TEMPORAL INTEGRATION IN COCHLEAR IMPLANTS AND THE EFFECTS OF HIGH STIMULATION RATES

Miguel Eduardo Obando Leitón

Graduate School of Systemic Neurosciences

LMU Munich

Dissertation der
Graduate School of Systemic Neurosciences der
Ludwig-Maximilians-Universität München

6th May 2019

Supervisor
Prof. Dr.-Ing. Werner Hemmert
Professur für Bioanaloge Informationsverarbeitung
Technical University of Munich

First Reviewer:	Prof. Dr.-Ing. Werner Hemmert
Second Reviewer:	Prof. Dr. Christian Leibold
External Reviewer	Prof. Dr.-Ing. Uwe Baumann

Date of Submission:	6th May 2019
Date of Defense:	21st October 2019

Abstract

Although cochlear implants (CIs) have proven to be an invaluable help for many people afflicted with severe hearing loss, there are still many hurdles left before a full restoration of hearing. A better understanding of how individual stimuli in a pulse train interact temporally to form a conjoined percept, and what effects the stimulation rate has on the percept of loudness will be beneficial for further improvements in the development of new coding strategies and thus in the quality of life of CI-wearers.

Two experiments presented here deal on the topic of temporal integration with CIs, and raise the question of the effects of the high stimulation rates made possible by the broad spread of stimulation. To this effect, curves of equal loudness were measured as a function of pulse train length for different stimulation characteristics.

In the first exploratory experiment, threshold and maximum acceptable loudness (MAL) were measured, and the existence and behaviour of the critical duration of integration in cochlear implants is discussed. In the second experiment, the effect of level was further investigated by including MAL measurements at shorter durations, as well as a line of equal loudness at a comfortable level.

It is found that the amount of temporal integration (the slope of integration as a function of duration) is greatly decreased in electrical hearing compared to acoustic hearing. The higher stimulation rates seem to have a compensating effect on this, increasing the slope with increasing rate. The highest rates investigated here lead to slopes that are even comparable to those found in persons with normal hearing and hearing impaired.

The rate also has an increasing effect on the dynamic range, which is otherwise taken to be a correlate of good performance. The values presented here point towards larger effects of rate on dynamic range than what has been found so far in the literature for more moderate ranges. While rate effects on threshold, dynamic range and integration slope seem to act uniformly for the different test subjects, the critical duration of integration varies strongly but in a non-consistent way, possibly reflecting more central, individual-specific effects.

Additionally, measurements on the voltage spread of human CI-wearers are presented which are used to validate a 3D computational model of the human cochlea developed in our group. The theoretical model falls squarely inside of the distribution of measurements. A single, implant dependent voltage-offset seems to adequately explain most of the variability.

Acknowledgements

This thesis would not have been possible without the continuous support of many different people to whom I owe my gratitude.

First, I want to give my sincerest thanks to Werner Hemmert for his continuous support throughout my stay at the BAI group, for his guidance on all aspects on this work, for never losing his confidence in me, and for his selfless warmth as a person.

I would like to thank Christian Leibold and Michael Pecka for being part of my thesis advisory committee and giving constructive criticism and guidance. In addition, I want to thank Prof. Uwe Baumann for kindly agreeing to serve as an external reviewer for the thesis, and Lutz Wiegrebe and Oliver Behrend for agreeing to participate in the examination committee.

To all members of BAI — Jörg Encke, Christian Wirtz, Robin Weiß, Steffanie Keller, Florian Völk, Sonja Karg, Ali Saeedi, Anja Kenn, and Heidi Schmitt-Irlinger – thank you for every single time you offered help or company, and for your advice in topics philosophical, neuroscientific, and personal. I want to specially thank Andrej Voss and Siwei Bai for sharing the noisy burden of the big-space office with me and making the ordeal enjoyable.

I would like to thank all the students that I was given the change to supervise during this work for trusting me with the opportunity and helping me grow: Zhongnan Qu, Peter Reiser, Andrew Curran, Patricia Schütz, Daniela Schwanda, Jakob Eberharter and Anna Dietze.

I would also like to express my gratitude to all test subjects who through their patience and diligent commitment made the experiments detailed here possible in the first place. Their selfless interest and curiosity were a main source of motivation for me.

I shout a big thank you to my families and my friends, to those near and those far away, for helping me become the person that I am. A special dedication is reserved for Aldo Rojas Supoy; despite being gone, his spirit carries on within everyone that was touched by him.

For his wisdom, which carried throughout the millennia never ceased to provoke insight, I want to thank Lucius Annaeus Seneca.

Lastly, for his unconditional and radical support, his indefatigable motivation during all kinds of moments, and for being a never-ending source of culinary and anatine inspiration I thank Dr. Fabian Wiescher.

A mi familia
Meiner Familie

Contents

List of Figures

Nomenclature

Φ normal cumulative distribution function, see equation (2.9), page 17

τ Time constant, page 14

T_C Critical duration, page 14

(e)ABR (Electrically evoked) Auditory Brainstem Responses, page 6

AB Advanced Bionics (Corporation), page 2

ABI auditory brainstem implant, page 19

AFC Alternative Forced Choice (method), page 9

AGC automatic gain control, page 58

ANOVA analysis of variance, page 46

BP Bipolar (stimulation, page 4

CI Cochlear Implant, page 1

CT computer tomography, page 47

DR Dynamic Range, page 2

ECAPs electrically evoked compound action potential, page 8

FEM Finite Element model, page 47

IFT impedance field telemetry, page 49

JND just noticeable difference, page 10

MAL Maximum acceptable level, page 2

MOA Method of adjustment, page 9

MP Monopolar (stimulation), page 4

MPI multi-pulse integration, page 19

NH normal hearing (as opposed to both electric hearing and acoustic hearing loss), page 10

OHC outer hair cell, page 6

pps pulses per second, page 5

RIB2 Research Interface Box 2, University of Innsbruck, page 46

SD, s (sample) standard deviation, page 56

SGN spiral ganglion neuron, page 20

TFS temporal fine structure, page 7

THR Threshold of stimulation, page 2

TI Temporal integration, page 13

Wf Weber fraction, $Wf = \Delta I/I$, where ΔI is the JND of a stimulus of amplitude I, page 56

WN white noise, page 46

Structure of the thesis

Chapter 1 – *Introduction* is the general introduction to the thesis. It includes an short introduction to the cochlear implant as a neural prosthesis, an overview of the response properties of the electrically stimulated auditory nerve, and some notes on psychophysical methods for experiments with cochlear implants.

Chapter 2 – *Literature Review* provides an overview of the literature on the phenomenon of temporal integration as well as models of loudness in both normal and electric hearing, and on the effects of stimulation rate in cochlear implants.

Chapter 3 – *Exploratory measurements of loudness integration in cochlear implant users at high rates* and Chapter 4 – *On the Effect of High Stimulation Rates on Temporal Integration in Cochlear Implant Users* take the form of publication manuscripts. Publication of the manuscripts is being worked on by the time of writing. They report on two sets of related psychophysical experiments on CI-wearers are reported. They had the goal of elucidating the integration of loudness in electric hearing for pulse trains of varying duration, as well as investigating the effects of a highly increased stimulation rate on loudness perception.

Chapter 5 – *Electrode Impedance and Voltage Spread* details on measurements of the voltage spread in the cochlea during electrical stimulation in humans, as a validation of a 3D-model that is currently being developed in our workgroup. The content of this chapter is best suited to be set in the context of a larger publication (Bai et al., 2019).

Chapter 6 – *Conclusion, outlook and closing remarks* presents concluding remarks summarizes the conclusions of the work presented here. The thesis is sent the context of existing literature, and some ideas for the outlook are given.

The introduction as a whole (chapters1 and 2), together with the conclusion (chapter 6) share a bibliography at the end of this work, while each of the manuscripts has a separate one of its own.

1

Introduction

1.1 The cochlear implant

Modern cochlear implants are considered to be the most successful sensory prosthetic, and promise to restore sound perception in otherwise deaf or hard-of-hearing persons. It can be used in cases of severe or profound hearing loss to restore the capacity of affected patients, as long as the auditory nerve fibres are still partially intact.

It consists of two parts, an external speech processor that also houses the microphone and batteries, and which can, in the optimal case, be replaced as technology progresses, as well as the internal implanted part, with an electrode array going into the cochlea. This internal part is designed to be passive and more robust, in order to avoid repeated surgery and the associated risks.

There, current pulses – usually biphasic – are responsible for stimulating the auditory nerve. In the purely ideal sense, the activation of the nerve should occur in a similar way that the original acoustic fibres would have caused, thus restoring hearing.

Both parts are separated from each other by the skin, and the external part is held in place magnetically. Radio frequency pulses sent and received by coils are responsible for energy and data transmission between the two parts.

The cochlear implant works by picking up external sounds through the external microphone and processing the waveforms into electric pulses which are then sent to the individual electrodes in the cochlear array. The exact steps of signal processing define the *coding strategy* and vary strongly between manufacturers. Implants from four manufactures have been introduced into the European market as of this work, those from Advanced Bionics, Cochlear, MED-EL and Neurelec. Because the software and hardware related to the CIs are not compatible with each other, the experiments portrayed in this thesis were all run on subjects with MED-EL implants.

Although individual results vary, in general CIs have proven to be highly beneficial for the affected population. Children born deaf or with a severe hearing loss can often develop spoken language in an age-appropriate speed if they are implanted at an early age (Wolfe and Schafer, 2014). Auditory stimulation during the sensitive period in the first years of early infancy is crucial for an adequate development of the neural pathways and the development of spoken language skills. As a result, the recommended age of implantation has continuously decreased during the last decades, nowadays being as low as 6 months (Valencia et al., 2008)[1].

Nevertheless, electric hearing is still far from being a complete replacement for normal capabilities. Although good speech understanding in silence is a realistic goal for new users (at least outside of tonal languages), there is much room for improvement concerning understanding in noisy or highly reverberant environments, the perception of music and tonality (crucial

[1] Besides the obvious benefit for the implantee, society at large also has a financial benefit from early implantation – implantation during the first year, as opposed to the second, results in net savings of over 20 000 € for society over a period of ten years, mostly due to costs for education and for the family (Colletti et al., 2011).

in languages like Mandarin), and the spatial localization of sounds. Additionally, many CI users still perform badly when identifying both speakers (Vongphoe and Zeng, 2005) and the emotion in their voices (Luo, Fu, and Galvin III, 2007).

