signals generated by the cochlear implant to electrodes, is fed to an *intra-cochlear current spread model* (S1.2), then the nerve impulses are estimated by the *electric nerve stimulation and firing model* (S1.3). Given the limited access to the algorithms used by the cochlear implant manufacturers, this approach was possible only using the ACE stimulation strategy.

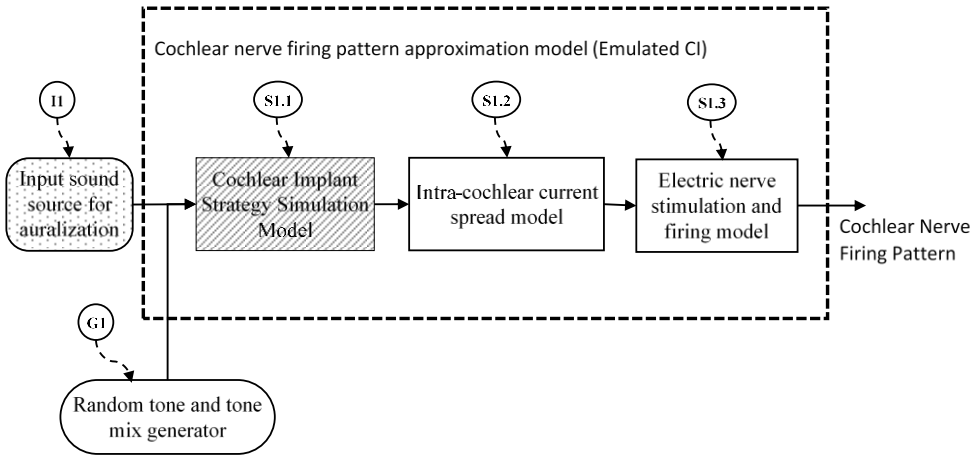


Figure 67: Block diagram – Cochlear nerve firing pattern approximation using emulated Cochlear Implant (Kuczapski, 2015)

iii) In order to circumvent the lack of cochlear implant stimulation strategy models, a 3rd implementation of M1 was developed, allowing direct interfacing with a real cochlear implant processor (Figure 68). In this case, both the input audio signal for auralization (I1) and the training signals (G1) are sent to a physical cochlear implant processor through a suitable audio interface (H2.1). The audio interface is realized using the computer audio output connected directly to the CI Processor (H2.2), if possible, or connected to a loudspeaker positioned in a soundproof chamber near to the cochlear implant. The stimulation signals generated by the CI Processor (H2.2) are captured by a specific hardware interface (H2.3) and the recorded signals are sent to the interfacing module (S2.1) within this implementation of M1. Further on, the nerve impulses are estimated the same way as in the 2nd implementation of M1.

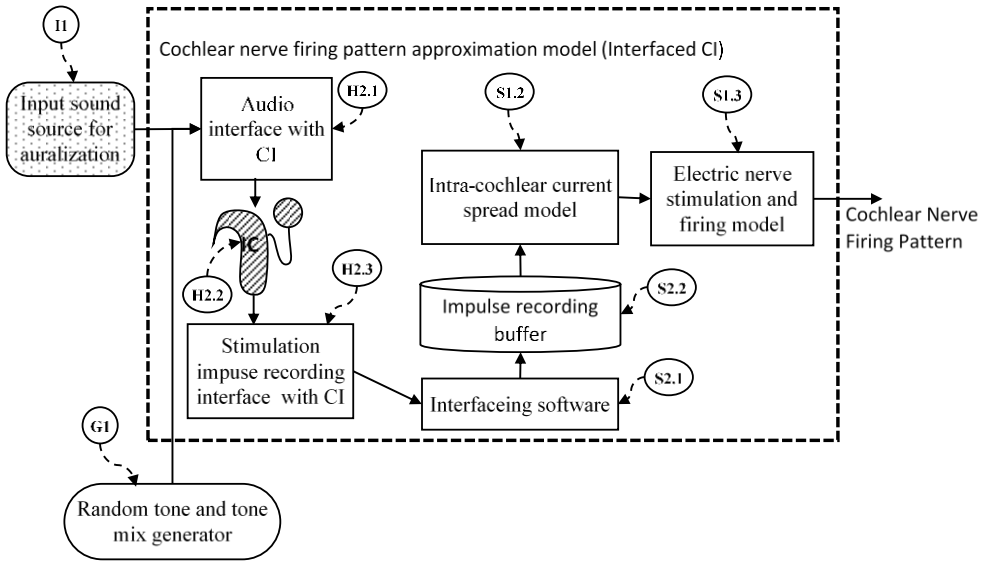


Figure 68: Block diagram – Cochlear nerve firing pattern approximation using Real Cochlear Implant Processor (Kuczapski, 2015)

Figure 69 depicts the physical setup needed to implement the 3rd variant of the M1 module. In this setup, H3.1 is a laptop used to run the auralization software including all software modules described above. A loudspeaker (H2.1) is used to convey the audio signal from the computer to the CI Processor (H2.2), and the CI Processor Hardware interface (H2.3) is used to retrieve the stimulation signals corresponding to the conveyed audio signal. In order to realize this experimental setup, such a CI Processor interfacing unit (H2.3) was developed and used (Kuczapski & Andreescu, 2016).

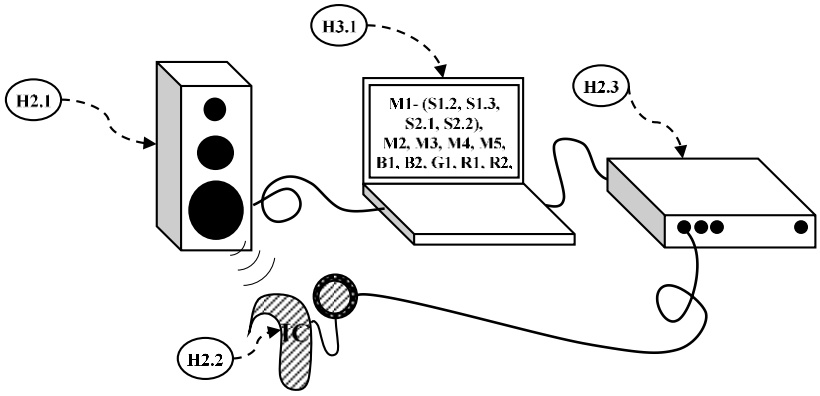


Figure 69: Physical setup of the system using real cochlear implant with an interface box. (Kuczapski, 2015)

4.4 Experimental Results to Validate the New Auralization Method - Cochlear Implant User Feedbacks

The main purpose of the methods developed in this thesis is to provide a *reliable tool to estimate and replay the sounds and their nature perceived by cochlear implant users*. Because hearing, like all sensory experiences, is very intimate for individuals, and no one can be sure that a certain stimulus is perceived in the same way by different persons, there are no ways to verify what an actual person is really hearing. Commonly, when sensations are verbally communicated between individuals, it is always done by comparisons to some commonly known phenomena or sensations commonly experienced by both interlocutors. For example, if someone tries to describe a perceived sound, they will use the expression "It sounds like …".

Unfortunately, many cochlear implant users have never experienced sounds without using the cochlear implant, therefore they are not sharing any common hearing experiences with healthy hearing people. Even more, it seems that the persons losing their hearing in a later stage of life, will not adapt completely to the cochlear implant, and they will always refer to the perceived sounds as artificial.

The *developed auralization algorithm* builds a bridge between the two otherwise completely isolated set of hearing experiences, by synthetizing sounds, which are perceived in the same way as the original sounds by the cochlear implant users, but carries a maximum amount of possible deviations/ distortions perceivable only by normal hearing persons.

In summary, the following statements are describing the *principles behind of the auralization algorithm*:

- The information contained in the sound generated by the auralization cannot be more than what was transferred through the electrodes.
- The cochlear implant is a consistent system that always deletes the same information from the audio signal.
- *If the audio signal does not contain information which would be deleted by the cochlear implant*, then the patient has chances *to perceive the complete information quantity carried by the audio signal*.
- If there is a method that deletes the same information from the audio signal as the cochlear implant would delete, and then the *original signal and the signal with deletion is replayed to a cochlear implant user*, *the user should not perceive the difference*.

Validation criteria of a successful auralization, based on these principles, are the following:

An auralization can be considered as an accurate representation of the sounds perceived through the modelled cochlear implant, if and only if:

1. **The generated sound was created only using the electric impulses created by the cochlear implant.**
2. **A cochlear implant user, using the same type of cochlear implant as the one modelled in the auralization process, cannot hear the difference between the original and the synthesized sound.**

An important prerequisite in providing an accurate auralization of the perceived sounds is to *capture or replicate the stimulation impulses* in a controlled way and *with high accuracy*. The best way to assure this fact is to directly integrate with the software algorithm used by the cochlear implant processor. During the development of the auralization algorithm, the ACE stimulation strategy was the only one available for us as software module and ready to be integrated with the auralization algorithm. Therefore, all experiments involving patients were executed using an auralization model trained with the ACE strategy, regardless of the patient's cochlear implant model.

Two sessions of validation experiments were done during the development of the auralization method:

I) In the first session, five patients were identified with MED-EL Tempo+ Cochlear Implants using coding strategies similar to the ACE but with only 12 electrodes. In these experiments, during an informal interview, a sample sounds were replayed both *in the original and synthesized form and the participants were asked to identify which version is more pleasant to hear*. In most cases, the synthesized sound was perceived by CI users as good or better compared to the original sound.

Based on these experiences a more elaborated method was developed to conduct a second session of experiments.

II) In the second session of experiments, we had 15 patients participating, with age varying between 9 and 18, and with various types of cochlear implants and settings:

#	Age	No. Implants	Brand	Implant Type	Processor Type	Uses FS4p
1	9	2	Cochlear	Freedom / Nucleus 6	Nucleus 5 / Nucleus 6	n/a
2	10	2	MED-EL	Pulsar / Pulsar	Opus 2 / Opus 2	N
3	10	1	MED-EL	Pulsar	Opus 2	Y
4	13	1	MED-EL	Sonata	Opus 2	Y
5	10	2	MED-EL	Sonata / Sonata	Sonnet / Sonnet	Y
6	13	1	MED-EL	Sonata	Opus 2	Y
7	9	1	MED-EL	Sonata	Sonnet	Y
8	18	1	MED-EL	COMBI 40+	Opus 2	N
9	9	2	MED-EL	Pulsar / Pulsar	Opus 2 / Sonnet	Y
10	9	2	MED-EL	Pulsar / Synchrony	Opus 2 / Sonnet	Y
11	13	1	MED-EL	Sonata	Sonnet	Y
12	9	2	MED-EL	Concerto / Synchrony	Opus 2 / Sonnet	Y
13	9	2	MED-EL	Sonata / Synchrony	Opus 2 / Sonnet	Y
14	9	2	MED-EL	Sonata / Synchrony	Opus 2 / Sonnet	Y
15	18	2	MED-EL	COMBI 40+ / Synchrony	Opus 2 / Sonnet	Y

Figure 70: Details of the patients participating in the 2nd experiment.

The 2nd experiment procedure: We have used a set of audio material registered in a studio consisting of *20 individual sentences* recited by both a *male and a female actor* resulting in 40 reference audio samples (Stanciu, et al., 2008). Using these samples, two *additional samples* sets were generated by *mixing them with white noise*, in the first case using a 15db signal to noise ratio and in the second case a 6db signal to noise ratio. All 120 natural samples were then processed by the auralization method trained for the ACE stimulation strategy, resulting a total of 240 audio samples.

Individual experiment sessions were conducted with each participant. During one session, between *15 and 25 randomly selected audio samples were replayed* and for each sample the following form was completed:

1. The ID of the audio sample – sentence ID, speaker gender and noise level
2. *Sentence understood* - by the participant
3. *Number of retries* – asked by the participant in case that it wasn't sure
4. Final *number of misunderstand words* – filled by the tester
5. A quality score (1-10) – *how natural is the sound*.

In addition, for each session the following participant data was recorded;

1. Age
2. Number of implants

3. Implant(s) type
4. Process(s) type

In total, considering all participants, 296 samples were replayed with the following distributions:

Sample Type	Count
Original	186
Auralization (ACE)	110

Speaker Gen.	Count
Male	150
Female	146

No. Words	Count
3	27
4	53
5	112
6	65
7	28
8	11

Signal/Noise	Count
Clean	120
15 dB	115
6 dB	61

Figure 71: Characteristics of the audio samples used in the 2nd experiment.

It is important to notice that the *ACE stimulation* strategy used in the auralization *does not match the cochlear implant types used by the participants*. All participants, excepting two of them, are recipients of cochlear implants with significantly better coding strategies. Never the less, the results of the experiments are still valuable and can validate the utilization of the proposed auralization method for better representation of the perceived sound.

First, in order to validate the plausibility of the test results, *overall speech understanding scores* are calculated and evaluated for plausibility. We have looked for two well-known factors in speech understanding: Signal to noise ratio, and speaker gender. *The results* are depicted in Figure 72 and Figure 73 and they *are in accordance with the similar statistics in the literature* (Nascimento & Bevilacqua, 2005), therefore *validating the setup of the experiments*.

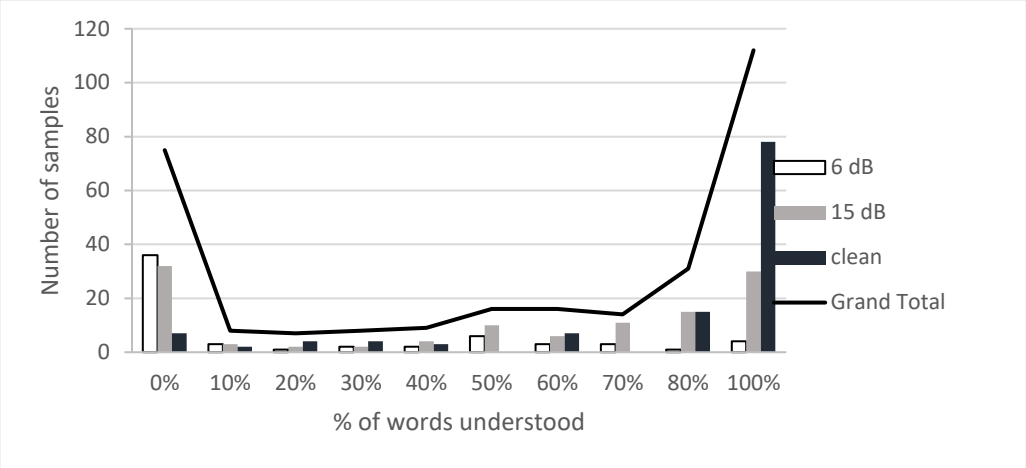


Figure 72: Speech understanding at different noise level: Clean; speech level with 15 dB above noise level; speech level with 6 dB above noise

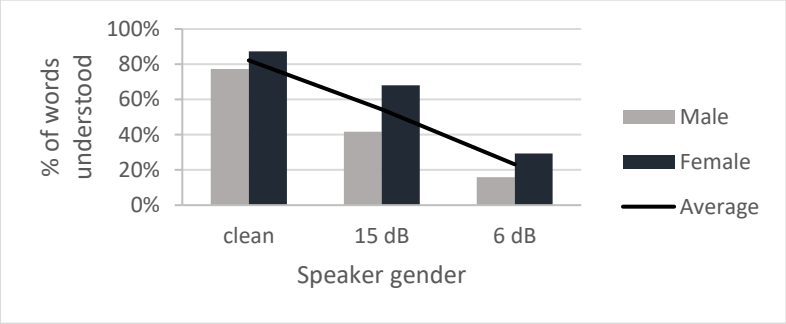


Figure 73: Influence of speaker gender on speech understanding

Figure 74 shows the average speech understanding scores per participant per sample type (auralization vs original audio sample). The results are highly variable from one subject to the other, however, it is completely understandable as the participants are using a wide variety of cochlear implant types and setups. Nevertheless, according to this chart, 6 out of 15 participants have performed the same or better when listening to the auralization. The same statistics were calculated considering only the audio samples without noise (Figure 75). In this case, the number of subjects scoring the same in both cases went down to 4.

Finally, the statistics of the subjective quality score is shown in Figure 76. In this case 4 participant were not able to give quality score (Subject 1, 10, 12 and 15). From the remaining 11 participants, 4 have rated the synthesized audio samples with and a higher average quality score.

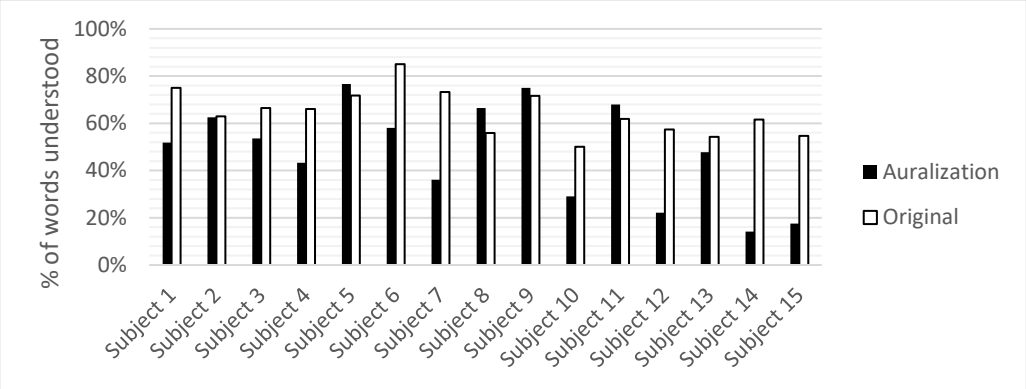


Figure 74: Average speech understanding per participant

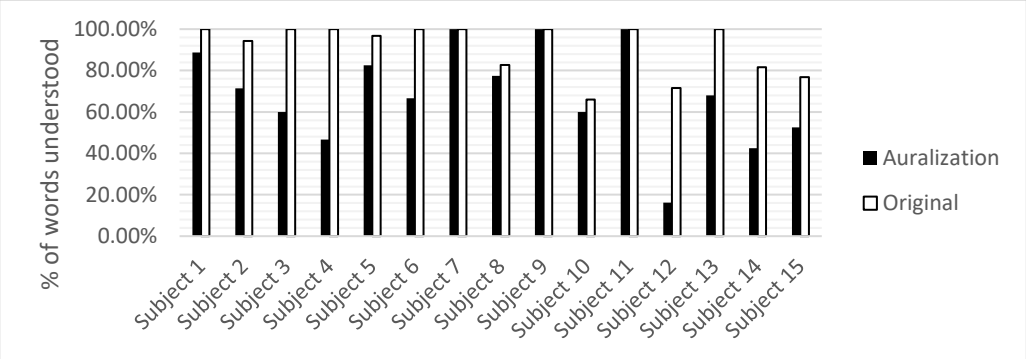


Figure 75: Average speech understanding per participant excluding samples with noise

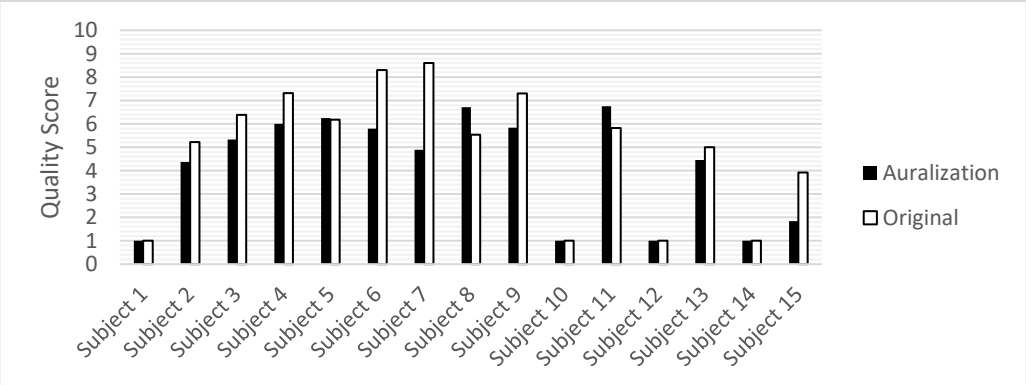


Figure 76: Average audio quality score per participant

Participant	% words	% words (clean)	Quality Score
Subject 1	--	-	?
Subject 2	=	--	-
Subject 3	-	---	-
Subject 4	--	---	-
Subject 5	+	-	=
Subject 6	--	--	--
Subject 7	---	=	---
Subject 8	++	-/=	+
Subject 9	+	=	-
Subject 10	--	-/=	?
Subject 11	+	=	+
Subject 12	---	---	?
Subject 13	-/=	--	-/=
Subject 14	---	---	?
Subject 15	---	--	--

Figure 77: Subject preference map, comparing auralization sound to original sound.

The table shown in Figure 77 displays an overall view on the results of the experiments. The following signs are used to mark the preferences between the synthesized and original audio samples:

"---", "--": Auralization scores *significantly poorer* compared to the original sound
"-": Auralization scores *poorer* compared to the original sound
"-/=", "=": Auralization scores *similarly or only slightly poorer* compared to the original sound
"+","++": Auralization scores *better or significantly better* compared to the original sound

Based on these experimental results, it can be concluded that Subject 8, Subject 9 and Subject 11 have consistently performed the same or better listening to the audio samples resulting from the auralization. Regarding the subjective appreciation of the sound quality, Subject 8 and Subject 11 have given in average higher scores to the synthesized sounds while Subject 9 had a slight preference toward the original audio samples.

In conclusion, the above experiments show the *capability of the auralization method to express the sounds perceived by cochlear implant users*, however it also highlights the *importance of a good calibration and accurate modeling of the targeted cochlear implant model* and configuration.

4.5 SUMMARY

In the proposed auralization method, the *row-wise image autocorrelation and cross-correlation* are used to create a *new self-learning method to detect perceived sound spectral components from cochlear nerve firing patterns*. The method ability to adapt to various types of CI nerve stimulations is demonstrated.

The proposed approach provides: i) an *objective method to predict hearing quality of CI users* with existing or under-development CI systems; ii) *foundation to develop novel auralization methods* to accurately replay sounds perceived through CIs.

CI patient feedbacks and experimental results indicate the utility and validity of the proposed method in *developing improved CI strategies by allowing rapid experimental testing*.