1.1.1 Some historical notes

Mudry and Mills (2013) summarize the history of cochlear implants, and Shah, Chung, and Jackler (1997) review the history of electric stimulation of the auditory nerve before CIs existed. One of the first records of the activation of the electric nerve through electricity are those from Alessandro Volta's experiments, as reported in a letter to the president of the Royal Society in 1800:

> Au moment que le cercle a été ainsi complété, j'ai reçu une secousse dans la tête ; et, quelques moments après, (les communications continuant sans aucune interruption,) j'ai commencé à sentir un son, ou plutôt un bruit, dans les oreilles, que je ne saurois bien définir ; c'étoit une espèce de craquement à secousse, ou petillement, comme si quelque pâte ou matière tenace bouillonnoit. Ce bruit continua sans relâche, et sans augmentation, tout le tems [sic] que le cercle fut complet, &c. La sensation désagréable, et que je craignis dangereuse, de la secousse dans le cerveau, a fait que je n'ai pas repété plusieurs fois cette expérience. (Volta, 1800)
>
> *At the moment when the circle was thus complete I received a shock in the head, and some moments after [...] I began to hear a sound, or rather a noise in the ears, which I cannot well define; it was kind of a crackling with shocks, as if some paste or tenacious matter had been boiling. This noise continued incessantly, and without increase, all the time that the circle was complete. The disagreeable sensation, which I apprehended might be dangerous, of shock in the brain, prevented me from repeating the experiment.* (as translated in Shah, Chung, and Jackler, 1997)

It would not be until 1961 until the first implantation of an electrode into the cochlea proper by William House and John Doyle. At this point, the implant consisted of a single electrode and could at most be used as a help in lip-reading.

In late 1977 and early 1978, the first patients were implanted with a modern, multichannel cochlear implant, invented independently by Hochmair and Clark.

The last major milestone came from the invention of the continuous-interleaved-sampling (CIS) coding strategy by Wilson, Finley, Lawson, Wolford, et al. (1991). Up to that point, analogue waveforms were presented simultaneously in all electrodes, which led to high interactions between the channels, and presented information in high-frequency waveforms that could not be effectively used by the cochlea.

Figure 1.1 shows the historical development of cochlear implant technology from the perspective of sentence recognition of patients. Especially noteworthy are the difference between single-electrode and multichannel implants, and the increase in speech understanding after the development of modern CIS-based coding strategies (SPEAK, ACE, and so on).

The cochlear implant is now widely adopted, and a routine measure in rich countries, with an increasing number of people helped by this technology. Two reports by the NIH set the number of implanted people world wide at 219 000 by December 2010, and at 324 200 by end of 2012 (NIDCD, 2013, 2016).

1.1.2 Fitting process

The *fitting* or the *programming* of a cochlear implant describes the process through which an audiologist sets the different parameters of the speech processor that define how sound waves are to be transformed into electrical pulses. An appropriate fitting has an important contribution on the later performance of implanted patients and their outcome (Geers, Brenner, and Davidson, 2003).

The most important and common parameters that are adjusted during the fitting procedure are the threshold of stimulation and the upper stimulation levels. These establish the electrical dynamic range (DR) onto which the incoming acoustic information has to be mapped.

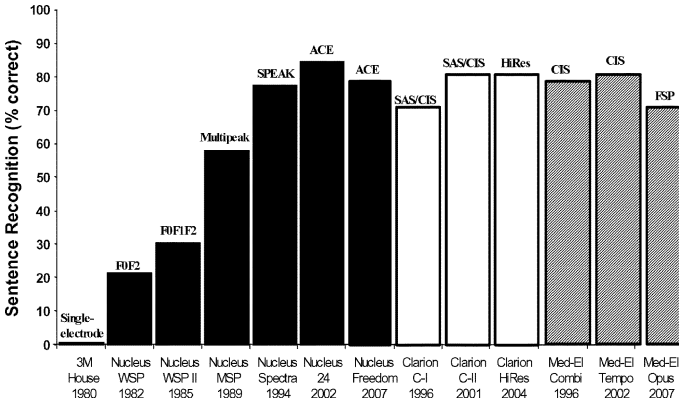


FIGURE 1.1: Historical sentence recognition scores for different cochlear implant models and coding strategies up to 2007. From Zeng, Rebscher, et al. (2008), © 2008 IEEE.

Other parameters of stimulation include stimulation rate, as well as pulse width, non-linear channel gain ("maplaw compression" in MED-EL systems), frequency allocation on the individual channels and input dynamic range, in addition to coding strategy-specific parameters, like the number of active channels in n-of-m strategies. Moreover, modern speech processors have additional settings and processing steps that can be included, ranging from microphone directionality to specific wind-noise suppression or an automatic gain control system. Wolfe and Schafer (2014, chapter 2) provide a comprehensive review of the fitting process and the relevant differences between the procedures for CIs of different manufacturers, as well as the manufacturer-specific nomenclature. Vaerenberg et al. (2014) report statistics on current practice in fitting centers worldwide, based on a questionnaire, and it includes the frequency with which different available parameters are usually set.

Because it is a time consuming process, the THR levels are often not measured in clinical practice. This is reflected in default or recommended settings by the manufacturers. In MED-EL fitting software, the T-levels are by default set at 0 CU, and for Advanced Bionics (AB) they are set at 10 % of the upper limits (Hughes, Goehring, et al., 2016). In the case of MED-EL, 70 % of fitting centres follow this default, as do 22 % of the centres in the case of AB implants (Vaerenberg et al., 2014). Because the large majority of patients have a dynamic range smaller than 20 dB (less than 2 %, see Van Der Beek, Briaire, and Frijns, 2015), stimulation near thusly selected THR levels is typically inaudible.

There is however evidence that setting the THR to values higher than these default settings leads, on average, to better speech perception (Holden, Reeder, et al., 2011). Rader et al. (2018) report on a method of determining THR levels that overshoots the values from the standard clinical procedure by 9 CU, and at the same time leads to better scores of speech understanding in noise and is subjectively preferred.

Van Der Beek, Briaire, and Frijns (2015) review the fitting data on 151 adult (Cochlear) CI-recipients. They report on the population-wide spread of T and M levels, as well as their dynamic range. In general, speech perception correlates significantly with dynamic range but the effect is not very strong ($r^2_{\mathrm{DR}} = 0.109$). A significantly better vowel recognition in subjects with larger dynamic ranges was also found by Cosendai and Pelizzone (2001) and Fu and Shannon (2000). However, there is some data by Bento et al. (2005) in which differences in DR were not found in CI-groups separated by speech understanding of sentences in quiet.

1.1.3 Nomenclature

As noted above, the most important parameters during clinical fitting are the lower and higher limits of electrical stimulation. However, the exact definitions and names of these limits differ

between the CI manufacturers as well as among researchers. In this thesis, I will adhere to the following definitions:

- The threshold, hence abbreviated as **THR**, is the minimal amount of current that a subject can detect with certainty, for a certain set of stimulation parameters. In the literature, these are also called "T-levels", and are sometimes abbreviated as "THS". Other alternative definitions used are the level of 50 % detection accuracy, or the highest level without a response.
- The level of maximum acceptable loudness, abbreviated as **MAL** in this work, determines the upper limit of stimulation. It is here defined as the maximum level of stimulation that is "loud, but not uncomfortable", as judged by the wearer. Other possible abbreviations include "MCL", where *MC* can possibly mean either "most-comfortable" or "maximum-comfort", and *L* can stand for either "level" or "loudness". It is also often called "C-level" or "M-level".[2]

The difference between these levels is called the (electric) dynamic range (DR), and can be expressed either logarithmically in dB or in absolute current units.

When describing the magnitude of stimulation, the different manufacturers also use their own units of magnitude. Some refer to the current amplitude or the total charge in one phase of the pulse in a linear scale, while others refer to a logarithmic scaling. Table 1.1 shows a summary of the different units and their equivalents.

Table 1.1: Different clinical units of stimulation magnitude, as used by some CI manufacturers. These values are all approximates. CU: Current unit. QU: Charge Unit. CL: Current Level (Units).

Manufacturer		Unit	µA	nC[3]	dB
Advanced Bionics	Before 2003	CU	1	–	–
	After 2003	CU	–	12.8447[4]	–
MED-EL		CU	1	–	–
		QU	–	1	–
Cochlear	CI22, CI24M/R	CL	–	–	0.176 (re 10 µA) [5]
	Nucleus Freedom	CL	–	–	0.157[6]

Quantifying levels through charge units is also a sensible choice because prolonging the phase duration at a fixed amplitude value increases the total charge and consequently the loudness of a stimulus. The loudness for a given charge is roughly constant for monopolar stimulation (Bonnet et al., 2012). For bipolar pulses with the same charge however, lower phase lengths produce louder percepts. The effect of phase duration is stronger at maximum loudness than near threshold levels (Zeng, Galvin III, and Zhang, 1998), and in general, perceived loudness grows more rapidly with current level for longer pulse phase durations (Chatterjee, Fu, and Shannon, 2000; Shannon, 1983).

1.1.4 Mode of stimulation

Another parameter of stimulation that plays an important role is given by the flow of current between different electrodes in the cochlea. The more normal case is called monopolar (MP) stimulation, which is a misnomer. There, only a single electrode in the cochlea is active at a

[2] Particularly confusing is the notation used by the Advanced Bionics' clinical fitting procedure. It includes the following loudness descriptions, in increasing order: "Most Comfortable", "Loud But Comfortable", "Maximal Comfort", "Uncomfortable". (see e.g. Landsberger and Galvin, 2011)

[3] $1\,\text{nC} = 1\,\mu\text{A} \times 1\,\text{ms}$

[4] Advanced Bionics Corporation (2003), Holden, Reeder, et al. (2011), and Zwolan et al. (2008)

[5] Drennan and Pfingst (2006), $I = 10 \cdot 175^{\text{CL}/255}\mu\text{A}$

[6] Azadpour, McKay, and Svirsky (2018)

time, and current flows to a ground electrode outside of the cochlea (sometimes directly to the encasing of the implant itself). In other modes (bi-, tri- and quadrupolar are some of the mentioned ones), one or more electrodes in the array are used for the return current, instead or in addition to the ground electrode.

In general, monopolar stimulation leads to a lower energy consumption, a broader spread of excitation (Snyder, Middlebrooks, and Bonham, 2008), with the benefit of reduced current amplitudes (e.g. THRs), and less variable thresholds between different electrodes (Bierer et al., 2015). This is attributed to a variable density of surviving neurons (or neural processes) which are more specifically stimulated by the focused stimulation. Although it is thus presumed that focused stimulation should lead to a better tonotopic representation of sounds and that a poor tonotopy is a main cause of impaired speech understanding with CIs (Friesen, Shannon, Baskent, et al., 2001), there is not much data showing perceptual benefits of focused stimulation.

1.1.5 Stimulation rates

When using current pulses, as opposed to e.g. sinusoidal current in older coding strategies, the rate with which stimulation pulses are output is usually given with pulses-per-second (pps) as a unit, instead of Hz. Commonly used rates in clinical settings range from around 900 pps to 3000 pps. The exact number will depend heavily on the manufacturer and the utilized coding strategy, as well as on the number of active stimulation sites – the total stimulation rate is usually fixed so the per-channel rate increases if electrodes are deactivated. The resulting rate is most commonly held equal across all the active electrodes (Hughes, Goehring, et al., 2016; Wolfe and Schafer, 2014, pp. 84–91). Exceptions are coding strategies that use different rates in the most apical electrodes in an attempt to better transmit temporal fine structure information, like MED-EL's FS4. In the strategies that explicitly avoid simultaneous stimulation on several electrodes in order to minimize electrical cross-over, the stimulation rate will also be limited by the width of the electrical pulses.

At the time of writing, the default speech coding strategy for MED-EL implants is FS4. In its default settings, the four apical fine-structure channels are each stimulated at a rate of 3000 pps, and the rest with a rate of 750 pps. In sum, the default total stimulation rate is 18 000 pps, for an average rate of 1500 pps per channel[7]. In the FSP coding strategy, the pulse rate for all 12 electrodes is usually chosen to be 1240 pps, 1485 pps, or 1635 pps.

1.2 Response properties of the electrically stimulated auditory nerve fibres

The cochlear implant works by direct stimulation of the auditory nerve fibres (ANF), which respond by firing action potentials. This process bypasses by design the usual pathway in normal ears via the basilar membrane and the inner hair cells (IHC). This leads to important differences in the response of the ANFs compared to acoustic stimulation. Broadly speaking, there are three types of differences: temporal effects of the electrical stimulation, the spatial effects due to the spread of excitation, and those arising due to the fact that the IHCs are missing.

1.2.1 Missing hair cells and synapses

In healthy humans, one inner hair cell has synapses to around 20 to 30 type I fibres (Pickles, 2012). Neurotransmitter release from the IHC is probabilistic in nature, and is modulated in response to voltage changes. In absence of sound stimuli, the auditory nerve fibres still fire spontaneously depending on their type, possibly through a tonic release of neurotransmitter. This has the positive effect of an increased sensitivity to near-threshold stimuli. In comparison, spontaneous firing and the associated IHC-caused variability are completely missing in the deafened ear (see Heffer et al., 2010, for nerve recordings in guinea pigs).

[7]Private communication with a clinical audiologist at Klinikum rechts der Isar.

The IHC-ANF synapse also is one of the sources of adaptation in the system, which is mediated by the production and recovery dynamics of the synaptic vesicles. After a sustained stimulus, less vesicles are directly available in the pre-synaptic ribbons, which leads to a decreased exocytosis and ultimately to decreased firing rates in the ANF. This effect is also absent in electrical hearing.

The outer hair cells (OHCs) are responsible for a non-linear compression, where only stimuli with a low amplitude are amplified. This is one of the mechanisms through which a large dynamic range can be achieved in the hearing system, and it is lacking in electrical stimulation. In general, the dynamic ranges of individual nerve fibres decrease from 20 to 50 dB SPL acoustically (depending on the fibre type, see Pickles, 2012) to as little as 1 to 4 dB at 200 pps of electrical stimulation (Javel and Viemeister, 2000). The rate-level functions of individual nerve fibres are thus much steeper in electrical stimulation.

Lastly, the implantation of CIs is an invasive surgery with strict indications – especially that the hearing loss is of high severity, such that the recipient could only benefit to a limited extent from the use of hearing aids, if at all. This automatically implies some deviation from the norm of the normal hearing population, be it through e.g. trauma or a pathological condition, which will have an impact on the response of the nerve fibres. In addition, the processes of the ANF are known to retract and degenerate in the prolonged absence of stimulation (Webster and Webster, 1981, guinea pigs), and longer term untreated hearing loss also leads to atrophy in the cochlear nucleus (Hardie and Shepherd, 1999, cats). In contrast, treatment with cochlear implants has a protective effect against degeneration. See Shepherd and Hardie (2001) for a review of both morphological and functional changes in the auditory pathway caused by deafness on humans and animal models. On a similar note, the review by Pfingst, Bowling, et al. (2011) summarises the effects of hearing loss on the human cochlea, as well as the effects of those changes on the later effectiveness of cochlear implants.

1.2.2 Spatial effects & spread of excitation

Near threshold, electrically stimulated neurons are stimulated at their dendrites, but with increasing electrical amplitudes, the place of stimulation shifts inwards along the nerve towards the axon (van den Honert and Stypulkowski, 1984). The electrode starts stimulating not only the peripheral processes (dendrites and soma) of the spiral ganglion neurons, but also parts that are central to the cell bodies, located in the modiolus. There, the fibres are more densely packed, so that a stronger recruitment with increasing current occurs (van den Honert and Stypulkowski, 1986; Javel, 1990). As an example, the growth of loudness as a function of stimulus amplitude tends to show a knee-point, after which it starts increasing much faster than at lower amplitudes. It has been suggested that this is caused when the modiolus starts getting stimulated by the electric pulses (McKay, Henshall, et al., 2003).

There is a large impedance difference between the bone and the cochlear lymph, which results in a broad spread of the voltage as a response to excitation (Ifukube and White, 1987; Malherbe, Hanekom, and Hanekom, 2016; see also Chapter 5). As a consequence, the effect of the activation of neighbouring electrodes is far from negligible, and it can well lead to a higher effective pulse rate for each individual fibre.

Lastly, while different fibre types respond differently to acoustic stimulation, with different spontaneous rates and single fibre dynamic ranges (Liberman, 1978), electrical pulses do not differentiate in this way between them, which further reduces the available dynamic range in electrical stimulation.

1.2.3 Temporal effects

Because electrical stimulation bypasses the middle ear and the conduction of the stimulus along the axis of the basilar membrane, electrical stimulation actually leads to faster auditory nerve responses compared to acoustical stimulation, specially with hearing aids. This can be easily seen in the latencies of the electrically evoked auditory brainstem responses (eABR),

which are lower than acoustically evoked responses by approximately 1 ms (van den Honert and Stypulkowski, 1986).

In addition, the envelope of the travelling wave on the basilar membrane produces an additional frequency-dependent delay that is also not present in electrical stimulation.

For low frequency acoustic stimuli (below $\sim$4 kHz), nerve fibres show a phase locking response, where the nerves preferentially fire at a particular phase of the sound wave. In CI stimulation, the stimulus pulses are much shorter, with durations on the order of tens of microseconds. As a response, the nerve fibres show much stronger and more precise phase locking and a less variable firing rate (van den Honert and Stypulkowski, 1984; Javel and Shepherd, 2000).

In the case of acoustic stimulation, the firing rate of neurons is also used to code the intensity of sound (Pickles, 2012). Electrically, since the response of neurons follows the stimulation so closely and because there is less adaptation in the system, this rate-coding loses reliability (Zhang et al., 2007). It is assumed that increasing electric stimulation amplitude leads to a stronger nerve response not through increased firing rates or probabilities, but through an increase in the number of activated neurons due to a larger spread of excitation (McKay and McDermott, 1998).

In the CIS strategy, only the envelope information is presented to the auditory nerve, while the temporal fine structure (TFS) is not. Although some current strategies try to present this by, for example, the modulation of stimulation with a high rate, the range of rates that can be processed and discriminated seems to be limited. This is surprising in light of the fact that ANFs under acoustic stimulation can use temporal cues (through phase locking) for frequencies up to 2 kHz, and that the electrically stimulated nerves of animals show a high temporal precision.

1.2.4 Response phenomena to electrical stimuli

Boulet, White, and Bruce (2016) review the literature on the most relevant response characteristics of the ANFs in the context of electrical stimulation, as well as the specific biophysical mechanisms that might mediate them. In a broader sense, four sets of phenomena are important:

- Refractoriness
- Temporal summation (or facilitation)
- Subthreshold adaptation (or accommodation)
- Spike-dependent adaptation (or spike-rate adaptation)

Although these characteristics are not unique to auditory nerve fibres since they are present in most active cells, they are particularly relevant in the context of electrical stimulation. In addition, this separation represents an abstraction, and is very hard if not impossible to observe in isolation from each other.

Refractoriness

After a neuron of the auditory nerve fires an action potential, it cannot fire despite further stimulation – the absolute refractory time. This is followed by a period of decreased sensitivity, called the relative refractory time, where the neuron shows an increased threshold to a second stimulus until it recovers back. The absolute refractory period lasts for a period of about 330 µs, while the relative refractory period has a recovery time constant of about 410 µs and may extend for up to 2 to 4 ms (Miller, Abbas, and Robinson, 2001). As a consequence of the refractory period, the nerve will show a reduced excitability for strong pulses for rates that go beyond around 250 pps (relative refraction) and 3000 pps (absolute refraction). It should be noted that the estimates of these refractory times vary among measurements, and seem to be sensitive to the number of data points and estimation methods used, among others (Boulet, White, and Bruce, 2016). Refractoriness has been posited to cause an oscillatory response of the auditory nerve at moderate stimulation rates, as measured with ECAPs (electrically evoked compound

action potential), which is taken to reflect the number of fibres that fire in response to each pulse (Hughes, Castioni, et al., 2012).