The *main original contributions* of the proposed method are as following:

- A new method is proposed for detecting spectral components of perceived sound from the cochlear nerve firing patterns, i.e., the *frequency detectors*, that mainly contains:
  a) a self-learning algorithm to determine the *frequency specific autocorrelation masks* for each desired spectral frequency component, and
  b) an *amplitude indicator* of each spectral component (amplitude frequency response characteristic) based on the *cross-correlation coefficient* between of the cochlear nerve firing autocorrelation pattern of an unknown sound, and the frequency specific autocorrelation masks.

- For the natural hearing, the frequency response of the frequency detectors at low and high audio frequencies are in accordance with the real hearing pitch perception that validates the proposed method.

- The proposed method is implemented using cochlear nerve firing patterns from:
  a) a cochlear implant with *simulated ACE strategy*, and
  b) an *experimental MED-EL Opus 2 CI processor using FS4P strategy*,
  with comparative results regarding frequency response of the frequency detectors versus the natural hearing.

- The proposed method is successfully used inside a new *auralization technique for sound synthetization* based on the cochlear nerve firing patterns. To validate this new auralization method, *both original and synthesized sounds were experimentally replayed to CI users*, and in most cases, the synthesized sound was perceived as good or better compared to the original sound.

5 APPLICATIONS (CONTRIBUTIONS) FOR MED-EL COCHLEAR IMPLANT USERS

5.1 THE MAESTRO COCHLEAR IMPLANT SYSTEM

During the experiments and research activities, we have mostly interacted with Med-El's cochlear implant system. Therefore, the peculiarities and fitting methods of this system have been extensively studied and the results have been published (Kuczapski & Stanciu, 2015) (Kuczapski & Andreescu, 2016).

The Maestro Cochlear Implant System is a family of cochlear implant solutions developed and manufactured by the MED-EL company. It is based on the *I100 electronic platform* that provides an implantable stimulator with 12 independent stimulation channels and delivers stimulation through 12 pairs of electrodes, where the electrodes are doubled in order to provide redundancy in case of mechanical failure of electrodes. The I100 platform is delivered in multiple physical packages, from ceramic to titanium casing, but all deliver exactly the same stimulation capabilities. The OPUS 1, OPUS 2 and SONNET external processors are part of the Maestro Cochlear Implant Systems, the OPUS 1 being released in 2006, the OPUS 2 in 2010 and the SONNET in 2014.

During the fitting procedures, the fitting specialist mostly interacts with the following parameters of the implant system:

MCL of each channel — The Maximum Comfortable Level (MCL) parameter specifies for each channel, which is the highest stimulation level that is still comfortable to the patient. The highest audio levels are mapped to this level, and the stimulator will never send an impulse with an electric charge higher than MCL;

THR of each channel — The Threshold (THR) parameter defines the lower bound of the stimulation impulse charges. When there is no sound registered as input, the implanted stimulator will still send stimulation impulses at the THR level. By default, the fitting software sets the THR to 10% of the MCL level;

Maplaw Compression factor — Defines a compression curve (usually logarithmic) that is used to map sound intensities to stimulation intensities. Similar to the natural function of the cochlea, intensity changes of soft sounds result in much higher changes in stimulation intensity than intensity changes of loud sounds.

AGC Compression Ratio and sensitivity — Automatic Gain Control (AGC) imitates the natural behavior of the ear, which, by the means of the stapedius reflex, reduces ear sensitivity when it is exposed to loud sounds. The compression rate and the sensitivity of the AGC can be set during the fitting procedure.

There are also other parameters that can be changed using the MAESTRO fitting software, but they are used only in exceptional cases and by highly qualified personnel.

The main target of a fitting specialist is to set the MCL, THR, maplaw and AGC parameters in such a way that they provide the best level of speech understanding in any ambient condition, and also provide comfortable hearing levels. MCL and THR levels are adjusted to obtain the best possible stimulation levels, whereas maplaw and AGC levels are adjusted to improve sound dynamics and clarity (Moctezuma & Tu, 2011) (Stöbich, et al., 1999).

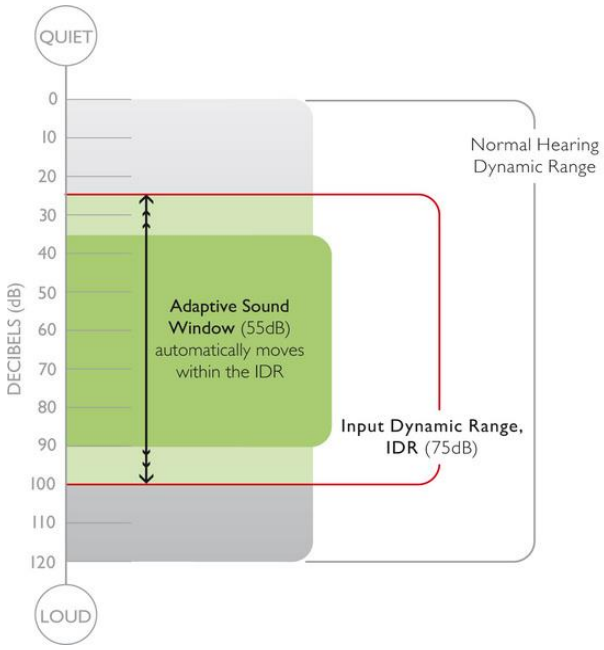


Figure 78: 55dB adaptive sound window of Automated Gain Control (MED-EL(2), n.d.)

According to MED-EL's two staged compression design, in 1st stage, the audio signal coming from the microphone is processed by the dual-loop AGC, which retains only the upper 75 dB dynamic range of the input signal, cutting off low amplitude components (Figure 78). In the 2nd stage, the middle 55 dB of the 75 dB dynamic range is mapped according to the maplaw to the THR–MCL stimulation range. Any signal under the 55 dB sound window will result in stimulation impulses of THR, and any signal above will result in stimulation impulses of MCL. It can be observed that if the input audio levels are well below 75 dB the adaptive sound window will be situated between 10 dB and 65 dB.

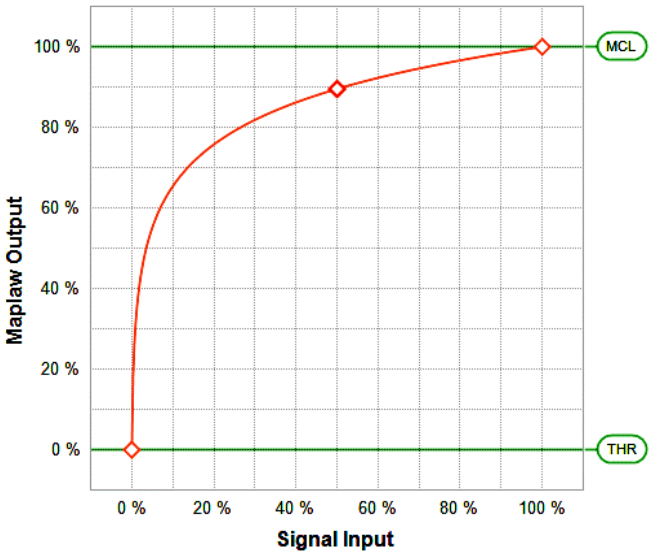


Figure 79: A typical logarithmic Maplaw compression curve, with a compression factor C=750 (from MAESTRO 4.0 fitting software), defined by equation (10)

Figure 79 depicts a typical logarithmic compression curve of the maplaw with the compression factor set to 750. The X-axis represents the level of the input signal resulted after the AGC, and the Y-axis shows the stimulation level between THR and MCL. In the given example a 50% input signal is mapped to a 90% stimulation level.

The MAESTRO fitting software maintains a database with each patient where the fitting specialist records the settings applied at each individual fitting session. The fitting software also allows the storage of audiograms.

5.2 Statistics of Typical Stimulation Levels

We have analyzed *historical fitting data of more than 150 pediatric patients collected over 10 years of fitting practice* (Stanciu & Hellmuth-Zweyer, 2015) (Stanciu, et al., 2011) (Stanciu, et al., 2008) (Stanciu, 2007). The analyzed data were exported from the MAESTRO Fitting software in XML format, and we have developed an application to read, interpret and calculate statistics over the exported data. The aim of this activities is to provide a reference for current and upcoming researches regarding cochlear implant fitting. All data were collected by a single fitting specialist, Dr. ing. Antonius STANCIU, during the initial routine fitting sessions. All patients went through a special incremental fitting procedure, aiming to preserve the integrity of the hearing nerves by providing the necessary time to learn and adapt to low stimulation levels, according to each patient learning curve.

Based on historical fitting database, the following useful statistics were generated:

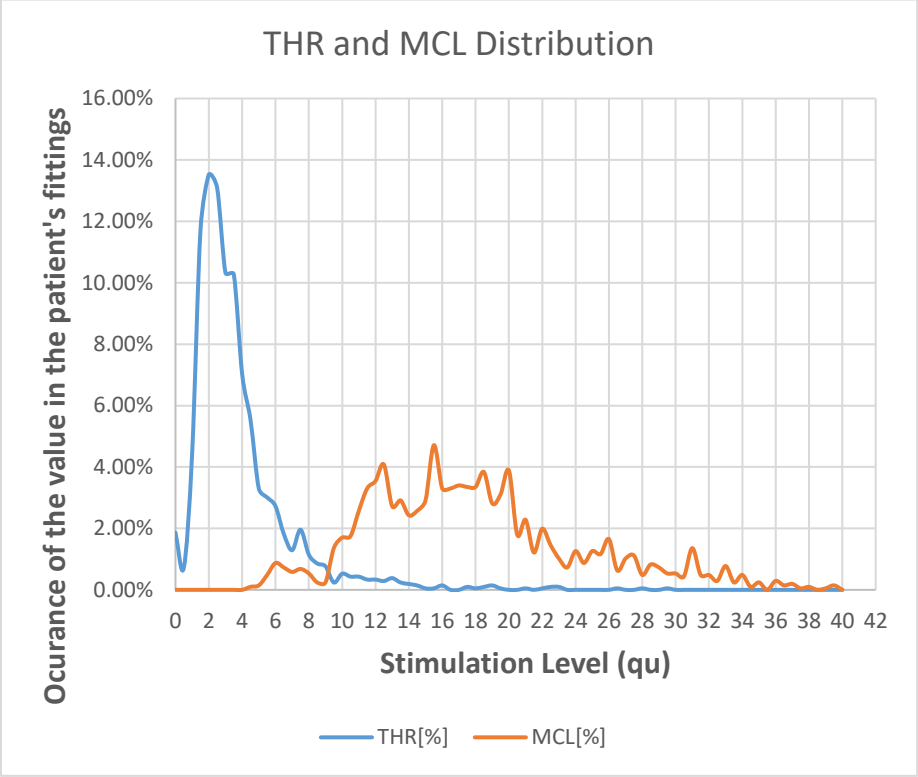


Figure 80: Distribution of THR and MCL levels for all channels. 1 qu (Charge Unit) is approximately 1 nC / Impulse

a) In the first chart (Figure 80), we have calculated and displayed the histograms representing the distribution of the used THR and MCL levels using 0.5 charge unit (qu) bin size. The distribution of the THR levels over the analyzed data set peaks at approximately 2 charge units. The shape of the distribution is rather narrow with only a small percentage of patients requiring significantly higher stimulation (more than 6 charge units). The MCL levels are more evenly distributed over a large domain of values, but similarly to the THR distribution, a small number of patients required MCL level over 25 units.

In our opinion, the shape of the THR and MCL distributions suggests that there is a typical range of THR and MCL levels, which should be aimed during the fittings. If satisfactory hearing levels can be obtained only with atypically high stimulation levels, it could indicate implantation issues or hardware defect of the cochlear implant.

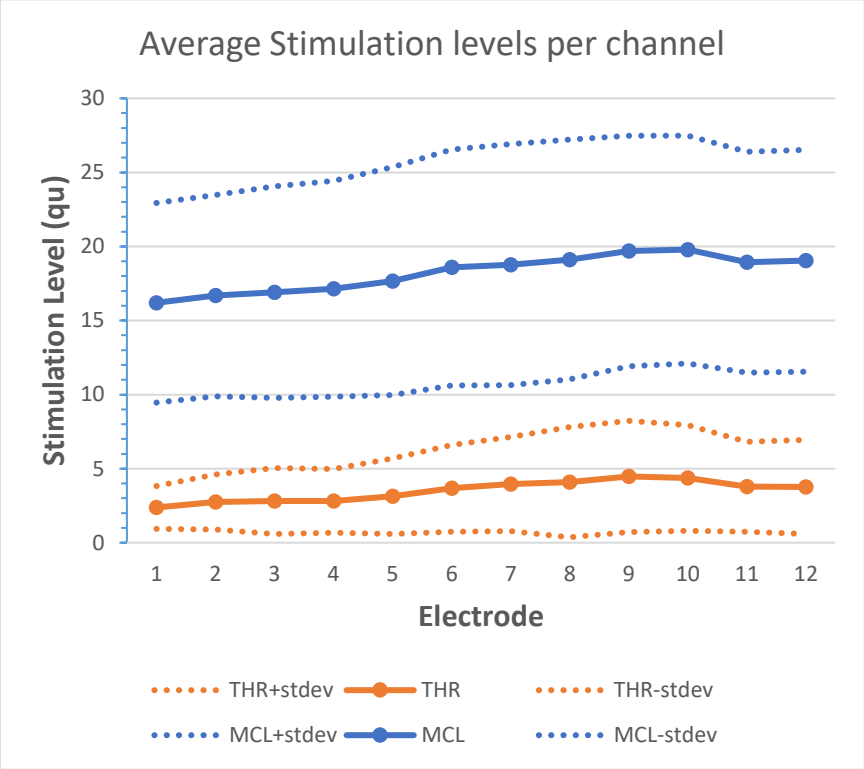


Figure 81: Average and standard deviation of THR and MCL stimulation levels per channel

b) Similar information can be seen on the charts depicted in Figure 81. In this case, the average and standard deviation of the THR and MCL levels is shown individually for each channel. A clear tendency can be observed showing that channels with *higher frequency requires slightly higher stimulation levels* (both THR and MCL).

Although typically hearing loss is installed first on the high-frequency range, this is due to the nature of the hair cells. In the case of cochlear implants, the hair cells are bypassed and the nerves are directly stimulated, therefore the corresponding frequency range should not affect the sensitivity. In our opinion, *this tendency could be created by the distance between the electrodes and the stimulated nerve fibers*. Indeed, the basal part of the cochlear duct, responsible for the high frequencies, is much wider compared to the apical part responsible for low frequencies. It is still a question whether the cochlear duct diameter variation can alone explain these tendencies.

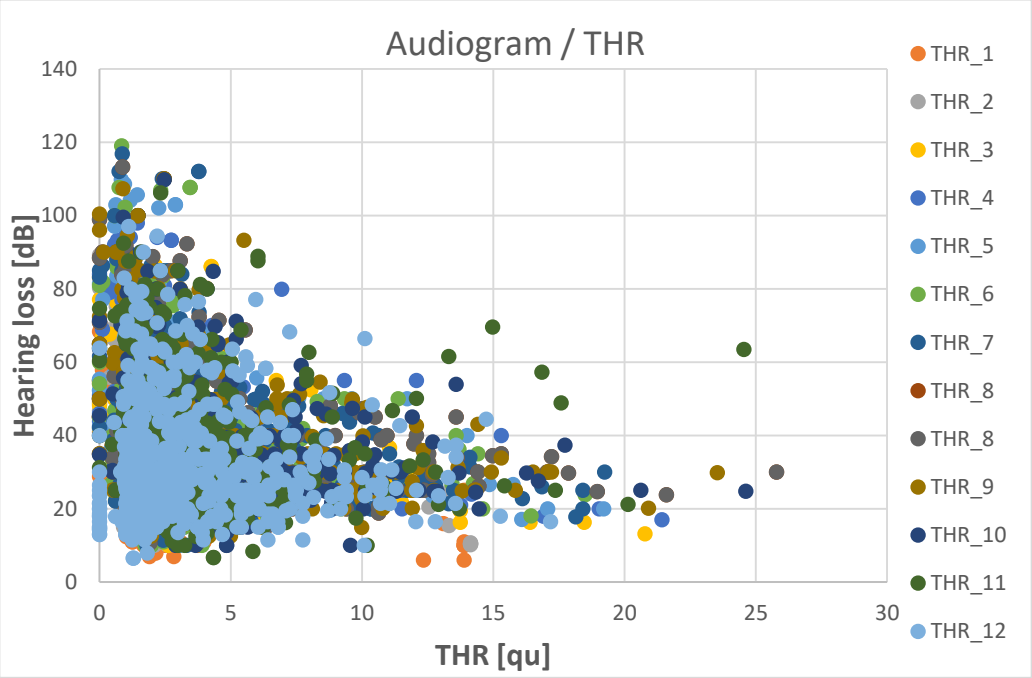


Figure 82: Hearing loss thresholds vs. THR plotted for all 12 individual electrodes considering fitting data of all patients

c) In the next figure (Figure 82), we have created a scatter chart showing the relation between the THR levels and the hearing loss threshold measured with free field audiogram. It was empirically observed, but now quantified, that *patients with low stimulation levels are better performers compared with those with high stimulation levels*. This observation is not easily explainable, given that all stimulation levels are fitted to match the needs of the individual patient, and the flat audiogram with minimal hearing loss is targeted. The content of the created scatter chart seams to underline the observations. It can be clearly observed that the *lowest hearing loss can be repeatedly achieved only with THR levels less than 5*. It is important to note that low stimulation levels not necessarily result in low hearing thresholds. In most cases, low THR levels are simply under the nerve stimulation threshold, making it impossible to perceive soft sounds.

Our interpretation of this chart is that the higher the stimulation level is, the higher the targeted hearing loss should be. In other words, one should not target restoration of hearing to 10- or 15-dB hearing loss once the THR levels are over 5, otherwise the overstimulation and even nerve destruction can happen.

5.3 Computer Assisted Fitting of Cochlear Implants – Tracking the Effective Stimulation Threshold

Auditory nerves are stimulated by an electrode array introduced in the cochlea along the basilar membrane. Because of the structure of the cochlea and the positioning of the electrodes in the cochlea, each individual electrode stimulates a different group of nerves generating different pitch sensations for the patient. In order to create meaningful hearing sensations, cochlear implants split sounds into frequency bands and translate the amplitude of the resulting signals into electric impulses of varying intensity and duration (Moctezuma & Tu, 2011). Typically, the impulse rates are fixed, but some CI models also transmit the resulting signal frequency by modulating the rates of the impulses (Moctezuma & Tu, 2011)

As the required stimulation intensities, to evoke hearing sensations, vary from patient to patient, and also vary over time, periodic fitting of the parameters of the cochlear implant is required. Also, in the period immediately after the implantation, the stimulation intensities must be increased gradually from a very low level to a comfortable one in order to prevent overstimulation and distress of the patient (Vargas, et al., 2012) (Gross, 2003).

The observations and problems discovered during the interviews and discussions with existing patients lead to the development of a software tool aiming to assist the work of the fitting specialists (Kuczapski & Stanciu, 2015). In the following, the details of this work are shown.

5.3.1 Fitting Techniques and Procedures

In the case of adult cochlear implant patients, THR and MCL levels are set using verbal feedback of the patient, whereas in the case of pediatric patients the fitting procedure is more difficult (Gross, 2003). There are three categories of procedures used in pediatric cochlear implant fitting:

Behavioral Measures — Behavioral measures require patients to indicate voluntary or involuntary that a sound stimulus was perceived. Young children, as part of their intensive auditory trading, are taught to react to sounds during specific games. This allows the fitting specialist to estimate adequate THR levels.

Objective measurements — A variety of objective measurement techniques were developed to estimate the correct THR and MCL levels. MCL level can be well estimated by the Electrically Evoked Stapedius Reflex Threshold (ESRT) measuring the *stimulation levels which result in muscle contractions in the middle ear*. In case of ESRT measuring, THR levels are set to 10% of the estimated MCL levels (Kosaner, et al., 2009). Another attempt to obtain objective THR and MCL measurement is by the means of Neural Response Telemetry (NRT). NRT records the variations of nerve action potential as a response to electric stimulation trying to estimate suitable THR and MCL levels. Studies showed that the results of NRT are poor and shall be used with precaution (Kiss, et al., 2003) (Caner, et al., 2007).

Speech Perception Measurements — Speech perception test measures the capacity of the patient to discriminate phonemes, individual words and words in context. The fitting specialist can fine-tune the stimulation levels and the compression AGC and maplaw by observing the frequency and intensities of phonemes which are not differentiated by the patient.

5.3.2 Fitting Based on Audiograms

The presented fitting techniques mainly focus on the correct setting and leveling of the MCLs. Although it is shown that THR levels are not very important regarding speech perception in quiet (Gross, 2003), according to our observations (Kuczapski & Stanciu, 2015) (Stanciu, et al., 2011) (Stanciu, 2007), it has an important role in the perception of speech and soft sounds in real-life conditions. The right setting of THR levels is difficult because, in order to avoid tinnitus, *THR must be under the hearing threshold level*. Also, the testing of electrodes individually does not reflect real-life condition, in which adjacent electrodes simultaneously stimulate nerve cells reciprocally increasing the resulting stimulation levels.

In order to find good THR levels, according to our experience, an audiogram obtained through *open air pure tone audiometry* is necessary. In case of young children, it is mandatory to teach them as soon as possible to cooperate through play and provide voluntary behavioral response.

After the implantation, the THR and MCL levels are gradually increased during several weeks, till the first behavioral response is recorded, and then the THR levels are adjusted gradually to improve the audiogram to normal levels. The *MCL level is set* through simple behavioral measures, although initially they are kept *at levels below those provoking stapedius reflex*.

It was observed that, at low levels of stimulation, the sensibility of auditory nerves may improve.