Temporal summation

Following a *sub-threshold* pulse that did not result in an action potential, the excitability of the neuron changes. During the first ∼300 µs, it is increased, which allows the neuron to fire to an otherwise sub-threshold pulse. It is mediated by capacitive charging of the neural membrane on one hand, and by changes in the activation of the Na^+ channels.

For prolonged stimulation, this summation can accumulate, so that many pulses occurring in a short period of time can combine constructively to provoke a spike. This effect increases with stimulation rate (Heffer et al., 2010), and is taken to be one of the main causes of multipulse integration (see chapter 2), where *psychophysical* thresholds decay with increasing stimulation rate, an effect that becomes more prominent at rates beyond ∼1000 pps (Zhou, Kraft, et al., 2015).

Recordings from animal nerves also support the notion that the threshold at high pulse rates is strongly mediated by integration over pulses. In the measurements of cat fibres of Zhang et al. (2007), thresholds of 20 µs phase/10 kpps pulse trains were even lower than 40 µs/5 kpps pulse trains.

Accommodation

For even longer intervals after a sub-threshold pulse, neurons show a reduced excitability. In the specific case of CIs and ANFs, not many measurements are available. The time range of accommodation extends to somewhere between 5 to 10 ms (Dynes, 1996). If the sub-threshold stimulation is repeated, the inhibiting effect seems to accumulate. This accumulating effect is closely related to supra-threshold (spike dependent) adaptation.

Adaptation

Spike-rate adaptation is the general name for the reduced excitability of a neuron after a prolonged stimulus provokes ongoing action potentials, and occurs on longer time scales than the previously mentioned effects. Several biophysical mechanisms combine to produce this phenomenon at different timescales, extending from tens of milliseconds (Zhang et al., 2007; Miller, Woo, et al., 2011) to at least minutes (Litvak, Smith, et al., 2003). Usually, the shorter term components can be well described as a decay with two-exponentials with time constants of around 10 ms and 100 ms, where the spike rate decays from a maximum rate towards a value where it stabilizes.

The amount of adaptation (usually quantified as the ratio between the maximum and minimum spiking rates) increases with stimulation rate. For simple pulse trains, the spiking rate usually falls and stabilizes at a lower value which decreases with rate. As the spiking rate decreases, the separation between accommodation and adaptation becomes fuzzier. Indeed, by investigating the *recovery* from adaptation, it has been shown that both sub- and supra-threshold maskers reduce excitability (Miller, Woo, et al., 2011), and that the ensuing recovery has a similar time course to adaptation.

Van Gendt et al. (2016) present a phenomenological model of the electrically stimulated ANFs, which they use to replicate physiological experiments with stimulation rates of up to 10 kpps. They model the adaptation as being composed of two components: a firing-dependent part and a stimulus-dependent one, both with time constants of 100 ms. At rates beyond 10 kpps and for stimulation with low amplitudes, the spike rates of the individual fibres can even lower down to almost zero during the first ∼50 to 100 ms, such that the neuron is blocked from spiking despite an ongoing stimulation. This was seen both in the physiological measurements as well as the model. It is also notable that their model is unable to accurately replicate the answers of nerve fibres to stimulation with 24 kpps, as measured by Litvak, Delgutte, and Eddington (2001).

1.3 Psychophysical methods for measurements with CI wearers

In the experiments presented in the following chapters, as well in the relevant literature, thresholds, maximum acceptable levels, and lines of equal loudness were measured. In these cases, the question arises as to which of several psychophysical methods to use. Especially for thresholds and equal-loudness, there are abundant possible methods described in literature (Zwicker and Fastl, 1999, chap. 1.3; Gescheider, 2013; or Gelfand, 2017, chap. 7, for a more detailed review).

Broadly speaking, three different procedures are particularly relevant for the tasks at hand here:

- Method of adjustment (MOA),
- Method of tracking (especially Békésy-Tracking), and
- Forced choice procedures (especially adaptive methods).

These methods all give a unique value as a result of the measurements, without explicitly giving the shape of the psychometric function as a result. The psychometric curve describes the probability of a correct/positive response as a function of stimulus level, and usually has a sigmoidal shape (also called "ogive" in older literature).

In the method of adjustment, the subject is given control over the stimulus, as well as a specific task. For example, they might be instructed to "set the stimulus to a level that is just barely audible" for threshold measurements. It has the main benefit of speed, in exchange for some loss of control on the stimulus by the experimenter.

In the method of tracking, introduced by Békésy (1947), the stimulus level changes at a set rate, with the subject controlling the direction of change. At the same time, a second parameter of the stimulus (e.g. tone frequency in the original method, but possibly also the duration of a pulse train) is steadily varied. In this way, the level as a function of the second parameter is tracked by the subject. It provides a reasonable balance of speed and precision, but depends on additional variables like the subjects' reaction time.

In typical forced-choice methods, the subject is presented with several intervals, only one of which is a correct response (e.g. has a signal, in measurements of threshold). The subjects are then forced to choose among the alternatives (thus AFC – alternative forced choice), and feedback might be given. Often, the stimulus level (or another relevant quality) is changed depending on the subject's response (or perhaps several responses), in what are called adaptive procedures (Levitt, 1971). They are described by the rules that dictate when and in which direction to change the level of a sound. For instance, a 2-down-1-up procedure will decrease the level of a sound only after two correct responses in a row, but will increase the level every time the response is wrong. Different procedures converge at different points of the psychometric curve, e.g. at the point of 70.7 % probability of a correct response for 2-down-1-up.

These methods have a high precision and completely avoid a particular kind of bias; subjects cannot reliably "cheat" into e.g. lower thresholds than real. They have the main drawback of increased duration and less control of the stimulus by the subject. These effects can lead to tiredness and boredom in participants, with associated worsening of performance, or worse, less subject participation.

1.3.1 Comparison of methods

For the experiments presented here, the method-of-adjustment was used. The measurement of maximum acceptable levels (MALs) do not allow to approach levels from above and below, in order to avoid overstimulation. This ruled out both tracking and adaptive methods. It seemed appropriate to not use different methods for similar measurements as are THRs and MALs, especially if a single method could otherwise be used for all experiments. In addition, the literature supported the method-of-adjustment as providing an acceptable balance between precision and testing-time, which was especially important considering that some experimental

sessions ran for about 4 hours. As noted above, it also has the added benefit of a more active participation and a heightened sense of agency by the test subjects.

Acoustic literature

In the context of frequency difference-limen, Wier, Jesteadt, and Green (1976) compared MOA and AFC, and found that MOA leads on average to markedly lower difference limens, hinting at a higher precision. Similarly, for measurements of masked intensity just-noticeable-differences (JND) Turner, Horwitz, and Souza (1994) measured lower values with the MOA than with AFC. In these two cases, it was not the amplitude of the stimulus that was the target measurement (absolute sensitivity), but its variability (differential sensitivity). In both cases, there is the possible caveat that MOA and AFC methods might measure a fundamentally different quality. However, it ought to be clear that the MOA is not intrinsically worse or less precise.

Boullet (2004) compares MOA and 2-AFC, (as well as magnitude estimation and multitrack-AFC) for measuring acoustic thresholds, and concludes that the MOA gives the best precision-duration compromise. For a similar standard deviation of repeated measurements, the 2-down-1-up forced-choice procedure had 4 times the duration of the MOA.

CI literature

Van Wieringen and Wouters (2001) compared MOA and 4-AFC (as well as a counting method) for measuring CI thresholds. On average, the found THR values were not different, but had more variability. While the test-retest variation of the MOA was larger by a factor of 2, the duration is only 25 % in comparison to the commonly used 2-down-1-up convergence (or 35 % of the 1-down-1-up procedure).

Several measurements in the workgroup of Bryan Pfingst also show agreeable results for the method of adjustment. In Pfingst and Xu (2004), it was found that the MOA tended to result in slightly higher THR-levels than forced-choice methods (by 1.5 dB in monopolar configurations), but there was no difference in across-site variation. Here, subjects were instructed to set a "barely audible" signal level, which the subjects seemingly tend to situate somewhat higher than the level of 70.7 % correct responses of the AFC. In addition, the results by Zhou, Kraft, et al. (2015) are taken to "[show] that the method of adjustment and adaptive tracking methods yielded very similar [MPI slope] results in these subjects" (cited in Zhou and Pfingst, 2016a).

Another possible method suggested in the literature is given by Bierer et al. (2015), who implement a method similar to Békésy-Tracking for measuring thresholds across electrodes in the CI-array, and compares it with AFC. It showed marginally lower test-retest reliability (rms errors were higher by under 1 dB re 1 µA), at 1/4 of the testing time. This method is a useful alternative when measuring a curve as a function of electrode position, and could in theory be extended to e.g. amplitudes as a function of duration, as were measured in chapters 3 and 4.

Rader et al. (2018) propose a different kind of force-choice, were subjects are tasked with counting two stimuli of varying amplitudes, and responding whether they perceived one, both, or none. However, they did not compare this method to other psychophysical method but only to an unspecified "standard clinical procedure", which does not have an exact meaning (Vaerenberg et al., 2014).

1.3.2 Biases and the Method of Adjustment

Although the method-of-adjustment is relatively fast and simple to understand for the subjects, there are known pitfalls that can lead to systematic measurement biases. Marks and Florentine (2011) shortly sum up the classical literature on the measurement of equal loudnesses with this method in normal hearing (NH) subjects: Depending on the intervals between successive stimuli, subjects might tend to judge either one of them as louder. Also, the fixed stimulus tends to be overestimated. There is a regression towards comfortable levels. To minimize bias, they recommend to randomize the order of fixed and variable stimuli, to use an unmarked and

continuously variable attenuator, to include tests of consistency, and to adjust the test stimulus to a reference and the reference to the test stimulus.

Subjects show a subtle asymmetry during loudness balancing tasks. If two sounds are compared, different results arise depending on which sound is fixed and which one is variable. In general, the variable stimulus will be set higher than the fixed stimulus. In acoustic experiments, this difference amounts to $\sim$2 dB (Zwicker, 1958; Scharf, 1961). When adjusting a magnitude to a standard, Poulton (1989, chap. 3) also recommends the variable to start alternately larger and smaller than the standard.

These recommendations were included in the second experiment, where a loudness balancing was conducted for stimuli of differing duration.

In addition, for thresholds subjects were explicitly instructed to raise the level of the stimulus until it could be confidently detected, and for both thresholds and balancing trials, asked to bracket the level (approach it from below and above). This was for example also the method in Zhou and Pfingst (2016a,b), who also measured THR in CI with MOA. Increasing the level first until it can be confidently detected should reduce the effect of tinnitus on stimulus detection, which is a common problem in CI-subjects when the microphones are switched off.

1.3.3 Balancing and qualitative aspects

When comparing and adjusting the loudnesses of stimuli with different durations, it is important to note that other qualitative aspects of the perception can and do change with duration. This has long been noted in the case of auditory signals (Ekman, Berglund, and Berglund, 1966), which can be explained, at least partially, due to the spectral spread of time-limited signals. Even though this effect does not play an important role with the short biphasic pulses of CI stimulation, the perceptive qualities of shorter pulse trains differ from longer pulse trains, and subjects often report that it is more difficult to determine the loudness of those shorter pulse trains.

Subjects are on general not always able to fully separate the loudness and the duration of a stimulus, so that either variable confounds the other. This effect can be reduced by running a loudness balance instead of, e.g., loudness estimation, and by emphasising in the instructions that duration and loudness are different aspects (Stephens, 1974).

1.3.4 A note on safety issues

Throughout the experiments presented in this work, a research interface was used to stimulate CI-wearers with specific combinations of pulses (see the Methods part of Chapters 3, 4 and 5). Because the use of a research interface bypasses the audio-processor and all its inbuilt safety mechanisms, it is up to the experimenter to ensure that patients will not be overstimulated. Because stimulation parameters were used that deviate from usual clinical settings (especially concerning the stimulation rate), utmost care was taken that the patients could always select comfortable levels of stimulation by themselves (which limits the available psychophysical methods, see subsection 1.3.1), that stimulation could be interrupted at any time, and that patients had ample opportunity to take breaks. For a summary of best practices regarding the use of research interfaces as well as possible safety concerns, the reader is referred to the work by Litovsky et al. (2017).

2

Literature Review

The literature on temporal integration is much more developed in the case of acoustic, and specially normal hearing, going back to the 1940s. For this reason, an overview of temporal integration in normal acoustic hearing is given, before elaborating on TI on CIs – see also the introductions of the manuscripts in chapters 3 and 4.

2.1 Temporal integration in normal hearing & phenomenological models

Temporal integration (TI) is the commonly used name for the general observation that in psychophysical experiments, in order to maintain a constant response, the amplitude of an otherwise identical stimulus decreases as its duration increases.

It has been observed in both acoustic hearing (normal and impaired) and electric hearing, as well as other sensory modalities. Moreover, it exists not only on detection thresholds, but also for supra-threshold stimuli. Shorter signals need a higher level than otherwise identical, longer signals, in order to be judged to have an equal loudness.

The first detailed descriptions of the phenomenon come from Kucharski (1927) and Békésy (1929). Because the first experiments of this type produced amplitude-duration functions of roughly the form $I_{\mathrm{THR}} \cdot t = E_0$ for some constant energy, it was at first taken to mirror an integration of intensity (or a neural correlate of intensity) over time – thus the name of temporal integration.

Later work quickly pointed out that the curves were systematically shallower, with a slope on the range of 7 to 9 dB per decade (Florentine, Fastl, and Buus, 1988; Gerken, Bhat, and Hutchison-Clutter, 1990; Plack and Skeels, 2007). In addition, the phenomenon does not hold for arbitrary durations – the slope of integration becomes shallower at longer durations.

Algom and Babkoff (1984) give a comprehensive review of the specific subject of the temporal integration at threshold on normal hearing, and Heil, Matysiak, and Neubauer (2017) briefly summarizes the existing literature, including more recent results.

In the literature described below, the word "model" has been used interchangeably to refer to merely descriptive functions for the $I(t)$-curve, to phenomenological descriptions that can be used for predictions, and in the least of cases, to biophysical or neural explanations.

In general, three different types of models have been proposed so far, which will be outlined in the following, along with their motivations:

- Perfect or leaky integration, with a long time constant, motivated by the finding of less integration for longer durations
- Short integration time constant, combined with compressive non-linearities, motivated by the fast temporal resolution of the auditory system
- Probabilistic models, motivated by experiments with short sound bursts and theories of integration in the visual system

2.1.1 Critical duration in NH

From the measured data, it is clear that the inverse relationship between amplitude and duration is not valid for all durations. For durations beyond the order of some 100 ms, the fall in amplitude is smaller, and in some of the existing reports, disappears at some point before 2 s. Two similar yet distinct concepts are commonly used to describe the measurement in integration and to reduce them to a single number in the literature: the *critical duration* and a *time constant*.

As a first approximation, the amplitude-duration curve can be described as consisting of two intersecting lines (in logarithmic axes): one with constant slope $I \cdot t = E_0$, and a constant line $I = I_\infty$. The time at which the lines intersect has been called the critical duration (T_C). Since the transition between maximum and no integration is not abrupt, it has also been used in cases where this linear approximation was not made. However, the definition becomes less exact, which has been criticized (Algom and Babkoff, 1984). A second option arises when using other functions to describe the form of the curve, like an exponential. In these cases, the decay is defined by the time constant (τ) of the function. In general, it will often be the case that, for the same data, an estimated time constant will be a lower limit on estimations of the critical duration: $\tau \leq T_C$.

The estimation of either value show great variability across the literature, ranging from 25 ms (Niese, 1959; Reichardt, 1970), up to at least 230 ms (Stevens and Hall, 1966). In addition, the estimate of the value changes with level, and tends to be smaller for supra-threshold stimuli (150 ms in Stevens and Hall, 1966; around 100 ms in Zwislocki, 1969). Scharf and Houtsma (1986) provide a review of works on temporal summation and the found values for T_C and τ.

According to Verhey and Uhlemann (2008), at least some part of the observed variability can be explained by the difficulty of subjects in comparing the loudness of sounds with very different durations.

2.1.2 Long integration models

The first way of explaining the observed results is to assume that the auditory system integrates either the intensity of the sound directly, or the resulting neural activity.

As mentioned before, a perfect, non-leaky integration, would result in the simple relationship:

$$I = \frac{k}{t}. \tag{2.1}$$

In order to explain the deviations for long durations, Garner and Miller (1947) proposed that there is a minimum effective intensity I_∞ for the auditory system, beyond which there is a perfect integration. This so called "diverted-input hypothesis" leads to a relationship of the form:

$$t \cdot (I - I_\infty) = k. \tag{2.2}$$

Although this might seem arbitrary and was quickly superseded by other explanations for the behaviour at long durations, corrections with the shape of diverted input have been revisited in recent models, either at the level of stimulus amplitude, or at the level of the neural response (Meddis and Lecluyse, 2011, see subsection 2.1.6).

An alternative interpretation which is mathematically equivalent is given by Hughes (1946). They proposed a TI function of the form:

$$\frac{I}{I_\infty} = 1 + \left(\frac{\tau}{t}\right), \tag{2.3}$$

which is equivalent to Equation 2.2 for $k = \tau \cdot I_\infty$. Here, τ is the time constant of an integrating element that stores energy, and should then be in the order of hundreds of milliseconds in order to describe the data[1].

[1] These measurements included durations up to 739 ms. Further measurements at longer durations could not be completed due to "the outbreak of war".

In a similar way, Plomp and Bouman (1959) described the phenomenon with a leaky integrator, in this case not unlike an RC circuit. This element integrates the intensity, or rather it integrates a neural effect in the auditory system that is proportional to the intensity. This continues until a critical threshold level is reached, which is when perception occurs. This leads to a relationship of the shape:

$$I = \frac{I_\infty}{1 - e^{-t/\tau}}. \tag{2.4}$$

Their results for the time constant $RC = \tau$ go from 150 to 375 ms, depending on frequency[2]. They compare their model to the diverted-input hypothesis of Garner and Miller (1947) and the probabilistic model of Crozier (1940). At this point it should be noted that these different models lead to very small differences in the exact shape of the TI curve – the reader is referred to figures 2–4 in the paper by Plomp and Bouman.

In general, fitting of the data with the leaky integration model will lead to time constants in this order of magnitude – see Gerken, Bhat, and Hutchison-Clutter (1990) for a review, which found mean values ranging from 68 ms for noise signals to 588 ms for low frequency tones.

Penner (1978) describes a more general model, of which both the linear approximation as well as Equation 2.4 can be seen as special cases. $I(t)$ is the sound intensity as a function of time, $y(t)$ the perception over time, and $h(t)$ is the kernel for averaging.

$$y(t) = \int_0^\infty I(t-\tau)h(\tau)\,\mathrm{d}\tau \tag{2.5}$$

Under the assumption that perception occurs when $y(t)$ surpasses a certain value, perfect integration is then the special case of $h(t) = k$. Similarly, the approximation with two lines is given for a rectangular window between 0 and T_C, and the leaky integrator of Plomp and Bouman (1959) is given for an exponentially decaying kernel $h(t) = k \cdot \exp(-t/\tau)$.