5.3.3 Effective Stimulation Threshold

Starting from the audiogram, the MCL and THR levels and the maplaw settings, it is possible to compute the Effective Stimulation Threshold (EST) for each individual electrode, which represents the *stimulation level that causes conscious hearing sensations* to the patient. We think that the resulting EST levels are good indicators of the sensibility and health of the stimulated auditory nerves. In order to compute the EST levels, first we have to formalize the maplaw curve and the AGC compression. We found that the logarithmic maplaw curves can be matched with the following equation:

$$o = \log_{C+1}\left(1 + C * i\right) \tag{10}$$

where *C* is the maplaw compression factor, *i* is the normalized signal input and *o* is the maplaw output. Both *i* and *o* take values in the range of [0…1]. Applying the AGC compression and adding the THR and MCL levels to the equation, we find the following formula to compute the EST levels:

$$EST = T + (M - T) * \log_{C+1}\left(1 + C * 10^{A-65/20}\right) \quad (11)$$

where *T* is the THR, *M* is the MCL, *C* is the maplaw compression factor, and *A* is the audio level in decibels indicated by the audiogram as hearing threshold. The settings of the AGC are not taken into consideration in the estimation of EST, because during the audiometry the patient is exposed to sound at low amplitudes, therefore the AGC does not attenuate the input signal.

Based on equation (11), we developed an application that collects the THR, MCL and maplaw settings from the MAESTRO fitting software database, and after the fitting specialist introduces the new audiogram, it *displays the EST values over the THR and MCL settings*.

The left side of Figure 83 displays the MCL and THR settings of each electrode in a form of a bar chart, where the bottom of the bar indicates the THR level and the top indicates the MCL level. The bars are also colored so that they are split in four regions according to the maplaw set: RED indicates the levels where sounds under 20dB are mapped; GREEN indicates stimulation interval for sounds between 20dB and 30dB; YELLOW indicates stimulation interval for sounds between 30dB and 50dB; and GREY indicates stimulation interval for sounds above 50dB. The calculated EST is displayed in a form of a red curve over the bar chart of the THR and MCL levels. The audiogram used to compute de EST levels is displayed on the right side of Figure 83a.

Through investigation of the recorded fittings and audiograms, it was observed that through fitting, the *EST levels shall be positioned in the GREEN area* of the stimulation levels, resulting in an *average hearing loss of 20–30dB*, which is equivalent to a mild hearing loss.

Furthermore, *using the software, the fitting specialist can propose new THR, MCL and maplaw settings based on the current EST levels*, and the *system computes the expected change in the audiogram* displayed with red on the right side of Figure 83b.

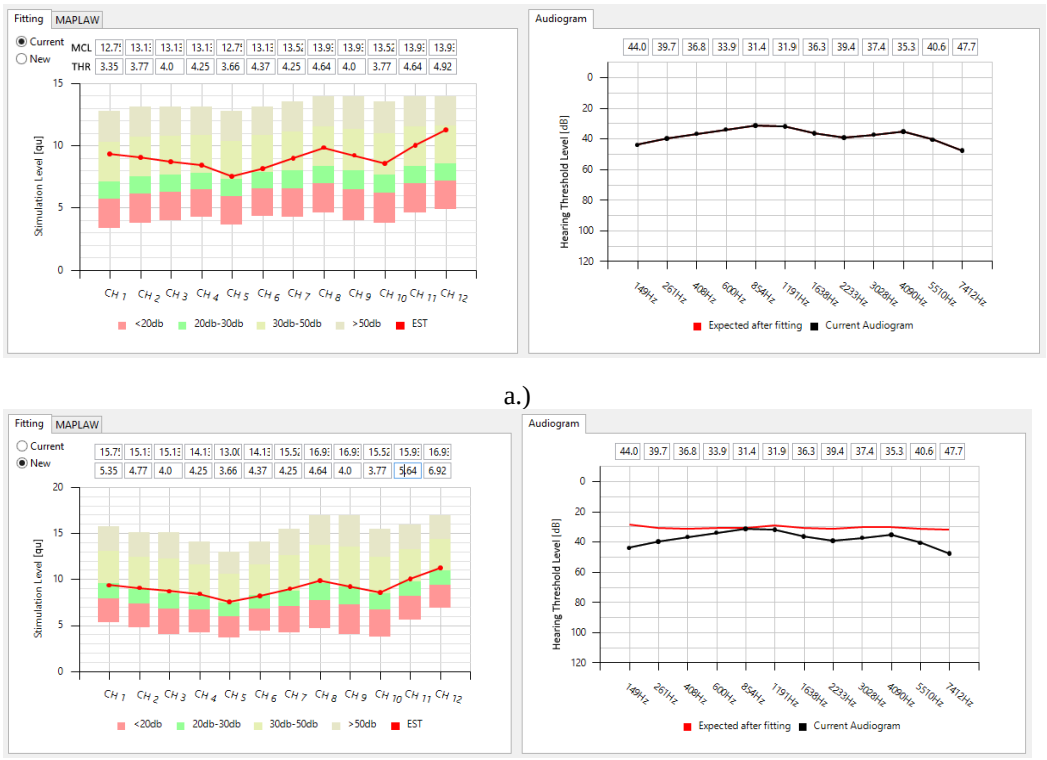


Figure 83: Screenshots of the assistive software tool displaying the MCL, THR and EST levels (left), and the audiograms (right): a.) Current fitting (left) and corresponding audiogram (right); b.) New proposed fitting (left) and predicted audiogram compared with the previous audiogram (right)

5.3.4 Monitoring of EST

Another function of the proposed software tool, is to compute and visualize the long-term evolution of the EST levels, with useful implications. When the stimulation levels of an electrode are selected to be edited, the application displays on a line chart the evolution of the THR, MCL and EST levels from the activation of the last entry in the database.

It was observed that typically *EST levels are decreasing* in the months following the first perceived hearing sensations and *finally become stable at a much lower level than originally appeared*, but over stimulation can stop this decrease of EST levels. Deterioration of the auditory nerves immediately increases the EST level, and also technical defects of the external processing is reflected though apparent EST level changes.

Figure 84 illustrates the typical evolution of the EST levels. For the first 3 fittings there are no meaningful audiograms, but starting with the 4th fitting, hearing sensation starts to appear and the EST level start to decrease, and finally after several months, the EST level is stabilized to a value much lower than the initial stimulation values which resulted in no hearing sensation.

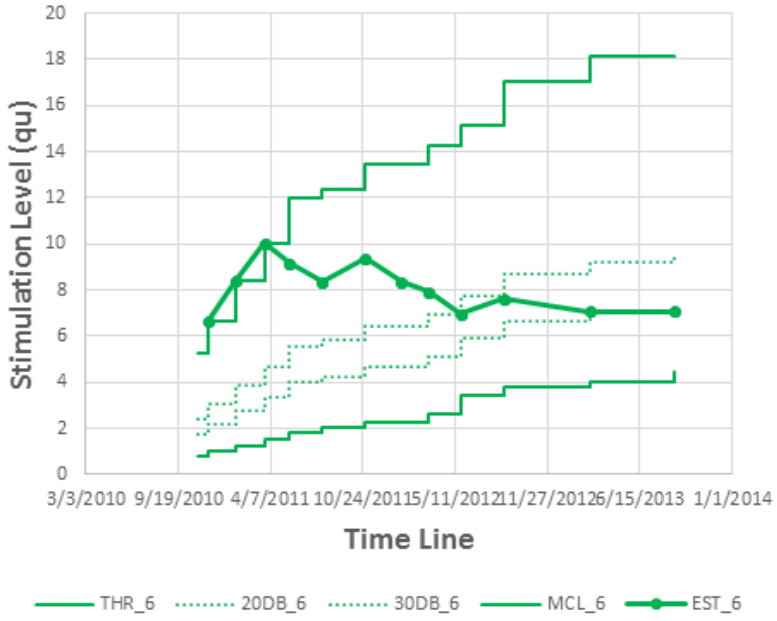


Figure 84: Typical evolution of EST levels

Figure 85 exemplifies an abnormal evolution of the EST levels. The *sudden increase of the EST level indicates a biological process that reduces sensitivity* or effectiveness of the electrodes. Besides other medical conditions, two possible causes can lead to such results: degenerative process of the nerve fibers or accelerated tissue growth around the electrode array.

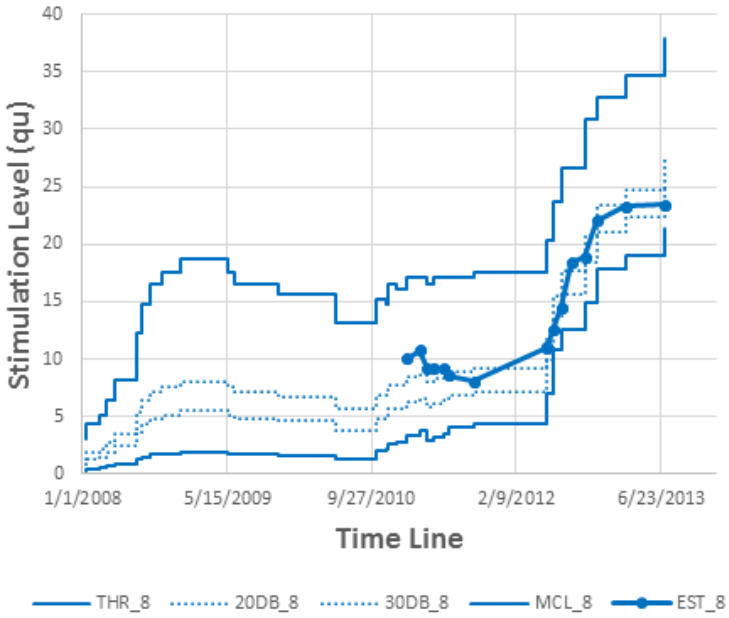


Figure 85: Degenerative process shown by increasing EST levels

5.3.5 Contributions

Cochlear implants (CI) are implantable electronic prosthetic devices developed to restore hearing in patients with severe to profound hearing loss, hearing sensation being restored by electric stimulation of intra cochlear nerve tissues.

In order to obtain useful hearing sensations, electric stimulation levels and signal processing parameters must be custom fitted for each individual patient. The correct fitting of the CI being one of the key factors for successful cochlear implantation. The fitting of the cochlear implant is done by trained fitting specialists, however, there are no well-established the proven protocols to maximize chances of success. Applied fitting protocols and principles vary from one specialist to the other, and largely depend on personal expertise. Difficulty of right CI fitting is increased when CI recipients are small children.

Various fitting methods, involving subjective and objective measures, are known in the literature, but these methods are mostly related to the fitting of the upper limit of the stimulation levels (MCL — Maximum Comfortable Level), making the fitting of the lower boundary of the stimulation levels (THR) less than optimal.

Based on the fitting experience of more than 150 children, we observed that, even if correct MCL level settings lead to good speech perception in optimal listening conditions, *correct THR level settings are required for hearing of soft sounds or understanding of speech in real-life listening conditions*.

This section first introduces a *new derived fitting coefficient - the effective stimulation threshold (EST)*, then describes the *developed assisted software for cochlear implant fitting* to obtain better results in fitting of Maestro Cochlear Implant Systems. Finally, the benefits of monitoring of the EST level evolution are shown.

The *developed software* loads the database of the MAESTRO fitting software and reads the MCL, THR, maplaw settings and registered audiograms of all fitting sessions of the selected patient, then *computes the history and current levels of EST*. After the fitting specialist introduces the *newly proposed fitting parameters, the software displays the expected changes of the audiogram*. EST history of each electrode is displayed, allowing the specialist to identify pathological processes or implant defects.

The main contributions in this section are:

- *New fitting coefficient - Effective Stimulation Threshold* (EST) (11) that represents the stimulation level to cause conscious hearing sensations to the patient. The EST is determined from open air tonal audiograms correlated with the THR, MCL and maplaw settings;

- *Development of an assisted software for cochlear implant* fitting to help fitting specialists to improve fittings. This software can be used to *predict new threshold* hearing levels and to monitor the long-term evolution of EST levels.

- Exemplification of the benefits of EST evolution monitoring in early *identification of CI technical defects or pathological* processes.

5.4 Case Studies of Fitting Evolution

Real field fitting experience suggests that the hearing quality and thresholds can be improved over time without changing the fitting parameters, but in the same time, over-fitting may cause impaired hearing and can permanently damage nerve sensitivity to electric stimulation. In order to study these suppositions, we have selected and presented a few patients, with long history records and known performances and pathology (Kuczapski & Stanciu, 2015).

We have created 2 chart types for each patient showing the evolution of the hearing levels from two different perspectives:

- **Time series of THR, MCL and EST levels** – showing the evolution in time of the monitored levels: trends, change velocities and convergence.
- **THR vs Hearing Loss scatter chart** – showing the relation between Hearing Loss and THR. Ideally, the points of the chart should be arranged in a monotonically descendant curve. If the curve starts to increase it could indicate overstimulation or other problems.

All case studies refer to children with congenital deafness and early implantation - between the ages of 1 and 2 years at the time of the implantation. In all case studies, a *fitting strategy with slow activation and increased simulation level* was applied.

5.4.1 Case 1 – Excellent performer with bilateral implants

Patient: 6-year-old, Female, Bilateral Implants.
Ear studied: Left
Electrode studied: 6

The first case study refers to a patient with bilateral cochlear implant with excellent levels of hearing restoration and performance.

Figure 86 show the evolution of the fitting level and the EST level for the 6th electrode computed based on audiograms. The chart comprises of 5 curves:

- THR levels
- MCL Level
- Electric Stimulation levels at 20dB Sound levels
- Electric Stimulation levels at 30dB Sound levels
- EST Levels

The first measurable response to pure tone stimuli was measured after the 5th fitting session with EST = MCL = 9. Within next three fitting, even though the stimulation levels were only moderately increased, the EST level dropped significantly. In the next four fittings, the EST levels remained unchanged and indicating a hearing loss of more than 30dB. According to the *fitting strategy, the THR and MCL levels were increased bringing the EST levels between the stimulation levels equivalent to 20 and 30 dB.* The EST levels were mostly descendant excepting the last fitting in the recordings.

The second chart (Figure 87) indicates, that up to THR = 3, each increase in the THR level was beneficial to the improvement of the hearing loss.

In Figure 88 and Figure 89 the evolution of the fitting levels and the relation between the *measured hearing loss and THR levels* is depicted for all 12 channels. It can be observed that all channels are displaying similar trends of evolution.

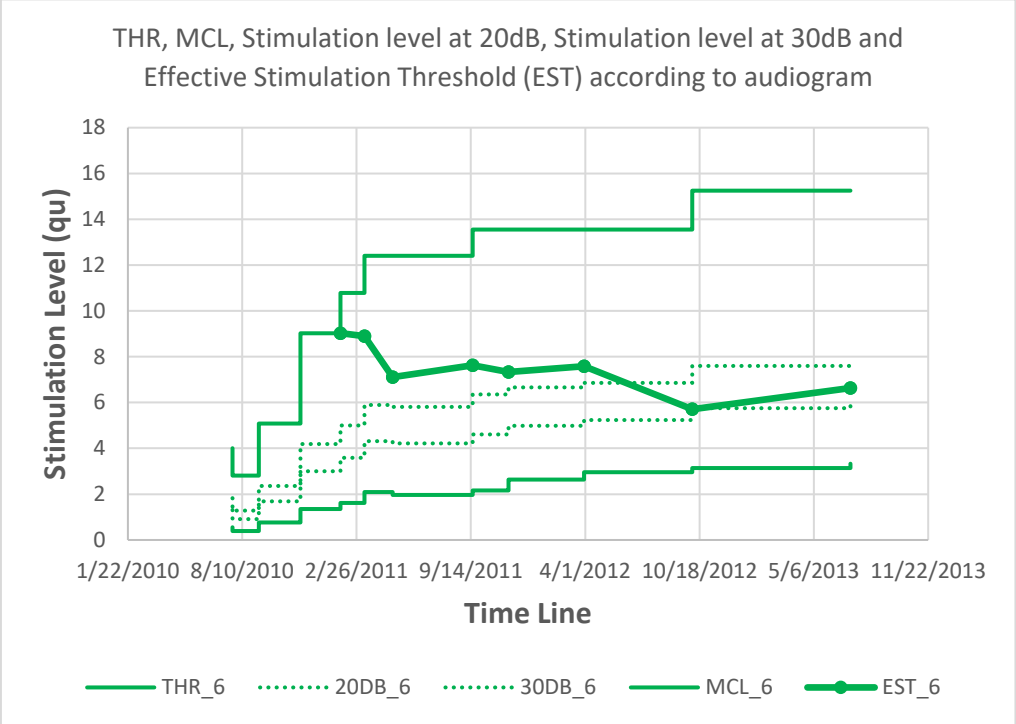


Figure 86: Case 1 – Evolution of THR, MCL and EST recorded for the 6th electrode

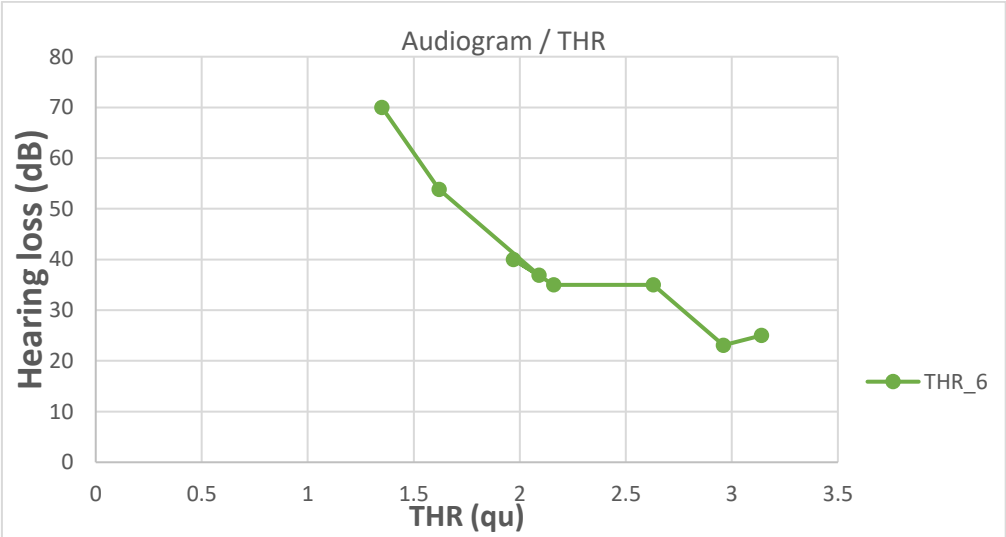


Figure 87: Case 1 – Relation between Measured Hearing Loss and THR level recorded for the 6th electrode

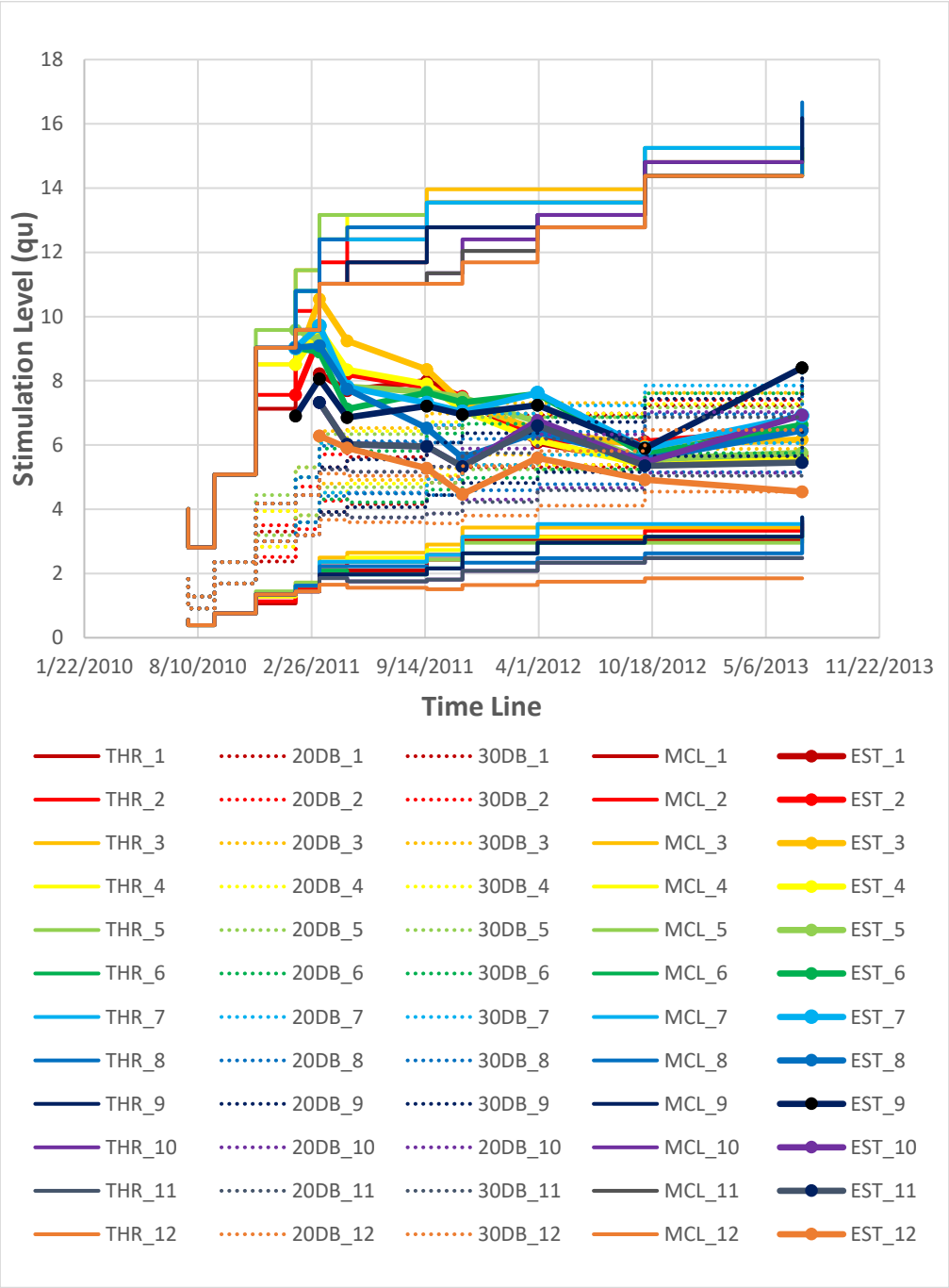


Figure 88: Case 1 – Evolution of THR, MCL and EST recorded for all electrodes.