This more general and mathematical approach to the subject was elaborated by Munson (1947) and Zwislocki (1960), by assuming that the input to the integrator is not the intensity of the stimulus directly, but a transformed neural correlate of it. Munson (1947) is one of the first papers to suggest an integration of nerve pulses for the loudness sensation, and to quantify the relationship between loudness and time with an electrical model with an exponential decay. The magnitude of sensation N as a function of time is given by:

$$N = kN' \int_0^t E_r(\tau)F(t-\tau)\mathrm{d}\tau, \quad \text{with} \quad r(t) = N' \cdot E_r(t). \tag{2.6}$$

$F(t)$ is the dissipation function, (equivalent to $h(t)$ in Equation 2.5), N' is a factor that depends on the number of excited fibres and the degree of excitations (i.e. on the intensity of the stimulus). $E_r(t)$ is an empirical function describing the adaptation effect due to continued stimuli[3], and $r(t)$ represents the rate of incoming pulses in the higher steps of the system.

With these mathematical models, it is possible to introduce an arbitrary amount of complexity into the theoretical framework, yet simpler forms are well able to reproduce the experimental data well (Algom and Babkoff, 1984; Robinson, 1974). Both simpler and more complex forms lead to an integration that occurs over longer time constants, and was confirmed for measurements of not only tones but also narrow-band noise (Zwicker and Wright, 1963). The relevant time constant usually is in the order of hundreds of milliseconds, with $\tau = 200$ ms being a commonly cited value for threshold stimuli, and around 100 ms for suprathreshold loudness (Zwislocki, 1969), but the value diverges in the literature.

2.1.3 Resolution-integration paradox

In a manner similar to Equation 2.6, Zwislocki (1960) developed the idea of a summation of quantal responses (neural excitation) with exponential decays into a theory that predicts

[2] Feldtkeller and Oetinger (1956) had already published a similar relationship, but without motivating the shape of the curve.

[3] In the specific case of Munson, the dissipation was exponential, and the adaptation had two terms, with $E_r(t) = 0.5 \cdot \exp(-t/\tau_1) + 0.5 \cdot \exp\left[-(t/\tau_2)^{0.15}\right]$, where $\tau_1 = 16.7$ ms and $\tau_2 = 1.33$ s.

the loudness of sounds of different durations, the threshold being a special case. After some simplifications, he also arrives at Equation 2.4, with a value of $\tau = 200\,\text{ms}$ for high frequency tones. He suggests that the integration process, however, occurs in rather central parts of the auditory system. This he saw as necessary in order to explain the short latencies and high discrimination abilities of the peripheral auditory system despite the long time constants of detection and integration.

In fact, when looking at the resolution of the auditory system and its discriminatory abilities such as modulation detection, gap detection, or masking, much shorter time constants of summation are the result – Moore et al. (1988) arrive at a temporal window consisting of two rounded exponentials, with an equivalent rectangular duration of 8 ms. On the other hand, a leaky integrator with a long time constant suggests a sluggish system with a low temporal resolution.

This contradiction between very short time constants regarding the resolution of the auditory system and the longer time constants of integration literature has been called the "resolution-integration paradox" (Viemeister and Wakefield, 1991). This problem allows several solutions, including one system with a variable time constant that is matched to the task or to the stimulus duration, separate systems with different time constants where which system gets used is task-dependent, or separate systems located at different stages of processing, with a slow integration being more central than a peripheral fast resolution, as Zwislocki suggested. Alternatively, a second type of integration model can also be motivated by this paradox.

See Eddins and Green (1995) for a review on the topics of temporal integration and resolution, as well as the areas where the results conflict.

2.1.4 Compressive short integration models

This second type of models has much shorter time constants, in the order of milliseconds, in order to reflect temporal discriminatory abilities such as modulation detection, gap detection, or masking. However, to explain the known decay of the threshold-duration function, further details have to be added. To this end, Penner (1978) extends Equation 2.5 with a compressive non-linearity, followed by the integration:

$$y(t) = \int_0^\infty [I(t-\tau)]^p \cdot h(\tau)\,\mathrm{d}\tau, \quad \text{with} \quad p < 1. \tag{2.7}$$

She mentions four motivations for the compression of the stimulus in the calculation of responses, i.e. the value of p:

1. Loudness as a function of intensity is compressed ($I^{1/3}$, Stevens' law, see Stevens, 1955).
2. This would also explain the failure of Weber's law.
3. Energy extends over 13 orders of magnitude, but firing rates over less than 3.
4. The cochlea itself has a compressive nonlinearity.

With the correct choice of $h(t)$, this function will result in an TI-slope of $-a$ in a way that can fit empirical data:

$$h(t) = \begin{cases} kap \cdot \hat{t}^{\,ap-1} & , \quad \hat{t} \geq 1 \\ k & , \quad \hat{t} < 1 \end{cases} \qquad \text{with } \hat{t} = \frac{t}{t_0} \text{ dimensionless.} \tag{2.8}$$

In addition, different shapes of the window function can create a flattening of integration beyond a certain duration. With an appropriate value for p, the exponent of t can be changed arbitrarily, which changes the effective width of the integration window $h(t)$. If $h(t)$ is compared to or approximated with exponential functions, much lower time constants arise than with models without compression. Penner claims that $p = 2/3$ leads to good fits for both results of temporal acuity (with time constants around 2 to 5 ms) and time-intensity (TI) data that would otherwise lead to time constants in the order of hundreds to thousands of milliseconds without the compression.

2.1.5 TI with short bursts

The short time constant of the integration window is in part motivated by results with repeated short bursts. As an example, Krumbholz and Wiegrebe (1998) measured the THR of short tone bursts with different interstimulus distances. The threshold of two short identical sounds separated by a silent gap does not change much if the gap grows beyond a few milliseconds – it stays constant in the range between 16 to 256 ms, roughly independent of the burst carrier frequency. This is at odds with longer time constants, and can in fact be seen as imposing a higher limit on their possible values.[4]

Moreover, they also demonstrate the existence of an effect of phase and its interaction with ringing in the auditory filters. Gerken, Bhat, and Hutchison-Clutter (1990) as well as Viemeister and Wakefield (1991) have similar results (albeit different interpretations) with clicks that have a broader frequency content. Another crucial observation by Viemeister and Wakefield (1991) is that the level of a masking noise presented in the interval between the individual bursts does not affect their threshold level.

As noted before, one of the main motivations of an exponential model of TI with a long time constant is the aforementioned critical duration (see subsection 2.1.1). However, not even this is uncontroversial in the literature. As an example, THR-levels for tone bursts in Florentine, Fastl, and Buus (1988) continued to decrease up to their maximum test duration of 500 ms – although the authors do comment that longer test durations would be necessary to estimate an integration time constant. Similarly, Penner (1978) did not find durations with constant threshold up to 1000 ms for broadband noises.

2.1.6 Probabilistic models

Attempts to explain threshold TI as a phenomenon without using an actual integration led to the development of more probabilistic models. Such were at first rejected (Algom and Babkoff, 1984), but have seen a comeback in the last decades.

One example of a probabilistic hypothesis was originally proposed by Crozier (1940) in the context of *visual* thresholds as a function of duration, given by a formula

$$I/I_0 = 1 \Big/ \Phi\left(\log_{10}\frac{t}{\tau}\right) \tag{2.9}$$

Here, Φ represents the normal cumulative distribution function. This equation was derived by assuming that the reciprocal value of the threshold intensity I^{-1} is a measure of the sensitivity of the system, and is based on the assumption that the thresholds of neurons follow a log-normal distribution. This was later tested in the context of audition by Garner and Miller (1947) and Plomp and Bouman (1959), but rejected by both.

Zwislocki (1960) also frames the question of temporal integration as having two different proposed solutions: one stemming from power summation (integration with large time constants) and a second one originating from statistical probability. He rejected the probabilistic hypothesis due to wrong predictions regarding experiments with pairs of clicks.

The probabilistic approach has seen a comeback in the last decades. A more recent type of model uses the concept of "multiple-looks" to combine the short and long time constants of previous models. A fast integrator ($\tau = 3$ ms), defining a single look, is joined with a short-term memory with a time constant around 300 to 500 ms. Stimuli with longer durations are not integrated, but instead allow for more looks to be observed, which increases the possibility of signal detection so that the threshold is reduced.

The model by Viemeister and Wakefield (1991) starts with a band-pass filtering of sound in a critical band. The output is passed by a nonlinearity (half-wave rectification), followed by a leaky integration with a rectangular window of width 3 ms. The running output of those windows is stored in a leaky memory with a longer time constant in the order of hundreds

[4] Interestingly, Zwislocki (1969) had cited that experiments with two pulses (e.g. Zwislocki, 1960) demonstrate that temporal summation *must* take place.

of milliseconds. The subject can then scan the whole vector of looks in memory to e.g. detect stimuli or modulations. They manage to approximate TI results with a long decay time by using a summation over the memory vector with a very specific weighting of the individual looks. Near the onset, the weights are minimal, increase rapidly to a maximum around 200 ms after onset and decay slowly after that.

The model by Heil and Neubauer (2003) expands on this by assuming in one "look" the stimulus generates an abstract *detection event* with a certain, low, probability that depends on the envelope amplitude. The effect of longer stimulus durations can then be explained by probability summation: if the criterion for detection is a fixed number of events, then a longer stimulus increases the chances of reaching detection. In their model, the *envelope* of sound amplitude is raised to an expansive power ranging from 3 to 5, and integrated over time. This work is later expanded, among others, in Heil, Matysiak, and Neubauer (2017), in the context of signal detection theory. The rate of a stochastic (Poisson point) process increases with the envelope of the bandpass-filtered stimulus amplitude. The subject bases her decision after a certain observation interval that is context-specific. In addition, they include a discounted amplitude – see the "diverted-input hypothesis" of Garner and Miller (1947) (subsection 2.1.2). They argue that the exponent of around 3 arises from the calcium-ion dynamics in the IHC–ANF synapse. Among other things, the model predicts "steeper psychometric functions for short than for long stimuli and, for a given stimulus duration, predicts that psychometric functions are steeper the higher the spontaneous event rate".

Meddis and Lecluyse (2011) follow along the same line. As Heil and colleagues, they assume that abstract events are generated with a rate that increases proportionally to the peak pressure of a stimulus (i.e. without exponentiation), minus a discounted rate.

2.1.7 Power-law and parameters affecting temporal integration

Despite the disagreement regarding different possible phenomenological models for TI, a simple description with a power-law (a single straight line in logarithmic axes) remains useful as a way to quantify individual measurements. While many parameters of the models vary widely across experiments (see, for example, the case of a critical duration), the actual value of the slope seems to be somewhat more consistent. For those datasets that show a flattening of THR at high durations, the slope can nevertheless be calculated for some set of short durations.