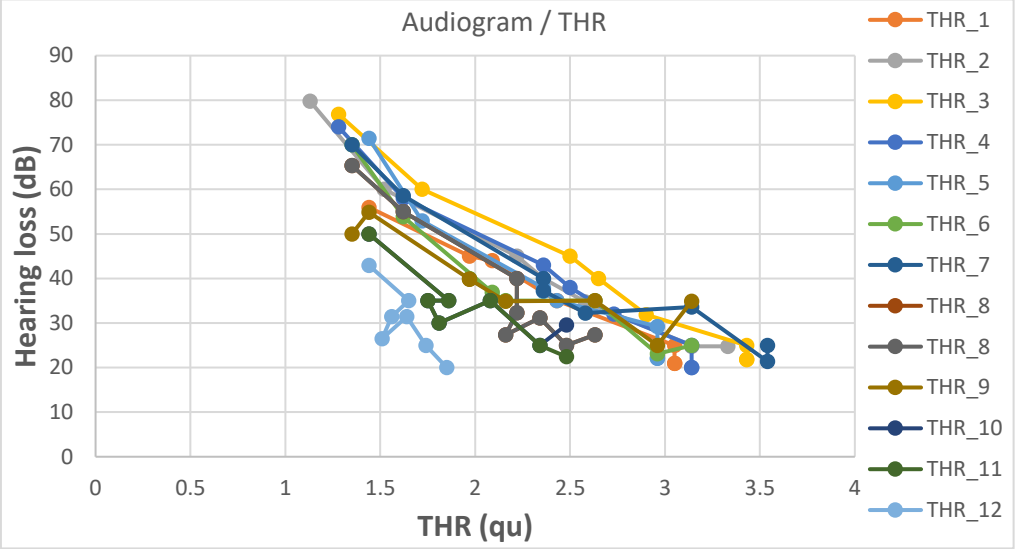


Figure 89: Case 1 – Relation between Measured Hearing Loss and THR level recorded for all electrodes

5.4.2 Case 2 – Average to good performer with unilateral implants

Patient: 6-year-old, Male, Unilateral Implant.
Ear studied: Right
Electrode studied: 6

In this case, we have studied the recorded fittings of a patient with unilateral implant, with average to good performance.

The evolution of the stimulation and EST levels (Figure 90) are very similar to those recorded for Case 1. A clear convergence can be observed on the EST levels, and the *scatter chart* (Figure 91) *clearly indicates by the monotonic decrease*, that the current hearing level were obtained without overstimulation.

Same observations can be made over the charts representing all channels (Figure 92) (Figure 93).

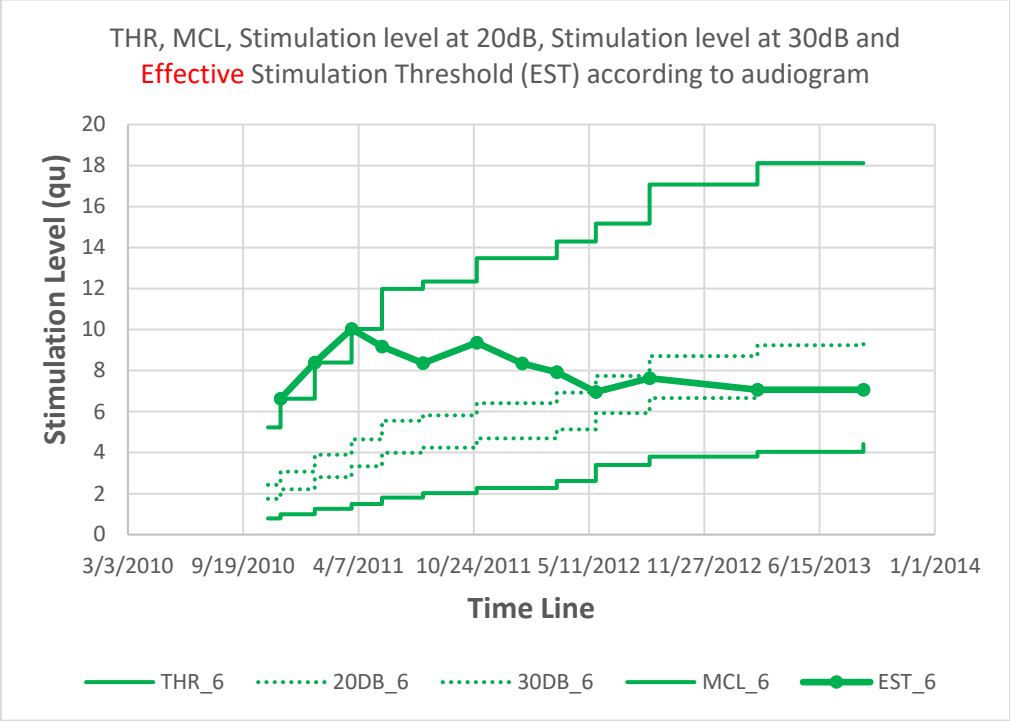


Figure 90: Case 2 – Evolution of THR, MCL and EST recorded for the 6th electrode

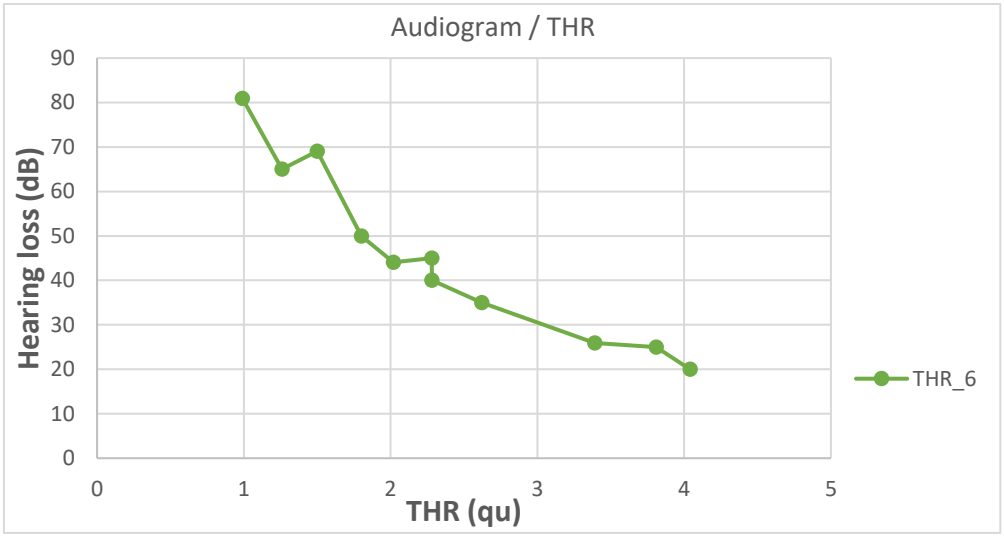


Figure 91: Case 2 – Relation between Measured Hearing Loss and THR level recorded for the 6th electrode

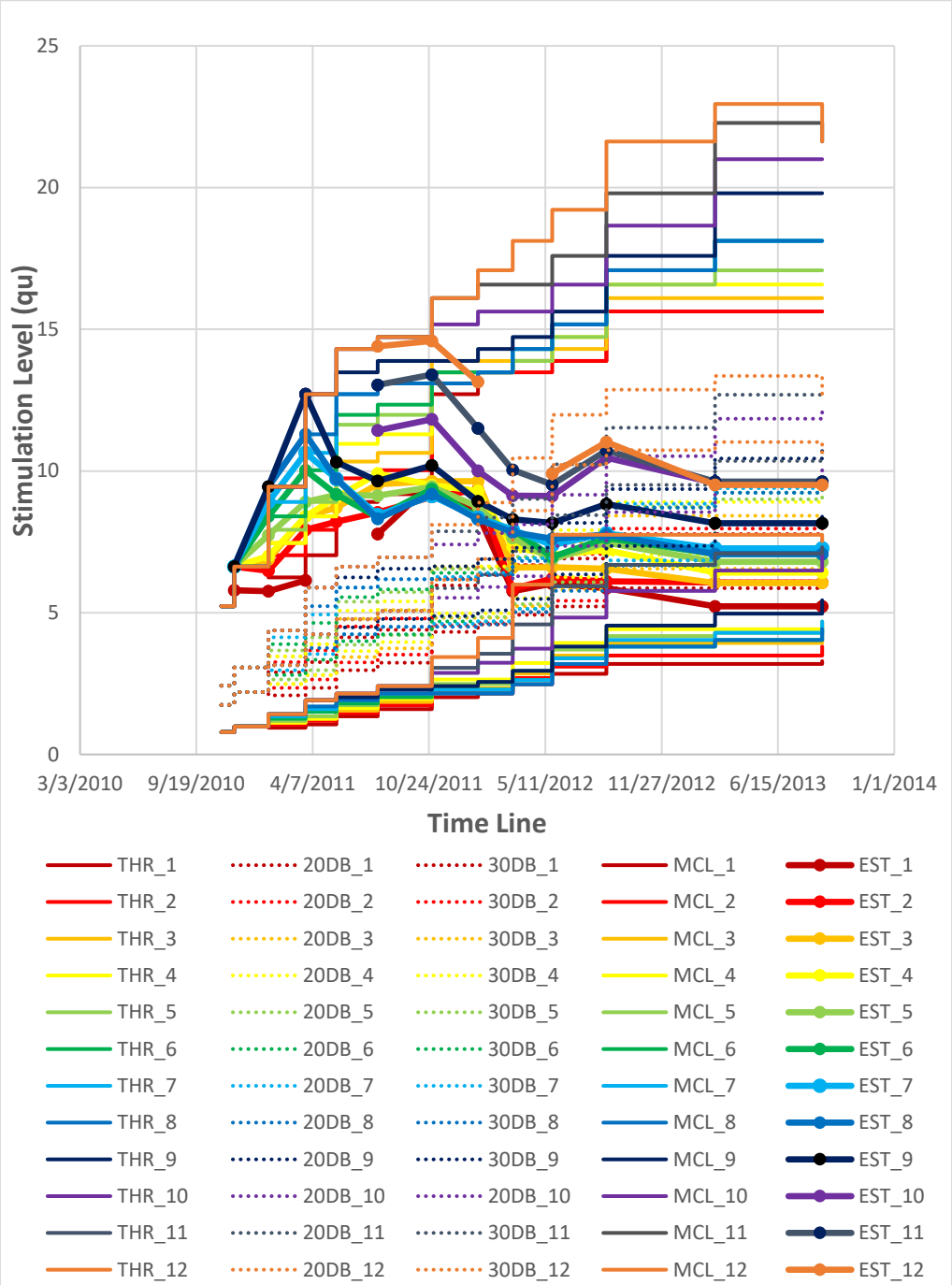


Figure 92: Case 2 – Evolution of THR, MCL and EST recorded for all electrodes.

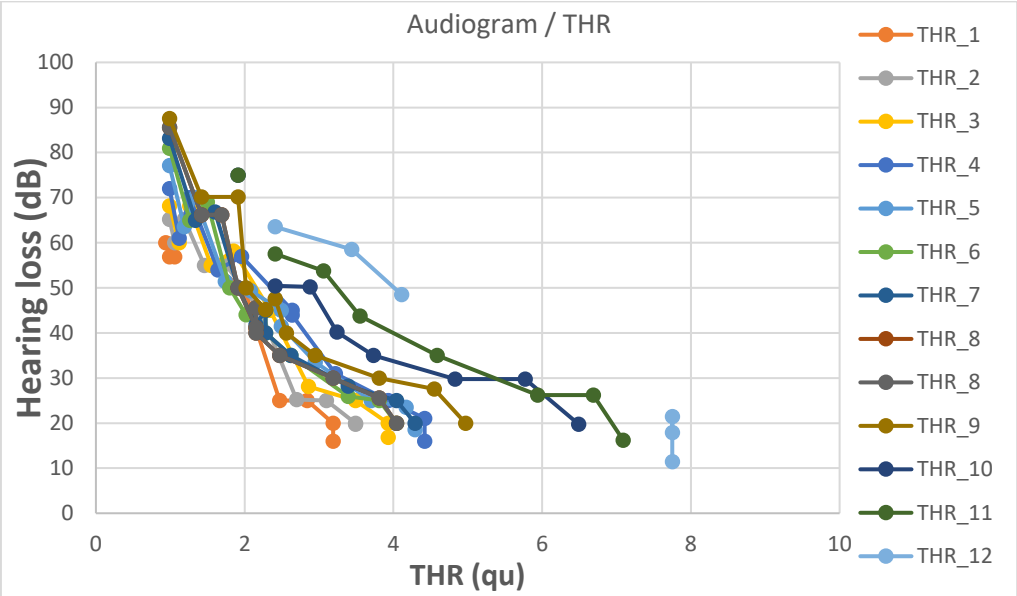


Figure 93: Case 2 – Relation between Measured Hearing Loss and THR level recorded for all electrodes

5.4.3 Case 3 – Good performer with bilateral implants, with sudden decrease in hearing quality

Patient: 8-year-old, Female, Bilateral Implants.
Ear studied: Right
Electrode studied: 8

We have studied the recorded stimulation levels of a patient, who had *good hearing performance for several years, and then it started to drop*. Any attempt to improve the worsened hearing levels were without any significant success. The complete record of the audiograms is not available for this patient, but what we can see in Figure 94 is that the evolution of the EST levels were normal at least for a period. After years of utilization, a significant increase of hearing loss was reported by the audiologist and immediately corrected by the fitting specialist by increasing the stimulation levels. The increase of the stimulation levels does not produce the expected results, and the EST levels were abruptly increased. This measurement and adjustment cycle has been repeated several times until the situation has been stabilized, leading to very high increases in the stimulation levels, especially the THR levels which were increased 4 times compared to the initial configuration. The struggle

to bring down the hearing loss levels is best seen in Figure 95. The EST levels shown in Figure 94 are increasing together with THR levels showing that the adjustments are not useful.
Most probably, the situation was caused by *some degenerative processes within the cochlea*, but it is not clear if the expectation of the patient to restore the hearing level to previous values has worsened it or not. If such charts were available during the fitting session of this patient, the fitting specialist might have chosen a different way to react to the increase of the hearing loss.
Figure 96 and Figure 97 shows that the *same degenerative process affects the other electrodes too*, therefore increasing the probability of cochlea wide degenerative process.

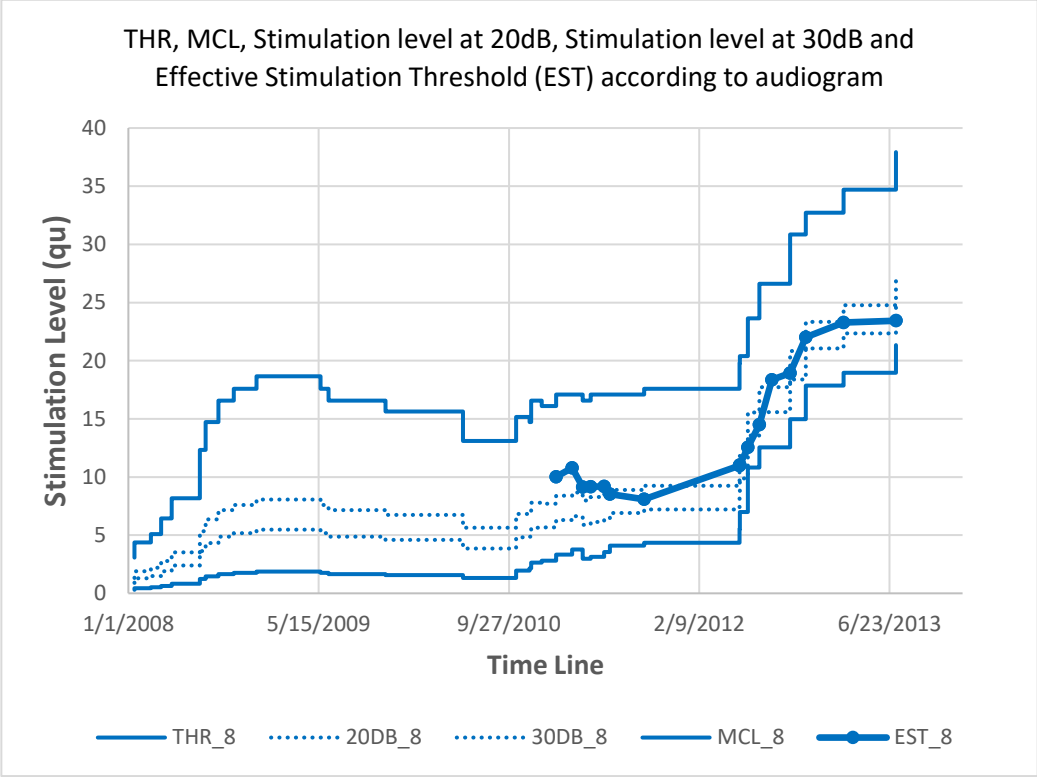


Figure 94: Case 3 – Evolution of THR, MCL and EST recorded for the 8th electrode

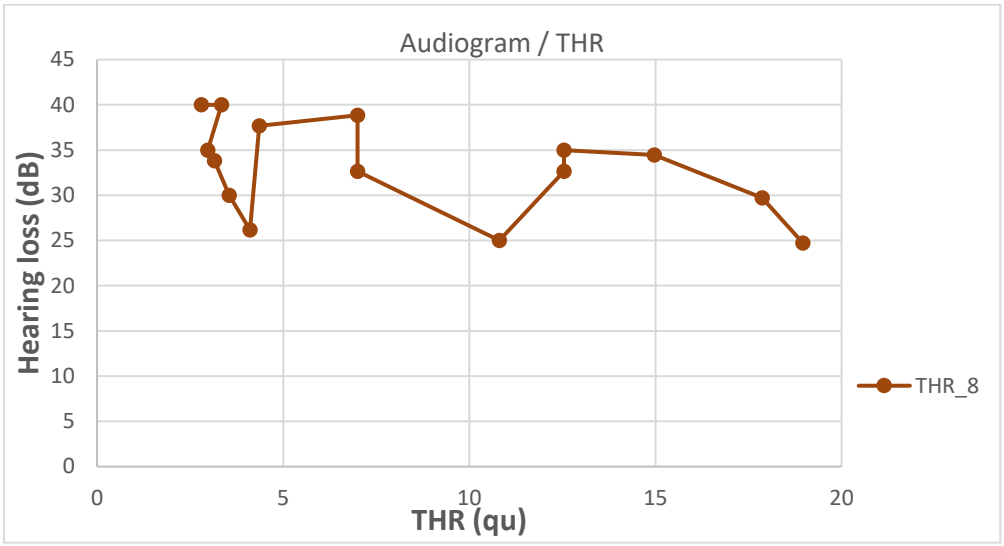


Figure 95: Case 3 – Relation between Measured Hearing Loss and THR level recorded for the 8th electrode

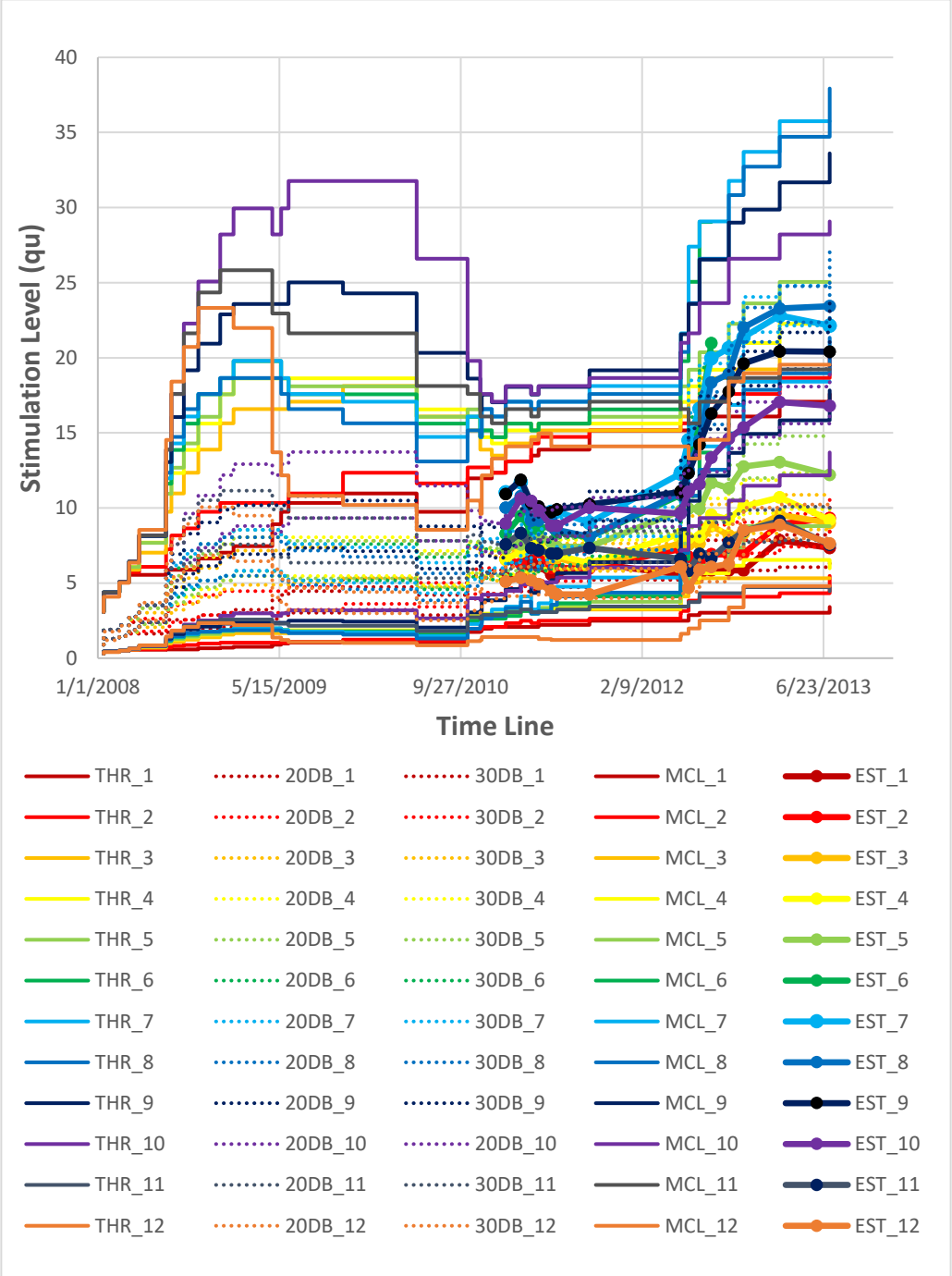


Figure 96: Case 3 – Evolution of THR, MCL and EST recorded for all electrodes

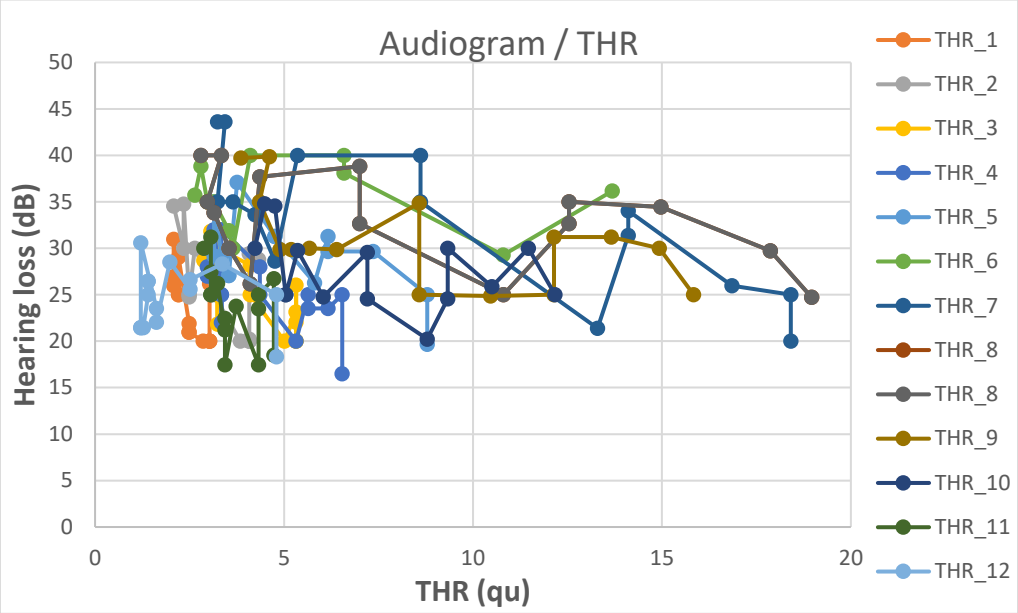


Figure 97: Case 3 – Relation between Measured Hearing Loss and THR level recorded for all electrodes

5.4.4 Case 4 – Good performer with unilateral implants, with sudden defect of the cochlear implant processor

Patient: 8-year-old, male, Unilateral Implant.
Ear studied: Left
Electrode studied: 6

The last case study shows the fitting recordings of a *defective cochlear implant processor*. After a *good evolution of patient, suddenly the audiologist reports significantly worsened hearing loss levels*. The fitting specialist tries to slightly increase the THR levels and in the meantime discovers that the MCL levels must be reduced, otherwise the stimulation is too loud and unbearable for the patient. After trying out the new fitting, the patient returned to a new session, and the processor was changed to a new one. After the processor change, the THR, MCL and EST levels are reverted to the previous values.

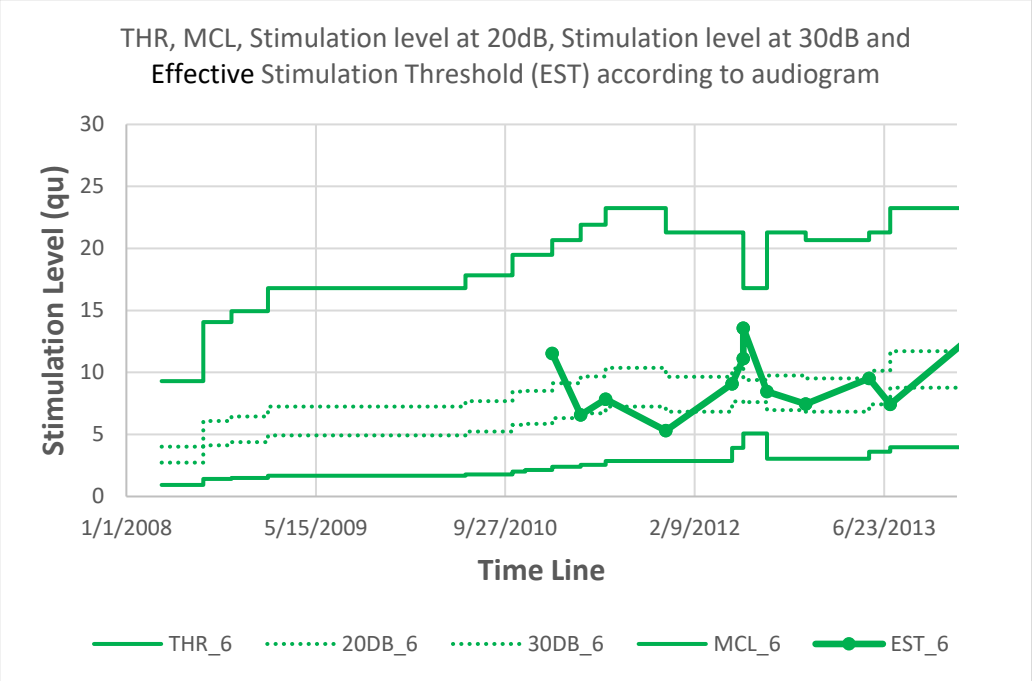


Figure 98: Case 4 – Evolution of THR, MCL and EST recorded for the 8th electrode

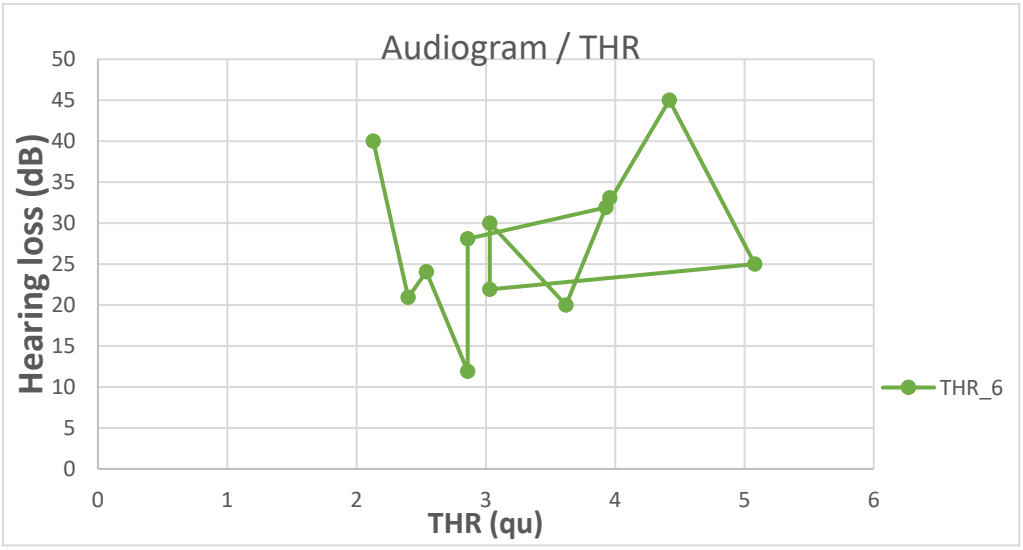


Figure 99: Case 3 – Relation between Measured Hearing Loss and THR level recorded for the 8th electrode

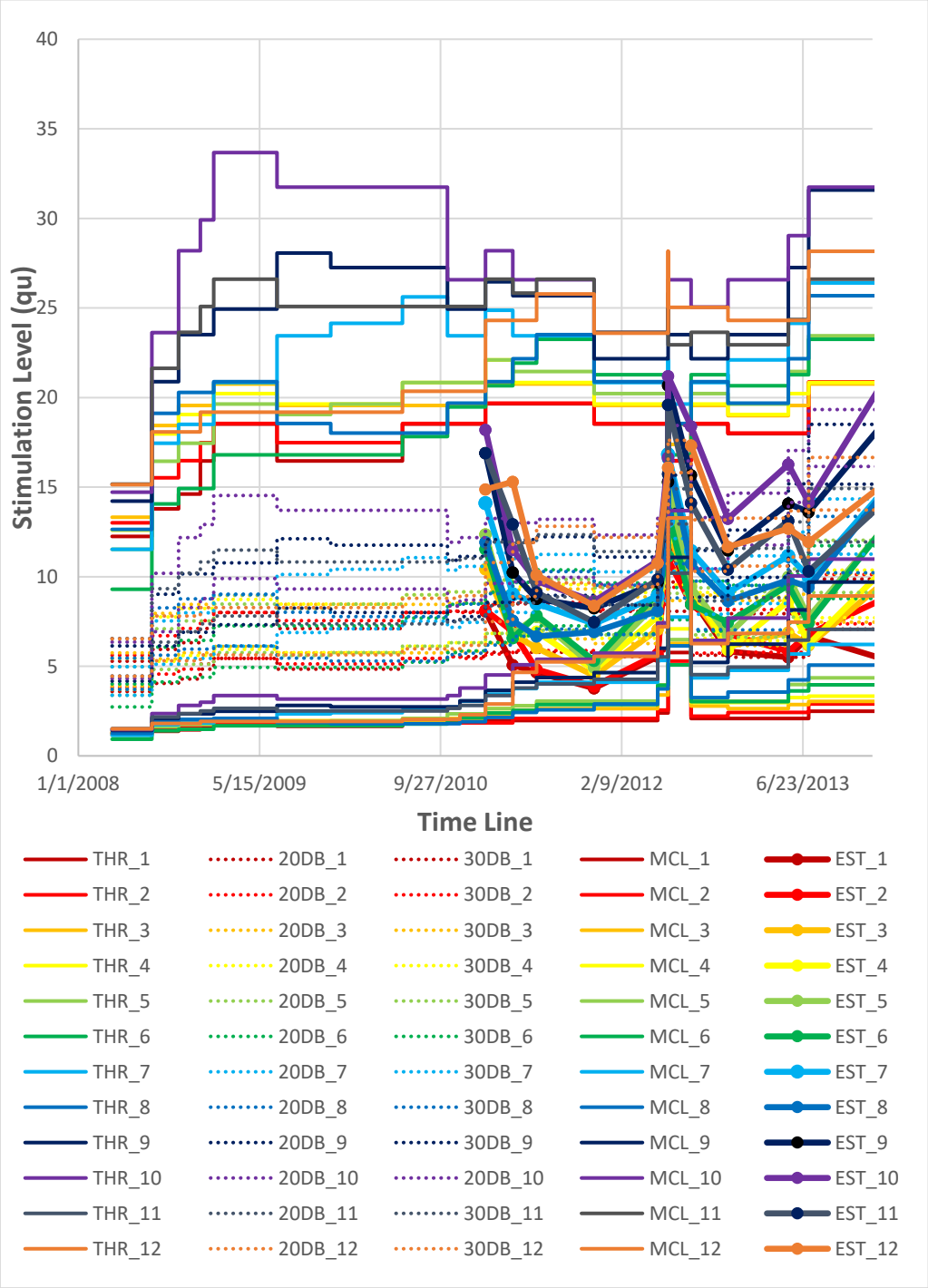


Figure 100: ***Case 4 – Evolution of THR, MCL and EST recorded for all electrodes***

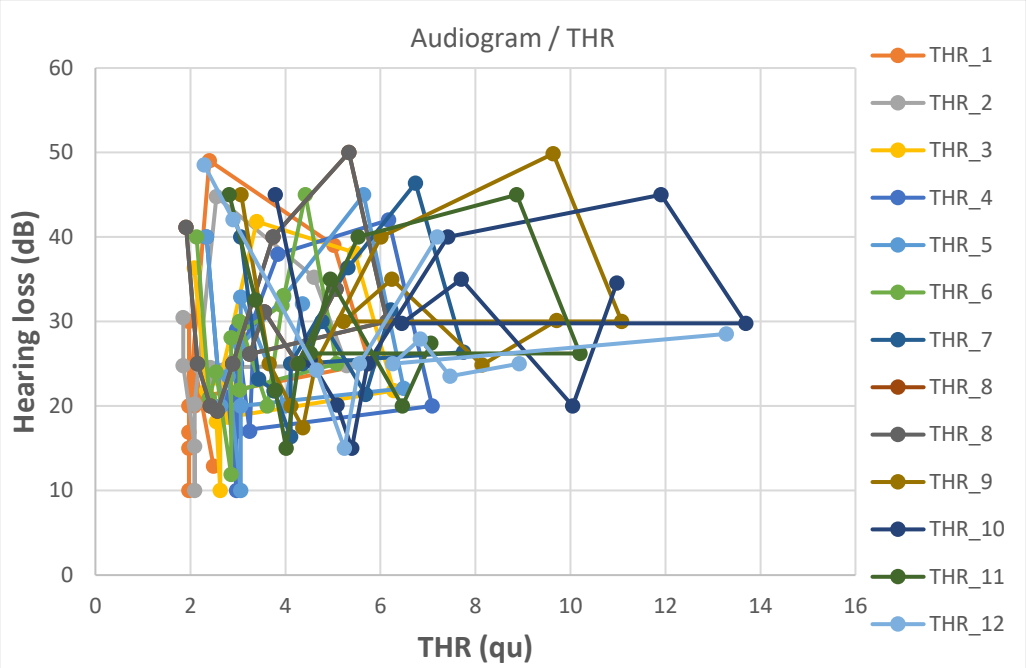


Figure 101: Case 4 – Relation between Measured Hearing Loss and THR level recorded for all electrodes

5.5 INTERFACING WITH MED-EL COCHLEAR IMPLANT PROCESSORS

5.5.1 Introduction

The present chapter presents the development of a tool in the cochlear implant area (Kuczapski & Andreescu, 2016) - a *real-time data acquisition system designed to register stimulation pulses from most MED-EL CI processors*. This system is used for:

- Technical support for further researches in CI;
- Fault detection used during fitting procedures;
- Simple auralization for demonstration purposes.

The proposed system mainly contains: i) I^{100} Detector Box of MED-EL that transduces the information received form the CI processor antenna into electric pulses, similar with an implanted receiver/stimulator; ii) 12 channel real-time analog data acquisition module developed using ADC from Arduino Due and a custom developed shield; iii) PC connected through USB to the data acquisition module, running a Java software developed for control, real-time visualization and auralization.

5.5.2 The I^{100} Detector Box

The I^{100} detector box is a table top device developed by MED-EL that is used by technicians in cochlear implants as an assistive tool for training or visualization. This device is an implant emulator that replicates the functionalities of a real cochlear internal implant and provides dedicated electric interfaces for measuring or registering generated electric pulses (MED-EL(3), n.d.). The main functions of the device are:

- Generate electrical stimulation pulses based on the information received from the CI external processor;
- Simulate a cochlear implant during telemetry;
- Testing and fault detection of external components of a MED-EL Cochlear Implant system.

Figure 102: I^{100} detector box (MED-EL(3), n.d.)

Figure 102 shows the I^{100} detector box with its connectors (1,2,3) and function selection commutators (4,5). There are 12+1 pins (2) representing the 12 electrodes and the ground of a cochlear implant, that generates electrical stimulation pulses. These pulses can be also collected through a DB15 female connector (1). It is possible to collect or measure the signal of a single electrode through the BNC connector (3). In this case the commutator S2 (5) is used to select which channel is connected to the BNC connector (3). The commutator S1 (4) can be used to simulate interruptions of the first electrode or short circuits between the first and second electrodes. By design, the I^{100} detector box is intended to be used as a standalone device for cochlear implant telemetry training, or in combination with an oscilloscope for visualization purposes.

5.5.3 Interfacing with CI Processors

For research purposes, a hardware and software system is developed, which is capable of collecting, registering and displaying the signals from all 12 electrodes. According to experimental observations on the generated electric pulses, the typical length is between 10 μs to 20 μs, the amplitude is less than 10V and, as already known in (Zeng, 2008) and (Wouters, et al., 2015), it is balanced biphasic with a 0V DC component. Therefore, in order to record meaningful signals, a 12 channel A/D converter is needed with a sampling rate around 100K samples per second per channel.

To realize the *data acquisition module*, an *Arduino Due* designed development board is used which has *12 analog inputs* connected to a single multiplexed A/D converter with *12 bits precision* and capable of a *1MSPS conversion rate*. The Arduino Due development board is capable of continuously transmitting data to a computer through the high-speed native USB interface managed to achieve a *1.5 MByte/sec transmission rate*.

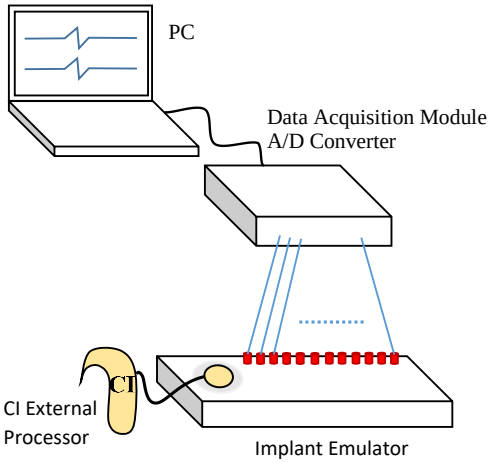


Figure 103: Block structure of the acquisition system

Figure 103 shows the structure of the real-time data acquisition system with the main modules. The cochlear implant external processor it connected to the I^{100} detector box through the electromagnetic couple that provides power to the detector box and also transmits the commands to generate pulses. The generated pulses are connected to the data acquisition module, are sampled and digitalized and sent to the computer by USB interface.

As the generated pulses are biphasic and the amplitudes are much higher than the maximum allowed 3.3V input voltage of the A/D converters, an *extension board* (shield) is developed as shown in Figure 104. It comprises of *12 voltage divisors with 10:1 ratio, and a voltage shift*, which provides a *virtual ground shifted to 1.65 V*.

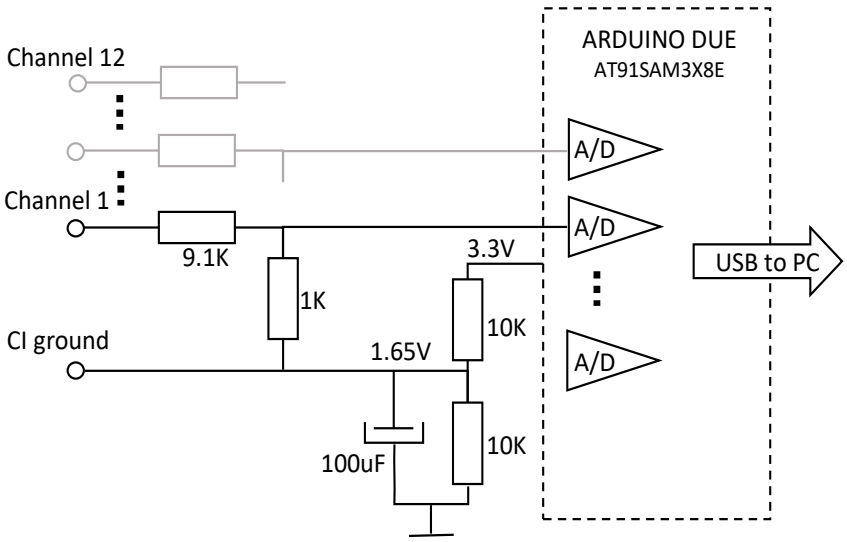


Figure 104: A/D converter array with level shifting

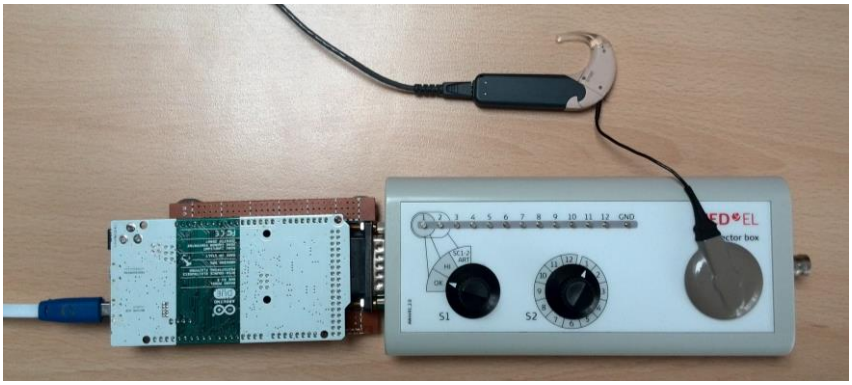


Figure 105: Photo of the acquisition system connected to a cochlear implant processor

The simple circuit shown in Figure 104 achieves to overcome the limitations of the A/D converter and to create a real-time acquisition module that can measure in *parallel 12 signals* with an *amplitude between -16V to + 16V* with a sampling rate up to *83 kSPS/channel*, *or up to 1 MSPS in single channel configuration*.

In Figure 106 and Figure 107, the registered waveforms are observed by using 12 channels, and respectively 1 channel configuration.. The first 4 channels in Figure 106 are different in shape and stimulation pattern because the cochlear implant processor used in the experiment *applies different coding strategy for the first 4 electrodes* responsible of delivering information *at low frequencies* (Wouters, et al., 2015) (Harczos, et al., 2013).