Gerken, Bhat, and Hutchison-Clutter (1990) argue for a power-law description rather than an exponential model because it best described their reviewed data, because it leads to a much smaller species difference between TI in humans and cats, and because the power-law model is less influenced by noise in measurements.

However, even with this there is still a high variability between the different experiments, and also across the subjects in a given experiment. In Plack and Skeels (2007, see Fig. 1a), data from different normal hearing subjects are shown as being equivalent to TIs ranging from 2.6 to 12.1 dB per decade.

Effect of windowing and the definition of duration Gerken, Bhat, and Hutchison-Clutter (1990) point out the necessity of a modified definition of duration to account for the reduced energy of the rise and fall of the signals and the ambiguity of its effect. In their results, the rise and fall of stimuli played an underproportional role in TI, although the lower levels of the auditory system would usually emphasize onsets. They take this discrepancy to be a sign that the "temporal integrator" is located in higher areas. In order to account for the results, individual sound bursts must be integrated before being exponentiated. The results of Heil, Matysiak, and Neubauer (2017) are also contingent on a modified definition of duration, in order to account for rise and fall times.

Effects of level Additionally, temporal integration seems to also be dependent on stimulation level. Florentine, Buus, and Poulsen (1996) measured the temporal integration for different levels of stimulation in subjects with normal hearing. TI, quantified as the level difference

between stimuli with durations of 5 ms and 200 ms, was not monotonic as a function of level, but was highest at moderate levels around 56 dB SPL. The authors conjecture that this can in part be caused by the mechanics of the basilar membrane. This relationship is further explored in Buus, Florentine, and Poulsen (1997) in the context of just-noticeable-differences – the same dependence on level appears.

Effects of hearing loss Hearing loss has been shown to reduce integration, with a negative correlation between the thresholds and the amount of temporal integration. Plack and Skeels (2007) tested temporal integration in NH and sensorineural hearing loss (SNHL) subjects at 2 and 4 kHz, at 4 and 24 ms (with 2 ms cosine windows). For NH-subjects, they measured slopes of temporal integration with an average (7.5 ± 2.9) dB per decade. SNHL-subjects had however reduced integration of 5.1 dB per decade. Oxenham, Moore, and Vickers (1997) showed that the difference between NH and HL lies mostly in long-duration stimuli. Temporal integration is not reduced in HL for durations of 2 to 10 ms.

Florentine, Fastl, and Buus (1988) measured TI in subjects with NH, SNHL and NH with simulated hearing loss with masking. They showed that the simulated hearing loss does not lead to the same reduction of temporal integration as the real hearing loss does. The effect is not caused by the shifting of thresholds, but by a loss in integration capability.

Effect of stimulation frequency Lastly, there is an effect of frequency on the slope of temporal integration. Gerken, Bhat, and Hutchison-Clutter (1990) reviewed 20 previous studies of temporal integration, and came to the conclusion that the data is best described as following a power-law function, again with an exponent of around 0.6 that varies with frequency.

This effect is subtle; for each doubling of frequency, temporal integration slopes decay by around 0.41 dB per duration decade (see figure 1 in the paper by Gerken et al.).

2.2 Temporal integration of electrical pulses & phenomenological models

2.2.1 Multipulse integration in electrical hearing and the effect of rate

A related concept to TI that arises only in the context of electrical stimulation is multi-pulse integration (MPI – sometimes also called "rate integration"), a term first proposed by the group of Bryan Pfingst to refer to the stimulation threshold decaying as a function of pulse-rate (Pfingst, Colesa, et al., 2011; Zhou, Kraft, et al., 2015).

Different mechanisms lie behind multipulse-integration. At very low pulse rates, ANF-fibres stimulated at threshold fire with a probability somewhere around 0.4 and 0.9 (McKay and McDermott, 1998). An increase in stimulation rates increases neural firing rates at both the individual fibre but also at the ensemble level by recruiting more fibres. Both of these processes can be assumed to lead to reduced psychoacoustical thresholds. At further increases of rate, beyond 1000 pps, facilitative integration effects at the level of the individual fibres play an additional role (see section 1.2).

In general, multiple studies have found a categorical difference in MPI for rates below and above 1000 pps – slopes at high rates are significantly steeper both in humans (Shannon, 1985; McKay and McDermott, 1998) and behaviourally in guinea pigs (Zhou, Kraft, et al., 2015). The latter group showed a high correlation between the TI slopes and the MPI slopes for low stimulation rates. They take this to indicate that TI and MPI below 1 kpps are mediated by the same mechanism, in a manner consistent with the multiple-looks model. The entrainment and refractoriness elicited by low rates would need a larger amount of neurons in order to elicit MPI. Only for higher rates would sub-threshold facilitation play the main role. This mechanism in turn reduces the dependency of MPI on neural density – and cochlear health in general.

Because of these effects, rate will have a strong effect on measures of temporal integration. Carlyon, Deeks, and McKay (2015) measured curves of TI between single pulses and 32 ms trains, at rates of 500 and 3500 pps in both CI-patients and auditory brainstem implant (ABI) users. Users with both devices showed temporal integration at both rates. However, CI users

additionally showed a drop of THR with rate for all durations, and an additional interaction effect, so that TI-slopes were steeper at a higher rate. Neither effect was prominent in the ABI-users – for the same duration, thresholds did not change much between the rates. Previous measurements had already shown that the effect of MPI in ABIs disappeared at rates of 200 pps (Shannon, 2011, as cited in Carlyon, Deeks, and McKay, 2015), indicating that the effect of MPI is particular to the neurons of the auditory nerve. The authors interpreted that this effect arises because individual brainstem neurons have a much smaller electric dynamic range before reaching firing saturation. When loudness increases as a function of rate, then it is due to recruiting of neighbouring neurons.

Middlebrooks (2004) measured TI in guinea pigs electrophysiologically via cortical thresholds. They measured little effect of stimulation rates below 1000 pps, and a fall of THR above. For instance, doubling the rate at 250 pps had little effect on THR, if any, but it caused a drop of several dB at higher (∼4000 pps) rates. Additionally, they also observed a significant fall in the THR of an electrode when applying sub-threshold trains at a fast rate at a different electrode. They interpret these results two-fold: first, there is an important effect of facilitation by which sub-threshold pulses result in a net residual charge, which reduces the threshold of following pulses; secondly, the electrical pulses are only integrated during a window of a few milliseconds.

In guinea-pigs, a stronger MPI at pulse rates below 1 kpps seems to correlate with spiral ganglion neuron (SGN) survival (Middlebrooks, 2004; Pfingst, Colesa, et al., 2011). However, these results do not easily extend to humans (McKay, Lim, and Lenarz, 2013). It must be noted that it is difficult to make cross-species comparisons and extend these results to humans for at least two reasons. On one hand, human auditory nerve fibres seem to be more robust than those of guinea pigs in the sense that the latter recede much faster in the absence of IHC stimulation. On the other hand, the innear ear of the guinea pig is much thinner, so that a cochlear implant can fill out the width of the scala, which might reduce variations caused by the position of the array. In the wider scalae of humans, there is more room for a variability of the distance between the electrode contacts and the neurons.

For more information on previous results on temporal integration in humans, the reader is referred to the introductions of the manuscripts in chapters 3 and 4.

As further expanded there, two important points are missing in the existing literature. First, the asymptotic behaviour of temporal integration of electrical stimuli at and above threshold, i.e. the critical duration, has not been investigated as it has in the normal hearing case. Second, TI – and for that matter, MPI – have been investigated in humans only for moderate to high rates which correspond at maximum to single-channel rates in normal CI use.

2.2.2 Loudness models in cochlear implants

At least three different phenomenological models have been suggested thus far in order to explain the process of temporal integration in cochlear implants. They all share an integration (of either intensity or nerve responses) over a relatively short period of time, and include a power-law non-linearity, not unlike the models in subsection 2.1.4.

The first models acted directly on the stimulus waveform. In the model of Shannon (1989), two different processes run in parallel, a compressive and an envelope extraction. However, only the latter is relevant for short pulse widths (below 400 µs, as was the case in the experiments of chapters 3 and 4).

In the envelope process, the stimulus is first half-wave rectified, and then expanded with a power-law function $y = k \cdot x^p$, with $p > 1$. This is then integrated with a time constant $\tau \sim 2$ ms. For the compressive process, which dominates for long pulse widths or slow sinusoidal currents and is thus less relevant for the daily stimulation of CI-wearers, no rectification is done, and $p < 1$. The maximum of the two outputs determines the output of the system.

As such, each process has three free parameters, although Shannon chose the same τ for both, based on previous measurements. A possible interpretation for the two processes is that the envelope process reflects the biophysical summation in individual neurons, while the

compressive process represents a more central integration of information coming from the spread of ANF activity, and could be related to cell survival or to synchronous activity.

As a limitation, however, the rectification step of the envelope process voids the effects of varying interphase gaps (Karg, Lackner, and Hemmert, 2013). It also does not allow for the differences caused by some different waveforms like pseudo-monophasic pulses, where the second phase has a longer duration but a lower amplitude (Macherey et al., 2006).

This was an important motivation for a second model, introduced by Carlyon, Van Wieringen, et al. (2005). In their model, the waveform is first passed through a low-pass filter that mirrors previous measurements of psychological threshold as a function of the frequency of a sinusoidal stimulus. The output of the filter is then multiplied with a Hanning window with a duration of 10 ms. This was done in time increments of 0.5 ms in the original paper, although the continuous-time limit should in theory work identically. For each of these windows, the RMS is calculated. The maximum RMS value is a measure of the magnitude of a stimulus, and its threshold is inversely proportional to it. This successfully predicts the results of experiments on interphase gaps, as well as by construction those with sinusoidal stimulation.