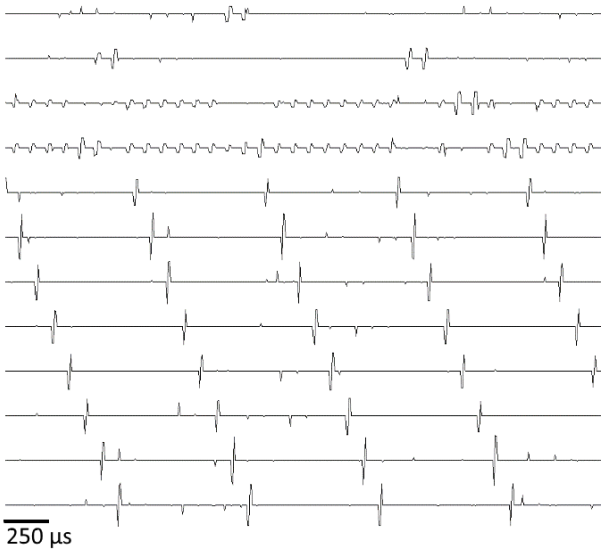


Figure 106: Simultaneous recordings of CI pulses for 12 channels (top to bottom: low to high frequencies) with 83 KSPS / Channel

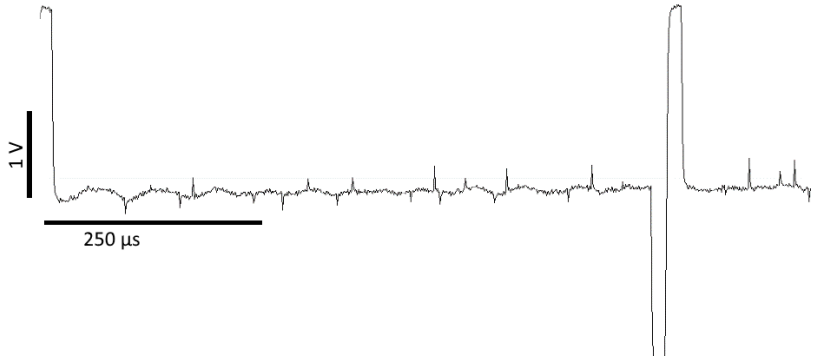


Figure 107: Single channel recording with 1MSPS

5.5.4 Stimulation Level Tool

Electric pulses are generated by mapping *sound levels* of different frequency-channels to *stimulation levels* (electric pulse intensities) according to the patient specific fitting parameters provided by the fitting specialist (Caner, et al., 2007).

During the *fitting procedure*, first the cochlear implant system measures the impedances inside the cochlea between each electrode and the implant ground. After that, for each channel, the fitting specialist sets the minimum (THR) and maximum (MCL) stimulation levels of the electric charge to be dispensed at each pulse. Using the measured impedances and the provided stimulation levels, the MAESTRO fitting software computes the *pulse durations* to be used, and the cochlear implant processor determines the *pulse amplitudes* corresponding to the sound levels (Wouters, et al., 2015).

An algorithm and a software application are proposed and implemented in this section assuring the following actions:

i) takes the fitting stimulation levels (THR and MCL) of the cochlear implant processor,
ii) takes the measured impedances,
iii) takes the utilized pulse durations,
iv) translates the measured pulse amplitudes to the stimulation level expressed in percentages (THR-0%, MCL-100%) by the maplaw compression curve,
v) displays in real-time the *stimulation levels* and the *sound levels* for each individual channel.

Such a visualization can be used for cochlear implant demonstrations and also for fault detections of the: microphone, automatic gain control and fitting. Moreover, the stimulation levels can be reverted to sound levels using the maplaw inverse.

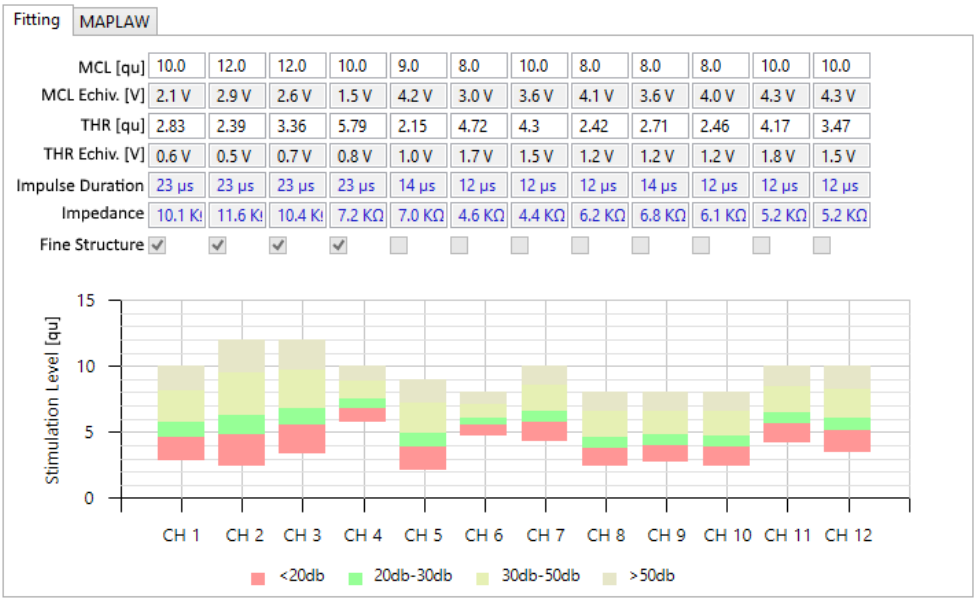


Figure 108: Stimulation level configuration tool for 12 channels

Figure 108 shows the *stimulation level configuration view for 12 channels*. It is possible to import the fitting parameters from the data base of MED-EL Maestro System Software, or it can be configured manually. Also, the *maplaw* compression curve can be set as displayed in Figure 109. This curve is also used in automatic gain control (MED-EL(1), n.d.).

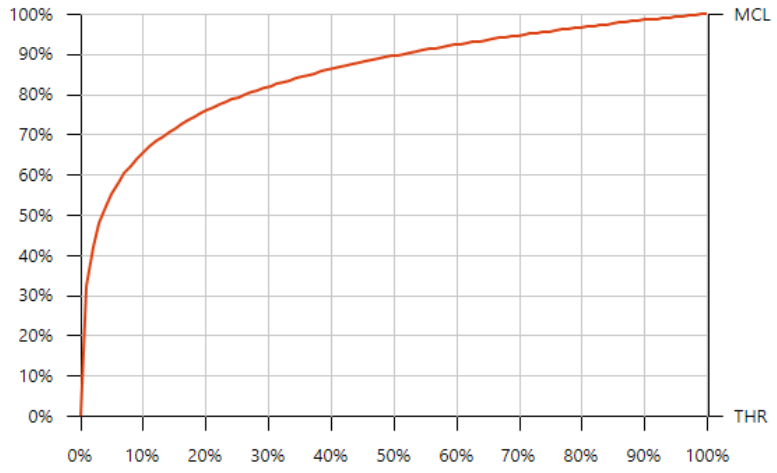


Figure 109: Maplaw compression curve used to map sound levels (x-axis) to stimulation levels (y-axis)

5.5.5 Basic Real-Time Auralization Method

The previous works (Zeng, 2008) (Kuczapski & Andreescu, 2016) (Harczos, et al., 2013) show that the auralization of the sound perceived by cochlear implant users can be an important research and development tool.

One purpose of the proposed real-time data acquisition system is to provide base for the realization of a self-learning adaptive auralization system that will be presented in an upcoming paper.

As proof of concept, a real-time auralization system is implemented based on the *simple auralization method* that achieves the following actions:

i) *collects the generated pulses* for each channel from the cochlear implant processor using the I^{100} detector box and the developed real-time data acquisition system,
ii) continuously approximates the *envelope of each channel* by computing the maximum pulse amplitude of the last 5 ms,
iii) *maps the channel amplitudes to the stimulation levels* using the known THR and MCL levels,
iv) transforms the *stimulation levels into sound levels* using the inverse of the *maplaw* compression curve,
v) ensures *amplitude modulation* with corresponding sound level of *sine wave generators having the central frequencies associated to each channel*; the sine wave amplitude before the modulation is inverse proportional to the frequency,
vi) *mixes all* modulated sine waves for auralized sound.

The implemented auralization method is shown in Figure 110.

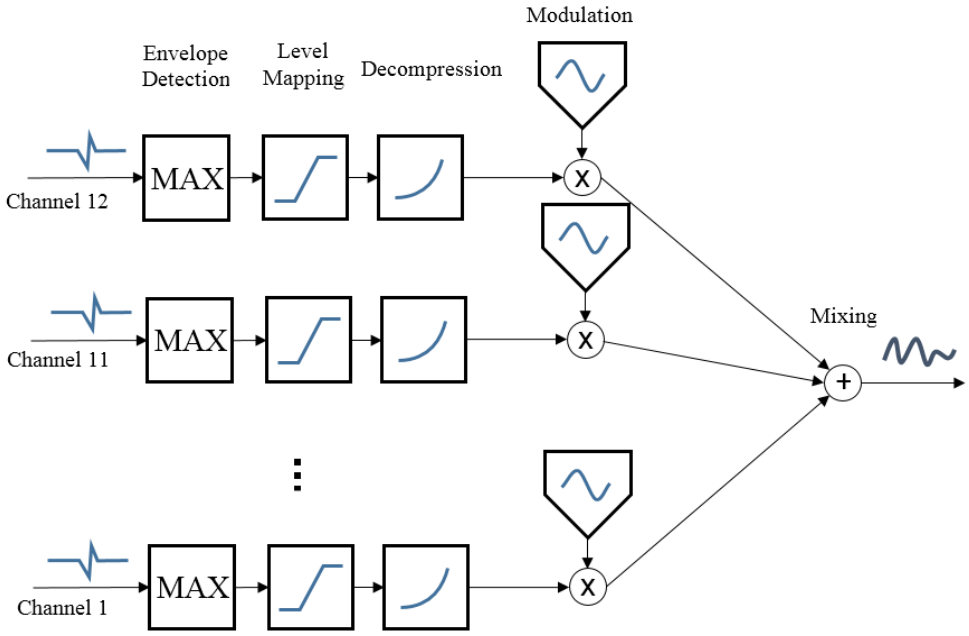


Figure 110: Schematics of the simple real-time auralization method implementation

Some experiments have been conducted to determine the quality and intelligibly of the auralized sound. The conclusions are: a) in the speech case, after a few seconds, the *auralized sound can be understood*, b) however, *music is not recognizable* and it is perceived mostly as noise.

As the simplified auralization method relies only on local cues, ignoring any temporal cue, these observations are consistent with the known literature (Moctezuma & Tu, 2011) (Zeng, 2008) (Wouters, et al., 2015).

Figure 111 and Figure 112 display the waveform and the spectrogram of a test sentence in the original form (top) and after auralization (bottom). In the waveform, it can be observed that the *envelope and dynamics of the sound are mostly preserved*, making the result understandable as speech. In the spectrogram representation it becomes clear that the *fine variation of the frequency components is lost*.

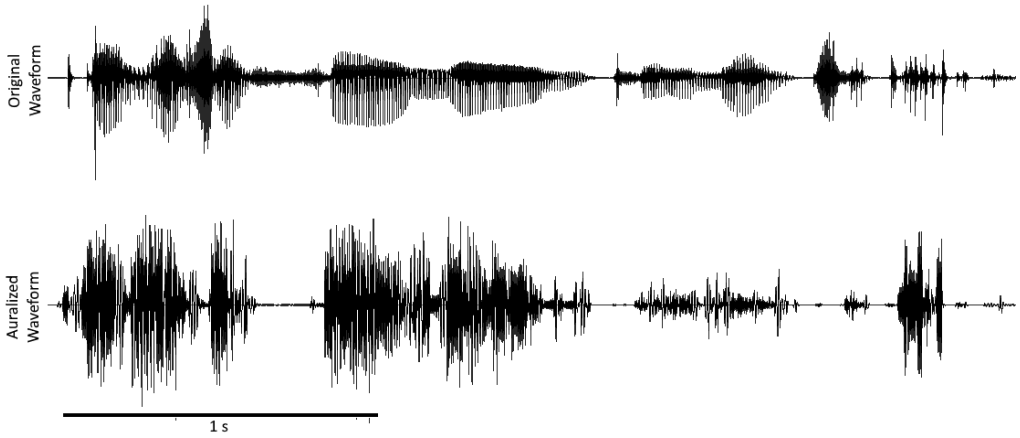


Figure 111: Envelope comparison of original sound (top) vs auralized sound (bottom)

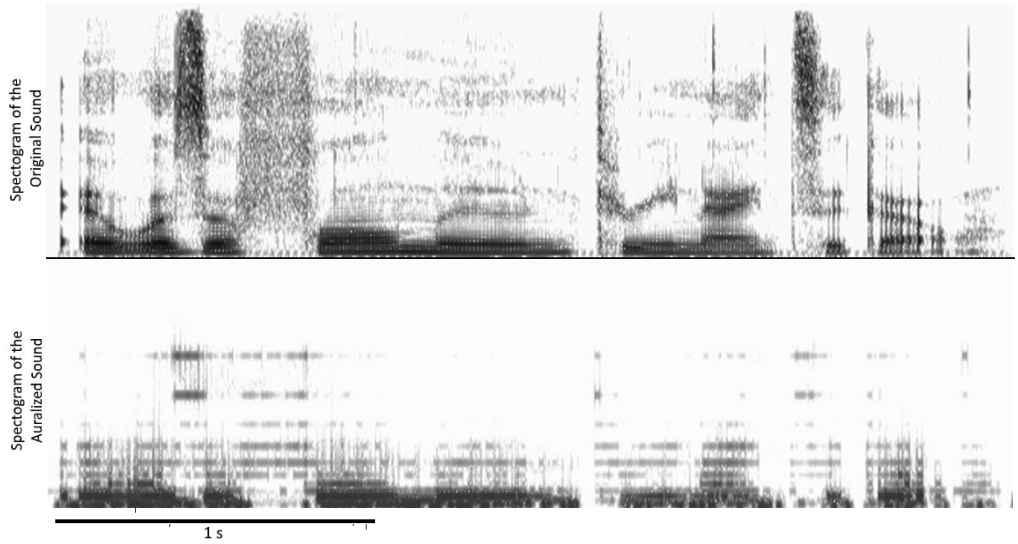


Figure 112: Spectrogram comparison of original sound (top) vs auralized sound (bottom)

5.5.6 Conclusions

Cochlear implants (CI) are surgically implanted electronic devices that provide a sense of sound to persons with severe to profound hearing loss. Hearing is restored by creating hearing sensations through direct electric stimulation of the auditory nerves.

In order to create meaningful hearing sensations, a cochlear implant fitting specialist has to set appropriate stimulation levels. As the hearing sensations are subjective and difficult to transmit, most fitting specialists never actually experienced on how changes in the fitting parameters are translated into perceived sounds. Moreover, the processor fault can be often overlooked and therefore, the degraded hearing quality is erroneously attributed to changes in the patient physiology. In this context, the contributions and conclusions of this section are as following:

1) A real-time data acquisition system is developed to be connected to a cochlear implant processor, in order to register and display the generated pulses on a computer display. The sampling rate is up to 83 kSPS/channel for 12 analog channels, or up to 1 MSPS in single channel configuration. The proposed system is developed for MED-EL Opus 1, Opus 2 and Sonnet cochlear implant processors with the following main parts: i) I^{100} Detector Box that transduces the information received from the CI processor antenna into electric pulses, similar to an implanted receiver/stimulator; ii) 12 channel real-time analog data acquisition module developed using Arduino Due and a custom developed shield; iii) PC connected through USB to the data acquisition module, running a Java software developed for monitoring, real-time visualization and auralization.

2) A useful assisted software tool is introduced that displays in real-time the *stimulation levels* and *sound levels,* and also the *main parameters and variables for each cochlear implant channel*. Such visualization can be used for cochlear implant specialists to optimize the fitting procedures, for demonstrations, and for fault detections of: microphone, automatic gain control, and fitting procedure.

3) A simple real-time auralization system is implemented. It approximates and *replays the perceived sound by using the registered pulses*. This is an important tool in the research and development area of cochlear implants, for *comparing the quality of the cochlear implant systems* delivered on the marked, and for *enhancing specific algorithms for CI processors*.

5.6 Intra-Cochlear Current Flow Model

In order to better estimate current spread inside the cochlea, a *3D current spread simulation model was developed*. High resolution cochlea models are not publicly available therefore, we have developed a computer program to generate an approximated *voxel model of the cochlea*. In such a model, the three-dimensional space is divided in a fine number of cubes of identical sizes called voxels (e.g. volume-pixel). Each cube represents a homogeneous space with various characteristics like color, material type, etc. Afterwards, using the generated voxel model of the cochlea, we have *developed an algorithm and implemented a software application to numerically approximate the current flows and potentials within each voxel*.

During the development of the simulation algorithm, we made the following assumptions:

- Cochlear tissues are bad isolators therefore inductances and capacities can be ignored;
- *Stimulation electrodes are modeled as current sources*, each electrode injecting current into the center of one or more voxels;
- *Ground electrode is modeled as voltage source*, creating a set of voxels with fixed potential set to 0V.

The above assumptions allowed us to define a reduced set of parameters and variables of the voxels:

$$\rho_i = resistivity\ of\ voxel\ i$$
$$U_i = Potetial\ of\ voxel\ i$$
$$IS_i = Current\ injected\ in\ voxel\ i$$
$$<?>_{i,j} = a\ variable\ around\ voxel\ i.\ (j = 1..6)$$
$$l = voxel\ size$$

The resistivity of the voxel (ρ_i) is determined by the simulated material type (e.g., bone, tissue, liquid, etc.), while the potential (U_i) and injected current (IS_i) of the voxel is calculated by the simulation algorithm. There are two types of voxels:

- *Voxels with fixed injected current* – in these voxels the divergence of the electric current densities (the sum of electric currents through all 6 faces of the voxel) is set to be a fixed value IS_i. Typically, IS_i is set to 0 A for most of the voxels, respecting Kirchhoff first law of current conservation. Non-zero IS_i is used to represent electrodes injecting current into the model.
- *Voxels with fixed potentials* – these types of voxels are used to create areas of fixed potentials like the area representing the ground electrode, which is set to 0 V. At least one voxel with fixed potential must be created to be used as reference potential by the simulation algorithm.

With these variables, it was possible to transform the current flow simulation problem into a problem of solving an equivalent network of electric resistances, current sources and voltage sources. The method to map the voxels to an electric network is depicted in Figure 113. We have assigned a circuit node to the center of each voxel and connected the nodes of neighbor voxels with a resistance ($R_{i,j}$). Considering that the voxels are perfect cubes (width, height and depth are equal to l) and applying the Pouillet law, we deduced the following formula to calculate $R_{i,j}$:

Resistance between center of voxel *i* and its neighbor *j* is:

$$R_{i,j} = \frac{\rho_i + \rho_j}{2l} \quad (12)$$

In addition to the resistances, based on the type of the voxel, a voltage or a current generator is connected to the center node.

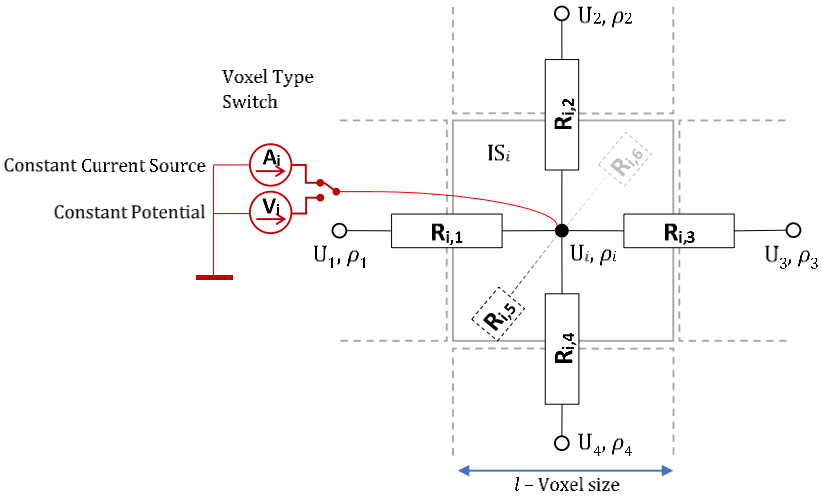


Figure 113: Electric model of a voxel with two possible working mode: a) Constant Injected Current; b) Constant Potential

Within this configuration, given that the potentials of the surrounding nodes/voxels are fixed, using Ohm and Kirchhoff laws and the superposition principle, the potential or the injected current is calculated, depending on the voxel type:

Equation for node with constant injection current (IS):

$$U_i = \sum_{j=1}^{6} U_{i,j} \frac{{}^{1}/_{R_{i,j}}}{\sum_{q=1}^{6} {}^{1}/_{R_{i,q}}} \quad + \quad IS_i * \frac{1}{\sum_{j=1}^{6} {}^{1}/_{R_{i,j}}} \tag{13}$$

Equation for node with constant potential (U):

$$IS_i = \left(U_i - \sum_{j=1}^{6} U_{i,j} \frac{{}^{1}/_{R_{i,j}}}{\sum_{q=1}^{6} {}^{1}/_{R_{i,q}}} \right) \sum_{j=1}^{6} {}^{1}/_{R_{i,j}} \tag{14}$$

The electrical model created using the described method translates into a *linear equation system*. Unfortunately, in order to have meaningful simulation results, we have observed that the cochlea model must be divided at least in a 100x100x100 matrix of voxels, resulting in *1 million equations*. This amount of equations cannot be solved accurately and efficiently with typically available Linear Programing software. Therefore, we had to *develop an iterative numerical solving algorithm*. The developed simulation algorithm relies on the electric equilibrium formulas described above, and has the following steps:

i) The *simulation is started* with all voxel potentials (U_i) and injected currents (IS_i) set to 0 or to the constant values specified based on the model configuration. By setting to 0 the U_i and the IS_i variables, the equations are validated in most of the voxels, excepting the voxel with configured current injection, the voxels with non-zero fixed potential and its neighbors.

ii) In order to satisfy the equations for all voxel, the algorithm starts a *sequence of iterations*, where in each iteration for each voxel, the *potential* (U_i) *or the injected current* (IS_i) *is calculated and replaced* into the voxel using the calculation formulas depending on the voxel type. Similarly, to a graphical 2D fade algorithm, from iteration to iteration, *the potentials will spread around the voxel* marked as current and voltage sources, converging to a stable configuration which validates the equations in all voxels. Once the iterations are reaching a convergent state, the algorithm can be stopped and the potentials of each voxel is calculated.

iii) After the potentials of each voxel are calculated, the *current through the resistances* is easily calculated using Ohm law. In order to determine the current density in a vector form for each voxel, we consider the currents through those 3 resistances which are pointing in the directions of the X, Y and Z axis.

Once we have implemented the *algorithm to calculate potentials and current flow densities within any arbitrary 3D model* described in the form of vertexes, we have *created a cochlea model* including implanted electrode array, using structure, geometry and resistivities found in the literature (Hanekom, 2002).

Material / Tissue type	Resistivity (Ω·mm)
Silicon	10^{10}
Electrode	1
Perilymph	700
Endolymph	600
Bone	6410
Basilar Membrane	4000
Mixed Tissue	2000

Figure 114: Resistivity values used for the cochlea model

Parameter	Value
Number of turns - Cochlea	2.5
Cochlea Basal Diameter	10 mm
Cochlea Apical Diameter	1 mm
Cochlea Height	4.5 mm
Cochlear Duct Diameter – Basal	2.5 mm
Cochlear Duct Diameter - Apical	1.5 mm
Number of Electrodes	12
Number of turns – Electrode array	1.7
Electrode Diameter - Basal	0.8 mm
Electrode Diameter - Apical	0.5 mm

Figure 115: Dimensions used to generate the cochlea model, including the electrode-array

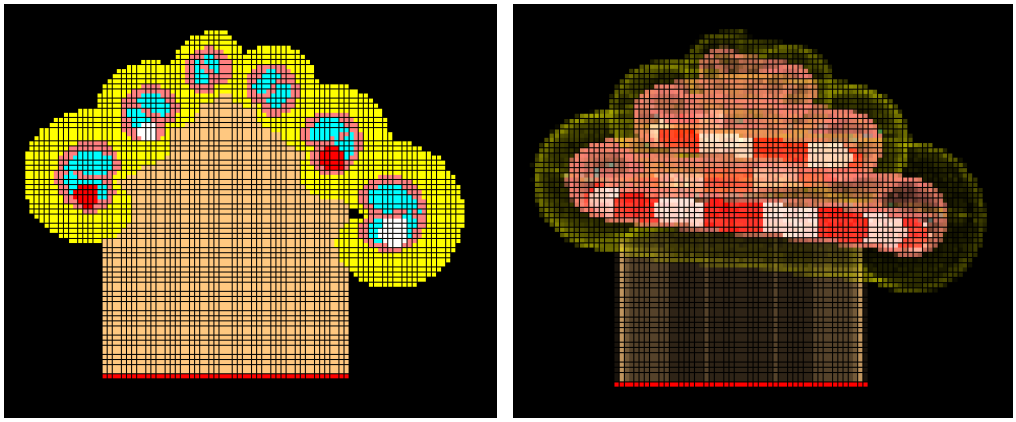


Figure 116: Programmatically generated 3D Voxel model of the cochlea representing body surfaces and current density. Vertical cross section – left; Transparent surface view – right; Model size - 100x100x100 Voxels

In Figure 116, the vertical cross section (left) and the transparent 3D model (tight) of the generated cochlea model can be seen. Each type of material or tissue is colored with a specific color:

- yellow – bone;
- orange – mixed tissue;
- pink – basilar membrane and cochlear duct walls;
- blue – perilymph;
- white – silicon (insulator);
- red – electrode (metal).

In the *vertical cross section image* of the cochlea model, the structure of the cochlear duct is clearly visible. The Cochlear duct is filled with the perilymph liquid, and it is divided by the basilar membrane. The electrode array is positioned in the lower part of the cochlear duct (i.e. Scala Tympani). On the left side, in the *transparent 3D rendering* of the cochlea model, the positioning of the electrode array is observable. The bottom part of the cochlea model was configured as metal (red) and it is used as the common ground point for the electrodes and these voxels must be configured with fixed potential $U_i = 0$.

In order to create a simulation of the current flows through the cochlea, for a given configuration of stimulation intensities, we have to configure for each electrode a *single voxel from the cochlea model with* IS_i *set* to the current provided by the electrode configuration. After the *initial configuration*, the simulation algorithm is executed as described above. The execution of such a simulation can take several hours, therefore it would be impractical to simulate the current spread in feasible time when evaluating the nerve stimulation of a stimulation sequence.

Due to the linear nature of the generated electric circuit network, we managed to *drastically reduce the time needed to generate current spread simulation* for any arbitrary configuration of electrode current intensities, *by initially generating the*

solutions for all configurations where only a single electrode is activated with exactly 1 Ampere. Once we generated these solutions, the solution of any arbitrary configuration of electrode current intensities can be easily calculated by applying the *principle of superposition* and combining the partial results weighted width the individual current levels.

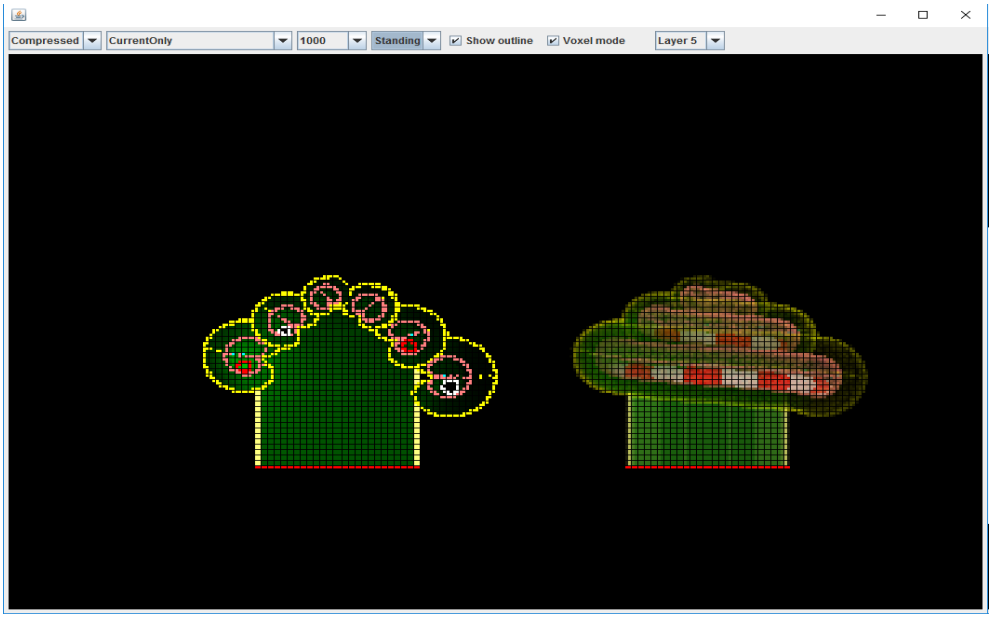


Figure 117: The application running the simulation, displaying the current density field over the 2D and 3D rendering of the cochlea model

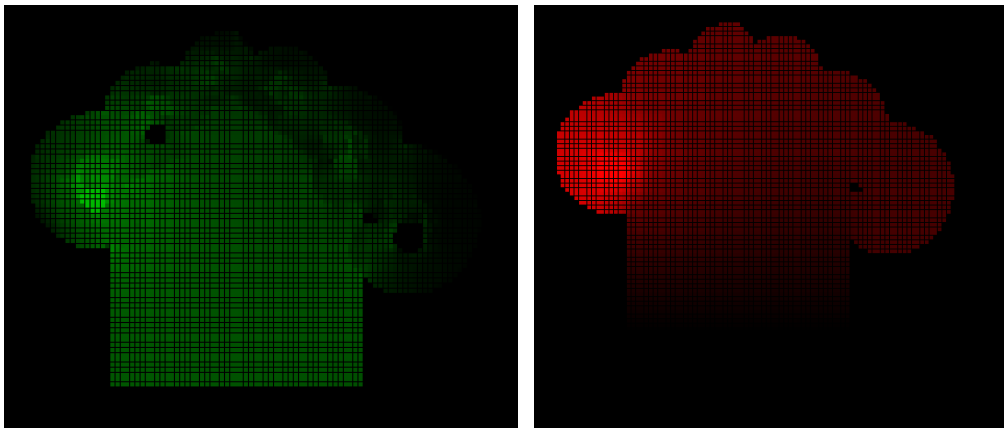


Figure 118: Vertical Cross Section of Simulation result for electrode #5: Current density field using logarithmic color scale (left); Electric field using linear color scheme (right)

Figure 117 shows the user interface developed to display the generated cochlea model and the progress of the simulation. The cochlea model is rendered in two ways: vertical 2D cross section and 3D transparent body. In both cases, only the outlines of the cochlea parts are displayed, the missing parts being filled with the current density field represented in green.

The result of the simulation for electrode #5 is depicted in Figure 118. The current density field (left) is shown with green color using a logarithmic scale to adjust color intensity based on the current density. The location of the electrode injecting current is clearly visible at the left part of the image, also black gaps can be observed marking the positions of the electrode array insulator. On the right, the electric potential field is represented with red, using linear color intensity mapping.

In order to better represent the result of the simulations, we have placed 1000 measurement points (probes) into the model, alongside of the base of the basilar membrane, the place where typically the nerves are stimulated. In Figure 119 and Figure 120 the *current density* and the *potential of each measurement point* is displayed.

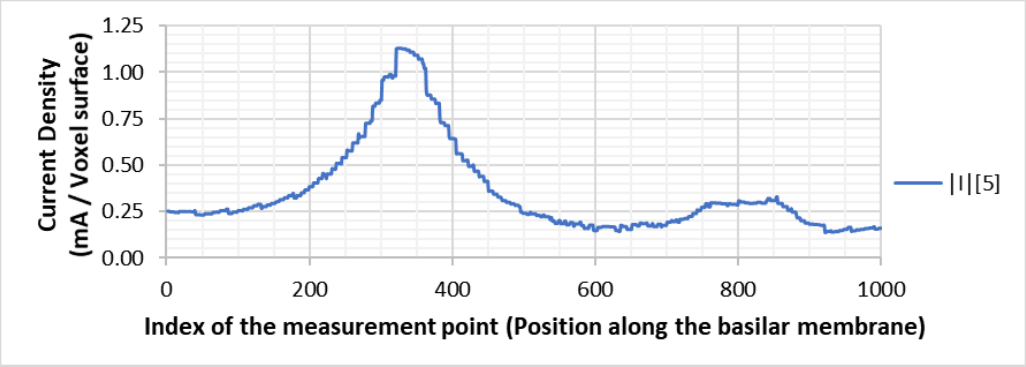


Figure 119: Curve showing the simulated current flow density alongside the Basilar Membrane generated by the 5th electrode injecting 1 A.

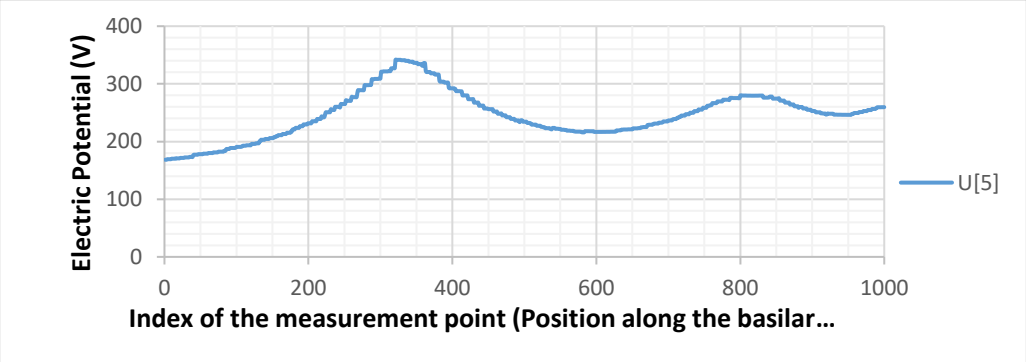


Figure 120: Curve showing the simulated potential alongside the Basilar Membrane generated by the 5th electrode injecting 1 A.

6 CONCLUSIONS

We have started the present research project with an urge to understand how patients, especially children, experience hearing through cochlear implants. Previously developed auralization models that simulate how sounds are perceived by patients were quite simple approaches. Although they are capable of giving a hint regarding the nature of the perceived sounds, the results were not dependent on the cochlear implant stimulation strategy, and have ignored the brain capability to adapt.

We observed that one critical condition to call an *auralization successful* is that a cochlear implant user, whose cochlear implant model was used to generate the auralization, *should not be able to observe the difference between the original and the synthesized sound (validation criterion)*. This condition was not met by the existing auralization approaches making it insufficient, especially for speech-therapists and the patient's family.

Another troubling issue within the cochlear implant user community is the wide variety in the efficacy in hearing restoration by cochlear implants. Even though similar cochlear implants are used with similar stimulation strategies, the *level at which hearing performance is restored is very specific to each individual,* and many times seams arbitrary. This puzzled both clinicians and parents leaving them almost helpless in the fight for quick habilitation of children born deaf, where the time is critical.

The first thing, we could think of to ease the situation, was the *development of a novel auralization method* with the *following characteristics*:

1. The auralization should relay on the actual *stimulation strategy* and *settings of a cochlear implant model.*
2. It must account for the *learning and adaptation capability of the brain.*
3. The auralization methods itself must be *completely agnostic to the cochlear implant stimulation strategy*, instead it should learn to interpret the signals.
4. The auralization method should *receive as input only the electric signals coming from the cochlear implant*, yet it must generate the *output sound in such a way that this sound is indistinguishable from the original sound* for a cochlear implant user.

Such an auralization method is valuable both for researchers - *developing new stimulation strategies,* and for *therapist* - working on (re)habilitating patients.

In the case of *researchers and cochlear implant developers*, the novel auralization method provides the *following advantages*:

- Objectively compare the expected hearing quality with existing sound coding strategies.
- Test bench for development of new coding strategies.
- Improve fitting procedures.

In the case of *speech- and other habilitation therapists*, the auralization tool shed *new light on the way by which patients are perceiving the exercises and habilitation programs*. Many times, the same therapies used in the case of hearing aid recipients are applied to cochlear implant users. Even worse, in the case of associated cognitive deficiencies, the therapists involved have no or limited knowledge on how to talk to these patients. This situation often leads to practices which are common sense in typical children, but are *very inefficient or maybe even harmful* in the case of cochlear implant users, for example:

- Talk lauder when it seems that the patient does not hear the therapists.
- Involve music or songs in order to help memorizing the words.
- Playing on instruments while singing.
- Thinking that if the cochlear implant is not in place, the child still hears something useful.
- Relying on the sound directionality while playing (for example hide and seek)
- etc...

Many therapists were shocked when we have shown the auralized sounds, and were even more puzzled when we have shown that those *"ugly robotic" sounds are heard as normal by the patients*. They have recognized that they had no knowledge about the nature of sound perceived by the patients, and that they think that some exercises they have applied before must be adapted.

Besides talking to professionals and parents, during the research and development of the current thesis, we have met hundreds of children with cochlear implants and we have monitored the evolution of them. Sometimes we were glad to see great evolution from one fitting to the other, other time we were completely puzzled by lack of it or even regress.

When *regress is observed*, the biggest question is what went wrong. Is it a failure of the processor? or even worse a failure of the implant? Is it a consequence of the previous fitting? or is it progress of the processes which provoked the deafness in the first place? In the case of advanced - verbal patients, many times they can give hints when and what went wrong, and they can help the fitting process. But when small children are fitted, they do not give any feedback on changes on hearing quality or defects of the implant.

Some defects can be easily found by the fitting equipment, like the interruption of an electrode, or changes in the intra-cochlear impedances, but *many other defects are hard to find*, like dirty microphones, defective or inoperative AGC stage. Also, the current fitting tools do *not monitor the evolution over time of the fitting parameters*, making it hard to observe slowly progressing fitting problems or degenerative processes. Facing these problems, *dedicated tools and methods were conceived and implemented* beside the development of the auralization method.

The structure and the contributions of the present thesis will be briefly presented in the following.

6.1 Contributions – Overviews and State-of-the-Art

Being an *interdisciplinary research project*, the present thesis must provide the *state-of-the-art and overview* of all relevant areas involved in the project, thus the following *main elements and conclusions* are synthesized:

i) Introduction to the ear anatomy and causes of neuro-sensory deafness (Chapter 2.1):

- Short description of the *ear anatomy*, describing the organs and parts involved in the propagation of the sound wave from outside to the inner ear.

- Detailed description of the *inner ear structure, function and mechanisms*.

- Examples of in vivo and in vitro measurements at different stimulation signals.

ii) Overview of the existing cochlear implant systems (Chapter 2.2-2.4):

- Description of the operation principles of the cochlear implants including presentation of its components.

- *Review and comparison of existing stimulation strategies,* available cochlear implant brands, and *current challenges in cochlear implant technologies*.

iii) Introduction of the existing ear mathematical models (Chapter 3):

- Presentation of *mathematical models of the ear most appropriated to the thesis* research objectives.

- Complete ear simulation models are quite complex, however the simulation of hearing through cochlear implant can be achieved with *simplified models because the cochlear implant bypasses most structures of the ear*.

- *Proposed simplified inner ear (cochlea) model* for initial experimenting of the auralization model.

iv) Overview of the Maestro cochlear implant system (Chapter 5.1):

- Overview of MED-EL Cochlear Implant System, describing its characteristics and the most important fitting parameters.

v) Review of existing auralization methods (Chapter 4.2):

- Short *review of two auralization methods* found in literature.

- Actual auralization methods rely on simple methods and do not approximate well the hearing experience of cochlear implant users. These auralization methods do not have the capability to simulate the brain ability to adapt to new stimulation patterns.

- *Concept presentation of a new auralization method involving machine learning methods*.

6.2 Contributions - Original Methods, Models and Results

Relevant original contributions are described in the following:

i) New auralization method using autocorrelation based pattern recognition (Chapter 4.3):

The *main research objective* of the thesis is to *develop a novel auralization algorithm capable to synthetize the sounds perceived by cochlear implant users* (pending patent). The implementation details, simulation results and experiments are detailed in chapter 4.

Main original contributions of the proposed auralization method are as following:

- ***New method is proposed for detecting spectral components of perceived sound from the cochlear nerve firing patterns***, i.e., *frequency detectors*, that mainly contains:
  a) a *self-learning algorithm* to determine the *frequency specific autocorrelation masks* for each desired spectral frequency component, and
  b) an *amplitude indicator* of each spectral component (amplitude frequency response characteristic) based on the *cross-correlation of the cochlear nerve firing autocorrelation pattern of an unknown sound, with the frequency specific autocorrelation masks*.

- For the *natural hearing*, the *frequency response of the frequency detectors* at low and high audio frequencies *are in accordance with the real hearing pitch perception* that *validates the proposed method*.

- The proposed method is *experimented using cochlear nerve firing patterns* from:
  a) a *cochlear implant with simulated ACE strategy*, and
  b) a *MED-EL Opus 2 CI processor using FS4P strategy.*
  Comparative results regarding frequency response versus the natural hearing are given for both a) and b) cases.

- The *proposed method is successfully used* inside the *new developed auralization technique for sound synthetization based on the cochlear nerve firing patterns*. *Both original and synthesized sounds were experimentally replayed to CI users*, and in most cases, the *synthesized sound was perceived as good or better* compared to the original sound, that is a *validation of the proposed auralization method*.

The *novelty of the developed auralization method* is highlighted by the following characteristics:

- The auralized sound is generated using the *electrical signals coming from a cochlear implant* or a cochlear implant model, in contrast with existing auralization methods, where the information loss caused by the stimulation strategy, is estimated using simple signal filtering elements.
- The auralization method is *completely agnostic at the stimulation strategy* used by the cochlear implant, being capably to learn to interpret any type of stimulation strategy in cochlear implants.

- Instead of typical signal processing approaches (e.g., filtering techniques), the developed auralization method *transposes the auralization problem into image processing and pattern matching terms*.
- Beside synthetizing the perceived sound during the *learning process*, *comprehensive quantitative frequency discrimination characteristics are generated*, without using analytic filters, making it possible to *objectively compare different stimulation strategies*.

In order to validate the developed auralization model, experiments with real patients are executed and discussed (Chapter 4.4).

ii) Frequency responses of the frequency detectors using stimulation through cochlear implants (Chapter 4.3.4), with the original contributions:

- The *frequency response of the frequency detectors, experienced by the CI users, are estimated and compared*, employing the cochlear nerve firing patterns stimulated through CI, using the developed auralization method with autocorrelation based pattern recognition. *Two well-known CI stimulation strategies are used*: simulation model for ACE strategy, and F4P strategy with hardware interfacing with a real CI processor.
- *New capabilities of the developed auralization method* are pointed out, comparing the experimental results with the known characteristics of the analyzed stimulation strategies:
  - Ability *to learn and adapt to unknown stimulation strategies.*
  - Ability to estimate *frequency discrimination characteristics with high accuracy*.

iii) Validation criterion of the auralization model - Cochlear implant user feedbacks (Chapter 4.4), with the following original contributions:

- *Introduction of a new universal validation criterion for any auralization method:* The *auralization synthesized sound should indistinguishable from the original sound for a cochlear implant user*.
- In some cases, the results of the auralization sound are perceived even better by the patients, opening the possibilities of *new pre-processing algorithms to improve hearing* performance.
- *Specific experiments with cochlear implant users to test and validate the developed auralization method*.

iv) Statistics of typical stimulation levels (Chapter 5.2):

Trying to understand the differences between CI patients, and looking for possible explanations in their differences, the *evolution of fitting and measurement data of 150 pediatric patients is collected and analyzed*. The *provided statistics* can be used by fitting specialist and medics to compare fitting levels of individual patients against typical fitting levels allowing to *more easily assess the evolution of the individual patient*.

v) Computer assisted fitting of cochlear implants – Tracking the effective stimulation threshold (Chapter 5.3), with the following *main contributions:*

- Review of current fitting procedures.
- *Introduction of new calculated fitting monitoring parameter - Effective Stimulation Threshold (EST)* (11), which represents the stimulation level that causes conscious hearing sensations to the patient. The EST is determined from open air tonal audiograms correlated with the THR, MCL and maplaw settings.
- *Assistive software for cochlear implant fitting* developed to help fitting specialists to improve fittings: *prediction of new threshold hearing levels*, and *long-term monitoring of the evolution of EST levels*. This software can directly use data exported from MED-EL fitting system making it immediately usable.
- Exemplification of the *benefits of EST evolution monitoring* *in early identification of CI technical defects or pathological processes*.

vi) Case studies of fitting evolution (Chapter 5.4):

- *Evolutions of 4 patients as case studies are presented and discussed using EST levels* and audiograms: 2 cases with typical good evolution, and 2 cases with sudden changes in the EST levels indicating defect of CI processor or degenerative processes.
- Monitoring of EST levels help the fitting specialist *to fine tune for individual fitting* and *easily detect progressive or acute degradation of the hearing performances*.

vii) Interfacing with Med-El cochlear implant processors (Chapter 5.5):

A *cheap and compact interfacing between CI processor and PC was developed*, to allow practical experiments, analysis of stimulation strategies, and easy troubleshooting of defective CI processors. It contains the following parts:

- *MED-EL I100 Detector box* used to translate the induced signals coming from the CI processor into 12 electric signals replicating the exact shape and amplitudes of the stimulation impulses delivered by the cochlear implant.
- *Arduino Due microcontroller* board with 12 high sampling rate ADC.
- *Own developed passive circuit adapter* to shift, divide and filter the signals between I100 Detector box and ADCs.
- *Own developed microcontroller program to capture and transfer high speed measurement data* (1 MSPS – i.e., 83 kSPS/channel for monitoring 12 channels) from the Arduino Due board to a PC through USB connection at *1.5 MByte/sec.*
- *Own developed Java application for real-time signal visualization, signal recording* and for controlling the microcontroller program parameters.

viii) Intra-cochlear current flow model (Chapter 5.6):

In preparing to future work, an *electric intra-cochlear current flow 3D model is developed* using own-developed Java library for electric distributed voxel model of the cochlea and current density field modelling generated by *current-injection of active electrodes*. The electric field along the basilar membrane *with electrode interactions* is estimated for *nerve stimulations*.

The *intra-cochlear current flow model* consists of the following own developed components and methods:

- *Computer algorithm to generate 3D cochlea shapes using 3D voxels* based on typical dimensions and conductivity values of the human cochlea.
- *Computer algorithm to generate 3D model of the cochlear implant electrode array.*
- *Efficient computation algorithm to estimate the 3D intra-cochlear current flow using voxels* with defined conductivity and specific injected currents by electrodes. This algorithm employs pre-calculated electric field along basilar membrane generated by each individual electrode normalized to 1A, and then determines by superposition the electric field for arbitrary weighted configuration of electrodes.

The advantages of the developed method is consisting in:
- More efficient current flow algorithm with configurable spatial resolution.
- Fast convergence using superposition of pre-calculated electric fields for each electrode.

6.3 Contributions - Developed Software Modules

All methods, algorithms and experiments were implemented and executed using Java 1.8 programming language. The main *developed Java libraries* during the research project are the following:

Generic signal processing framework in Java: Generic Java library designed to allow structured implementation of signal processing schematics, having data streaming and real-time processing in mind. It provides *generic classes and interfaces* for signal processing in time-domain and also a set of filter implementations including, parametrized FIR and IIR filters, and Gamma-Tone filter.

The developed signal processing framework is *characterized* by:
- Ability to be used in *real-time signal streaming application*.
- Ability to be used in *buffered/non-real-time processing*.
- Supports all *OOP concepts* to allow the development of *well-structured and reusable* signal processing models.

Adaptation of Zhang auditory model implementations from C++ to Java: Implementation details and C++ source code is provided in the literature for the *auditory simulation models* (Chen & Zhang, 2006). The *inner hair cell and the synapse models*, used in the experiments within the thesis, were Java adaptations of these models and algorithms.

Simplified ear model: Java implementation of the *proposed simplified ear model* using the own-developed signal processing framework and the Java adaptation of Zhang auditory models.

Image processing library: *Simple image processing library implementing basic operations used in the experiments*: *Image autocorrelation*, normalization, blur, cut-off above or below limits.

Library for electric and current density field modelling in non-capacitive medium: Own developed library using *finite differences modelling approach, developed to estimate of electric fields in any 3D* object presented as a *voxel model*.

Cochlear model generation method: Java method to generate a 3D voxel model of a human cochlea, using parametric curves and geometric shape.

6.4 Future Work

During the research and development of the subjects included in the present thesis, many new problems and opportunities were discovered. Some of these are related to the day-to-day problems faced by cochlear implant users, other are meant to improve current cochlear implant technologies. In the following, the topics planned to be addressed in the feature are presented.

Complete auto-diagnosis box for MED-EL Cochlear implant processors

Partial hardware failure of the cochlear implant processor is a common issue among cochlear implant users, and unfortunately it too often goes unnoticed for too long periods. This is a serious issue for pediatric patients, and for many defects there is no current approach to detect them.

Using the interfacing device developed and presented in the present thesis, we intend to develop a *portable device capable of registering a reference catalogue of test sounds and stimulation impulses specific to a certain fitting configuration*, which can be used *to identify any changes in the cochlear implant processor responses*, thus highlighting *any possible defect*. This complete auto-diagnosis box for CI processors contains: an isolated sound chamber, high quality speakers, the I100 Detector Box and the developed interface, all connected to a PC or tablet.

Realtime pre-processing of sounds

During the auralization experiments, patients using a certain type of cochlear implant have continuously reported the *synthesized sound as being more enjoyable and clearer than the original, thus CI processors should pre-process the sound as auralization sound.* This effect might be caused by the simplification of the frequency spectrum considering the actual auditory masking characteristics of the cochlear implant. Although the actual proposed auralization process is complex and computationally heavy, we think that it is possible to obtain similar effect by applying the *cochlear implant specific auditory masking characteristics* in real-time using less resource intensive algorithms.

Vocoder based on neural networks

One key piece for a successful auralization is the sound synthesis part - vocoder, where the detected spectral components are transposed in sounds. This is also one of the most computationally intensive components in the current implementation. This computational complexity prohibits the utilization of such component in real-time

applications, although it could be a key component in audio enhancement applications for cochlear implant users.

With the *advance of deep-learning technologies* (He, et al., 2019) (Cui, et al., 2018) and the appearance of *hardware accelerated neural network*, it could be possible to *develop a vocoder* suitable for low power real-time applications.

Generating new stimulation strategies using generative adversarial neural networks

Recent breakthroughs using *generative adversarial neural networks* have demonstrated the capability to *generate photorealistic portraits with arbitrary characteristics of the human face*. The setup of such approach is based on two neural networks competing between each-other, one trying to generate a random face, and the second one trying to guess whether it is a generated face or a real photo.

Inspired by this approach, it could be possible to setup a chain consisting of *two neural networks connected through* the *developed current spread* and *nerve excitation models*. The first neural network is being trained to transform audio inputs in a train of stimulation impulses, while the second neural network is trained to revert the nerve firing patterns back to the original sound.

With this approach, it might be possible to *discover completely new ways of stimulations* which are still transmittable by the auditory nerves, but might transfer significantly more information compared to the current approaches, *without the need to increase the number of the electrodes* (Pascual, et al., 2017).

REFERENCES

A, D. & C., E., 1957. Auditory prosthesis by means of a distant electrical stimulation of the sensory nerve with the use of an indwelt coiling. *Presse Med. 65(63):1417, Aug. 1957.*

Blake S. Wilson, M. F. D., 2008. Cochlear implants: Current designs and future possibilities. *Journal of Rehabilitation Research & Development, 45(5):695-730, Dec. 2008.*

Caner, G., Olgun, L., Gultekin, G. & Balaban, M., 2007. Optimizing fitting in children using objective measures such as neural response imaging and electrically evoked stapedius reflex threshold. *Otology & Neurotology, 28(5):637-40, Aug. 2007.*

Chen, F. & Zhang, Y.-t., 2006. A new acoustic model incorporating temporal fine structure cue for cochlear implant. *Proc. 5th Int. Special Topic Conf. on Information Technology in Biomedicine (ITAB 200), 4p., Sep. 2006.*

Chilian, A., Braun, E. & Harczos, T., 2011. Acoustic simulation of cochlear implant hearing. *Proc. Int. Symp. on Auditory and Audiological Research (ISAAR 2011), 3:425-432, Dec. 2011.*

Choi, C. T. M. & Lee, Y.-H., 2012. A review of stimulating strategies for cochlear implants. *in Cochlear Implant Research Updates, Intechopenp, pp. 77-89, Apr. 2012.*

Cui, Y., Xi Wang, L. H. & Soong, F. K., 2018. A new glottal neural vocoder for speech synthesis. *Proc. Interspeech, pp. 2017-2021, Sep. 2018.*

Dhanasingh, A. & Jolly, C., 2017. An overview of cochlear implant electrode array designs. *Hearing Research 356:93-103, Dec. 2017.*

Drennan, W. R. & Rubinstein, J. T., 2008. Music perception in cochlear implant users and its relationship with psychophysical capabilities. *Journal of Rehabilitation Research & Development 45(5):779-790, 2008.*

Eriksson, J. L. & Robert, A., 1999. The representation of pure tones and noise in a model of cochlear nucleus neurons. *The Journal of the Acoustical Society of America 106(4 Pt 1):1865-1879, Oct. 1999.*

Eshraghi, A. A. et al., 2012. The cochlear implant: historical aspects and future prospects. *Anat. Rec. (Hoboken) 295(11):1967-80, Nov 2012.*

Falcone, J. D. & Bhatti, P. T., 2011. Current steering and current focusing with a high-density intracochlear electrode array. *Proc. 33rd Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society, pp. 1049-1052, Sep. 2011.*

Fayad, J. N., Otto, S. R., Shannon, R. V. & Brackmann, D. E., 2008. Cochlear and brainstem auditory prostheses "Neural interface for hearing restoration: cochlear and brain stem implants". *Proceedings of the IEEE 96(7):1085-1095, Jul. 2008.*

Ghildiyal, S., 2016. Cost of cochlear implant operation may go down to Rs 1 lakh. *The Times of India, Mar. 2016.*

Gross, A., 2003. Fitting techniques for the pediatric cochlear implant patient. *http://www.audiologyonline.com/, May 2003.*

Hallpike, C. S. & Rawdon-Smith, A. F., 1934. The "Wever and Bray phenomenon." A study of the electrical response in the cochlea with especial reference to its origin. *J Physiol. 81(3):395–408, Jun. 1934.*

Hanekom, T. H., 2002. Modelling of the electrode-auditory nerve fibre interface in cochlear prostheses. *PhD Thesis, University of Pretoria, Jun. 2001.*

Harczos, T., Chilian, A. & Husar, P., 2013. Making use of auditory models for better mimicking of normal hearing processes with cochlear implants: the SAM coding strategy. *IEEE Trans. Biomed Circuits Syst. 7(4):414-425, Aug. 2013.*

Harczos, T., Szepannek, G., Katai, A. & Klefenz, F., 2006. An auditory model based vowel classification. *Proc. 2006 IEEE Biomedical Circuits and Systems Conf., Dec. 2006.*

He, Y., Zhang, H. & Wang, Y., 2019. RawNet: Fast end-to-end neural vocoder. *Proc. Interspeech, arXiv preprint arXiv:1904.05351, Apr. 2019.*

Hochmair, I. et al., 2015. Deep electrode insertion and sound coding in cochlear implants. *Hearing Research 322:14-23, Apr. 2015.*

Hochmair, I. et al., 2007.). MED-EL cochlear implants: State of the art and a glimpse into the future. *Trends in Amplification 10(4):201-219, Jan. 2007.*

House WF, 1976. Cochlear implants. *Ann Otol Rhinol Laryngol 85 suppl 27(3Pt2):1-93, May-Jun. 1976.*

House WF, U. J., 1973. Long term results of electrode implantation and electronic stimulation of the cochlea in man. *Ann Otol Rhinol Laryngol 82(4):504-517, Jul.-Aug. 1973.*

Kiss, J. G. et al., 2003. Neural response telemetry in cochlear implant users. *International Tinnitus Journal 9(1):59-60, 2003.*

Kosaner, J., Anderson, I., Turan, Z. & Deibl, M., 2009. The use of ESRT in fitting children with cochlear implants. *Journal of Int. Advanced Otology 5(1):70-79, Jan. 2009.*

Kuczapski, A., 2015. Metodă pentru auralizarea sunetelor percepute prin intermediul implantelor cohleare. *OSIM Romania, Patent No. a 2015 00879, Dec. 2015.*

Kuczapski, A. M. & Andreescu, G.-D., 2016. Modelling and simulation of hearing with cochlear implants: A proposed method for better auralization. *Soft Computing Applications. Advances in Intelligent Systems and Computing, Springer, 357:753-767, Nov. 2016.*

Kuczapski, A. M. & Andreescu, G.-D., 2016. Real-time interfacing for fault detection and auralization with MED-EL cochlear implant processors. *Proc. IEEE 11th Int. Symp. on Applied Computational Intelligence and Informatics (SACI 2016), pp. 191-195, May 2016.*

Kuczapski, A. M. & Andreescu, G.-D., 2017. New autocorrelation based self-learning method to detect sound spectral components in cochlear nerve firing patterns in case of cochlear implants. *Proc. 40th Int. Conf. on Telecommunications and Signal Processing (TSP 2017), pp. 429-434, Jul. 2017.*

Kuczapski, A. M. & Stanciu, A., 2015. Assistive tool for cochlear implant fitting: Estimation and monitoring of the effective stimulation thresholds. *Proc. IEEE 10th Jubilee Int. Symp. on Applied Computational Intelligence and Informatics (SACI 2015), pp. 307-311, May 2015.*

Kuczapski, A. M. & Stanciu, A., 2015. Computer Aided fitting, estimation and long term monitoring of effective stimulation thresholds for children with cochlear implants. *Proc. 12th European Symp. on Pediatric Cochlear Implantation, Toulouse, Jun. 2015.*

Lin, Y. et al., 2011. Chinese disyllables tone perceptual characteristics and the effect of stimulation rate on tone recognition in cochlear implants. *Proc. 4th Int. Conf. on Biomedical Engineering and Informatics (BMEI 2011) 3:1300-1304, Oct. 2011.*

Loebach, J. L., 2007. Cochlear implant simulations: A tutorial on generating acoustic simulations for research. *Progress Report No. 28, Indiana Univ., pp. 359-368, 2007.*

Lopez-Poveda, E. A. & Eustaquio-Martin, A., 2006. A biophysical model of the inner hair cell: The contribution of potassium currents to peripheral auditory compression. *Journal of Association for Resarch on Otolaryngology 7(3):218–235, Sep. 2006.*

Mahalakshmi, P. & Reddy, M. R., 2012. Investigation of the envelope and phase information for improved speech perception using an acoustic simulation model for cochlear implants. *Proc. 2012 IEEE-EMBS Int. Conf. on Biomedical Engineering and Sciences, pp. 555-558, Dec. 2012.*

Mahalakshmi, P. & Reddy, M. R., 2012. Speech processing strategies for cochlear prostheses - the past, present and future: A tutorial review. *International Journal of Advanced Research in Engineering and Technology 3(2):197-206, Jul.-Dec. 2012.*

Meddis, R. & Lopez-Poveda, E. A., 2010. Auditory periphery: from pinna to auditory nerve. In: *Ch.2 in Computational Models of the Auditory System, Springer, 35:7-38, Feb. 2010.* s.l.:s.n.

MED-EL(1), n.d. Automatic gain control provides a carefree listening experience. *http://www.medel.com/technology-automatic-sound-management.*

MED-EL(2), n.d. FineHearing™. *https://www.medel.com/int/show2/index/id/1362/title/FineHearing/.*

MED-EL(3), n.d. I100 detector box - User Manual. *MED-EL Elektromedizinische Geräte GmbH.*

Miller, D. A. & Matin, M. A., 2011. Modeling the head related transfer function for sound localization in normal hearing persons and bilateral cochlear implant recipients. *Proceedings of 14th International Conference on Computer and Infonnation Technology, Dec. 2011.*

Moctezuma, A. & Tu, J., 2011. An overview of cochlear implant systems. *BIOE 414:1-20, 2011.*

Nascimento, L. T. & Bevilacqua, M. C., 2005. Evaluation of speech perception in noise in cochlear implanted adults. *Brazilian Journal of Otorhinolaryngology 71(4):432-438, Jul.-Aug. 2005.*

Noble, J. H., Labadie, R. F., Gifford, R. H. & Dawant, B. M., 2013. Image-guidance enables new methods for customizing cochlear implant stimulation strategies. *IEEE Trans. Neural Systems and Rehabilitation Engineering, 21(5):820-829, Mar. 2011.*

Nogueira, W. & Buechner, A., 2012. Conveying low frequency information through analog electrical stimulation in cochlear implants. *Proc. 20th European Signal Processing Conf. (EUSIPCO 2012), pp. 509-513, Aug. 2012.*

Nogueira, W., Buechner, A. & Edler, B., 2005. Fundamental frequency coding in NofM strategies for cochlear implants. *Audio Engineering Society Convention 118, May 2005.*

Nogueira, W. et al., 2007. Automatic speech recognition with a cochlear implant front-end. *Proc. 8th Annual Conf. of the International Speech Communication Association (INTERSPEECH 2007), 4 pp, Aug. 2007.*

Pascual, S., Bonafonte, A. & Serra, J., 2017. SEGAN: Speech enhancement generative adversarial network. *Proc. INTERSPEECH 2017. arXiv preprint arXiv:1703.09452, Mar. 2017.*

Patterson, R. R. K. H. J. M. D. Z. C. a. A. M., 1992. Complex sounds and auditory images. *in Auditory Physiology and Perception, Pergamon, 83:429-446, Jan. 1992.*

Ramsden, R. T., 2002. Cochlear implants and brain stem implants. *British medical bulletin 63(1):183-193. Oct. 2002.*

Rebscher, S. et al., 2008. Considerations for design of future cochlear implant electrode arrays: Electrode array stiffness, size, and depth of insertion. *Journal of Rehabilitation Research & Development 45(5):731-748, 2008.*

Schnupp, J., Nelkel, I. & King, A., 2011. Auditory Neuroscience: Making Sense of Sound. *MIT Press, 2011.*

Smith, Z. M., Parkinson, W. S. & Long, C. J., 2013. Multipolar current focusing increases spectral resolution in cochlear implants. *Proc. 35th Annual Int. Conf. of the IEEE Eng Med Biol Soc. (EMBC), pp. 2796-2799, Jul. 2013.*

Somek, B., Fajt, S. & Ana Dembitz, M. I. J. O., 2006. Coding strategies for cochlear implants. *Automatika 47(1-2):69-74, May 2006.*

Stanciu, A. N., 2007. Small Details with Significant Impact in Fitting Prelingual Children. *5th European Balkan Congress - Hearing Implants and High Tech Hearing Aids, Zagreb, Oct. 2007.*

Stanciu, A. N., Cosgarea, M. & Necula, V., 2011. First Results Using Parallel Stimulation in MED-EL Cochlear Implants. *13th Symposium on Cochlear Implants in Children, Chicago, Jul. 2011.*

Stanciu, A. N. et al., 2008. Setting a Baseline for Vocal Audiograms in Romanian Language. *10th International Conference on Cochlear Implants and other Implantable Auditory Technologies, San Diego, Apr. 2008.*

Stanciu, A. N. & Hellmuth-Zweyer, U., 2015. Telefitting as part of the rehabilitation management. *12th European Symposium on Pediatric Cochlear Implantation (ESPCI 2015), Toulouse, Sep. 2015..*

Stöbich, B., Zierhofer, C. & Hochmair, E., 1999. Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy. *Ear and Hearing 20(2):104-116, Apr. 1999.*

Taitelbaum-Swead, R. et al., 2005. Speech perception of children using Nucleus, Clarion or Med-El cochlear implants. *International Journal of Pediatric Otorhinolaryngology 69(12):1675-1683, Dec. 2005.*

Tan, Q. & Carney, L. H., 2003. A phenomenological model for the responses of auditory-nerve fibers. II. Nonlinear tuning with a frequency glide. *The Journal of the Acoustical Society of America 114(4):2007-2020, Oct. 2003.*

Vargas, J. L. et al., 2012. Long-term evolution of the electrical stimulation levels for cochlear implant patients. *Clinical and Experimental Otorhinolaryngology, 5(4):194-200, Dec. 2012.*

Voigt , H. F. & Zheng, X., 2010. The cochlear nucleus: The new frontier. In: *in Computational Models of the Auditory System, Springer, 35:39-63, 2010.* s.l.:s.n.

Vondrasek, M., Sovka, P. & Tichy, T., 2008. ACE strategy with virtual channels. *Radioengineering 17(4):55-61, Dec. 2008.*

Wang, et al., 2016. The relative contributions of temporal envelope and fine structure to mandarin lexical tone perception in auditory neuropathy spectrum disorder. *in Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing, Springer, 894:241-248, Jan. 2016.*

Wang, S., Xu, L. & Mannell, R., 2011. Relative contributions of temporal envelope and fine structure cues to lexical tone recognition in hearing-impaired listeners. *Journal of the Association for Research in Otolaryngology 12(6):783-94, Dec. 2011.*

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

Wilson, B. S., Lopez-Poveda, E. A. & Schatzer, R., 2010. Use of Auditory Models in Developing Coding Strategies for Cochlear Implants. In: *Computational Models of the Auditory System.* s.l.:Springer, pp. 237-260.

Wouters, J., Doclo, S., Koning, R. & Francart, T., 2013. Sound processing for better coding of monaural and binaural cues in auditory prostheses. *Proceedings of the IEEE 101(9):1986-1997, Jul.2013.*

Wouters, J., McDermott, H. J. & Francart, T., 2015. Sound coding in cochlear implants: From electric pulses to hearing. *IEEE Signal Processing Magazine 32(2):67-80, Feb. 2015.*

Xu, Y. et al., 2019. Design, fabrication, and evaluation of a parylene thin-film electrode array for cochlear implants. *IEEE Trans. Biomedical Engineering 66(2):573-583, Jul. 2019.*

Yushi Zhang, W. H. A., 2006. Gammatone auditory filterbank and independent component analysis for speaker identification. *Proc. 9th Int. Conf. on Spoken Language Processing, 2006.*

Zeng, F.-G., 2008. Cochlear implants: System design, integration, and evaluation. *IEEE Reviews in Biomedical Engineering 5(1):115-142, Nov. 2008.*

Zeng, F.-G.et al., 2015. Development and evaluation of the Nurotron 26-electrode cochlear implant system. *Hearing Research 322:188-99, Apr. 2015.*

Zhang, W., 2010. Measurement and modelling of head-related transfer function for spatial audio synthesis. *PhD Thesis, Australian National Univ., Aug. 2010.*

Zhang, X., Heinz, M. G., Bruce, I. C. & Carney, L. H., 2001. A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression. *Journal of the Acoustical Society of America 109(2):648-670, Feb. 2001,* pp. 648-670.

Zirn, S. et al., 2015. Perception of interaural phase differences with envelope and fine structure coding strategies in bilateral cochlear implant users. *Trends in Hearing 20:1-12, Sep. 2016.*