More modern models were introduced and developed by Colette McKay and colleagues over a series of at least 4 different publications:

- McKay and McDermott (1998)
- McKay, Remine, and McDermott (2001)
- McKay, Henshall, et al. (2003)
- McKay, Lim, and Lenarz (2013)

They introduced a *phenomenological* model with a short time constant (3 to 10 ms). In this model, they do not integrate the electric waveform, but an abstract measure of the ANF response ("peripheral neural excitation"). This type of model is capable of making predictions not only for TI in threshold, but also for loudness, as well as modulation detection and discrimination.

These series of models explicitly see themselves as extensions of previous acoustic models with a short integration and a *compressive* non-linearity (see subsection 2.1.4). This is in contrast to the models mentioned which act directly on the waveform and have an *expansive* step.

The first step of the model looks into the contributions of individual pulses and the temporal interaction of two pulses due to refractoriness (McKay and McDermott, 1998). The effect strength of a second pulse (E_2) relative to the first is modelled as a sigmoid function. The magnitude of neural response as a function of the interpulse interval (t) is given by

$$E_2(t) = E_1 \left(1 - \frac{R}{1 + e^{(\mathrm{IPI} - t_0)/D}}\right). \tag{2.10}$$

R represents the proportion of neurons firing on the first of two pulses, and is dependent on subject and stimulus level. Values of $t_0 = 7.3\,\mathrm{ms}$ (average refractory period) and $D = 0.8\,\mathrm{ms}$ (standard deviation of the refractory period) are able to fit the data well.

This output is then sent to a central temporal integrator (exponential, equivalent duration of 7 ms). A further paper specified a more complex shape for the summation window (McKay, Lim, and Lenarz, 2013). Their TI window has the same form as in papers by Oxenham and colleagues (Oxenham, 2001)[5], as obtained mostly through measures of masking:

$$W(t) = \begin{cases} (1-w)\exp(t/T_{b1}) + w\exp(t/T_{b2}) & t < 0 \\ \exp(-t/T_a) & t \geq 0, \end{cases} \tag{2.11}$$

where T_{b2} is longer and takes forward masking into account. Best fits for normal subjects are $T_a = 3.5\,\mathrm{ms}$, $T_{b1} = 4.6\,\mathrm{ms}$, $T_{b2} = 16.6\,\mathrm{ms}$, and $w = 0.17$. Importantly, these same TI window parameters can successfully be used for the experiments with CI users. The authors' interpretation is that processing of ANF activity seems to be largely unaffected in CI users; the difference would then lie mostly in the firing patterns themselves.

[5]See also Oxenham and Moore (1994) as well as Plack, Oxenham, and Drga (2002).

After temporal integration of *spike activity* at each position in the cochlea, the excitation density is scaled to a loudness value. This process introduces another non-linearity to the system. McKay, Henshall, et al. (2003) measured the current-loudness function, and found it to have two sections. For low amplitudes it follows a power-law, and above a subject- and electrode-specific point, the exponent of the power-law increases. Their explanation for this behaviour is that higher exponents at higher amplitudes might reflect recruitment in areas with more dense neural material (axons in central auditory meatus in comparison to cell bodies and dentrites in the spiral ganglion)[6].

At the last point in the models, overall loudness is obtained by spatial integration of the loudness values at each cochlear position (McKay, Remine, and McDermott, 2001).

2.3 Other psychophysical effects of high stimulation rates

2.3.1 Adaptation in electric stimulation

Higher stimulation rates of CIs increase the amount of adaptation in single nerve fibres. It seems that on electrical stimulation, the time course of adaptation is better known than the time course of temporal integration. In general, three different time scales of adaptation have been reported for both acoustic stimulation and electrical stimulation (Moxon, 1967, as cited in Litvak, Delgutte, and Eddington, 2001).

Litvak, Delgutte, and Eddington (2001) measured the short-term adaptation (in the order of hundreds of milliseconds) in cat ANFs as a function of a wide range of stimulation rates, going from 1200 to 24 000 pps. Electrically stimulated nerve fibres showed a strong response at onset and a gradual decrease over 100 to 300 ms (short-term adaptation). The amount of adaptation, quantified as the ratio between initial[7] and final spiking rates, increased with stimulation frequency. The authors state that the mechanism of adaptation is highly dependent on the voltage changes evoked by stimulation even if these do not evoke spikes – making this more of an accommodative adaptation, following the nomenclature of Boulet, White, and Bruce (2016). Similar results were later obtained by Miller, Hu, et al. (2008), who characterized the changes in the response properties of ANFs as a function of the level and rate of pulse trains.

Zhang et al. (2007) investigated the time course and time constants of adaptation as a function of stimulation rate and amplitude, also in feline ANFs. They found that the firing rate of ANFs is well described by a two-exponent model: a rapid and a short-term component (in the order of 10 ms and 100 ms respectively), both with time constants comparable to, though larger, than those found in acoustic studies.

The spiking rate of fibres stimulated with a pulse train tends to reach an asymptote well below 300 ms. They also found that the amount of adaptation (the normalized rate decrement, as in Litvak, Delgutte, and Eddington, 2001) increases with stimulation rate. It also decreases with stimulation amplitude – strong enough stimulation can partially overcome the effects of adaptation at pulse rates over 1000 pps. Lastly, the time constants of the exponential curve fits also decrease with rate, and are largely independent of stimulus level (cf. Fig. 9 in their publication).

2.3.2 The effects of increasing pulse rate on ANF responses

Temporal integration can also be readily seen in cochlear implant users, with some important differences. One of the main reasons behind the differences in CIs is that the electrical nature of the stimuli leads to much different aggregate responses in the ANFs (see section 1.2), depending on the stimulation rate.

In particular, electrical stimulation of the deaf ear with stimulation rates much below the refractory period of the nerves (<800 pps) lead to a very high synchrony of the firing times (Javel and Shepherd, 2000). This is partly caused by the absence of the input from the IHCs,

[6] This explanation was first postulated by van den Honert and Stypulkowski (1984), who conducted measurements of single ANF activity after electrical stimulation of (non-deafened) cats.

[7] as opposed to onset, 10 to 20 ms after it

which normally leads to spontaneous firing in the nerve fibres even in the absence of sound. Without this input, the timing of the action potentials of the electrically stimulated fibres will tend to cluster tightly in a short span of time following the electrical stimulation (Wilson, Finley, Lawson, and Zerbi, 1997).

The refractory period also plays a role at higher rates. It reduces the response to stimuli inside the refractory period, so that the neurons tend to respond again to a later following pulse. For a certain range of stimulation rates corresponding roughly to the refractory period (from around 300 to 2000 pps), this creates an alternating response during a pulse train. This behaviour can be observed in both direct recordings from the ANF of cats (Zhang et al., 2007), as well as in an alternating pattern in the amplitude of human ECAPs measurements (Hughes, Castioni, et al., 2012; Wilson, Finley, Lawson, and Zerbi, 1997).

Increasing the rate beyond around 5000 pps leads to yet another regime with reduced synchrony. Some 20 ms after onset, a so called "pseudo-spontaneous activity" arises. The alternating pattern in the ECAPs measurements decreases and disappears with increasing stimulation rate (Hughes, Castioni, et al., 2012; Wilson, Finley, Lawson, and Zerbi, 1997)[8]. This is attributed to the different refractoriness characteristics of the individual fibres and to the fact that the nerves adapt much more strongly at higher rates. These combined effects cause a desynchronization of the fibre population. This has been named "stochastic" (Hughes, Castioni, et al., 2012) or "pseudospontaneous" activity (Rubinstein et al., 1999). Apical electrodes keep showing alternating patterns for higher frequencies than basal electrodes (Hughes, Castioni, et al., 2012; Wilson, Finley, Lawson, and Zerbi, 1997).

Rubinstein et al. (1999) suggested that the effect at higher rates is more similar to auditory behaviour and might lead to a better representation of the temporal fine structure of sound and to an improved temporal resolution of the system.

They also propose the use of a conditioning high-rate stimulus which would induce a desynchronization of the ANFs. In their experiments, it was shown to eliminate the alternating ECAPs response caused by a 1016 pps stimulus, given a high amplitude. The addition of such a conditioning stimulation has also been shown to at least increase the electric dynamic range of sinusoidal bursts (Hong and Rubinstein, 2006). However, in the presence of a desynchronizing pulse train, not all temporal response properties resemble those of spontaneous firing (Litvak, Delgutte, and Eddington, 2001). The post-stimulus time histograms deviate strongly, showing a distribution with two peaks (see also Zhang et al., 2007). In addition, although the variability across stimuli presentations can be made comparable to the spontaneous case, this only occurs on a narrow range of rates and stimulation levels. This poses a problem, since the dynamic range of individual fibres in a single cochlea is very variable.

Despite these expected changes in the firing properties of the nerve fibre, it is not clear that an increased stimulation rate could lead to a better speech understanding.

2.3.3 Dynamic range

Besides the decreases in threshold with both duration and rate mentioned thus far, the maximum acceptable levels of cochlear implant users also show the same relationships. However, the slopes of both TI and MPI are shallower at maximum levels than at threshold, with the consequence that the electrical dynamic range increases with either pulse train duration or stimulation rate.

Previous measurements have quantified this increase in DR at 1.28 dB per doubling of pulse rate (octave) up to 5000 pps (Zhou, Xu, and Pfingst, 2012); 1.30 dB per octave to 3868 pps (Bonnet et al., 2012); and 1.2 dB per octave up to 6500 pps (Kreft, Donaldson, and Nelson, 2004b). In addtion, the latter group measured a *shallower* slope of MPI above 3250 pps for both THR and MAL, interpreting it as some saturation of the effect. The effect of duration on DR has not been explicitly investigated in this way.

[8] During the first milliseconds after the pulse-train onset, responses of individual fibres still bunch in time at specific intervals (around 4 ms for the measurements of Zhang et al., 2007, see Fig. 1 in their paper).

2.3.4 Stimulation rate and speech recognition

Due to the effects discussed above, the stimulation rate is also capable of strongly altering the performance of individual CI-users – usually measured as their ability to understand speech in either silence or noise.

The cochlear implant manufacturers (especially Advanced Bionics and MED-EL) encourage using higher and higher stimulation rates, and use as a selling point that their implant systems are particularly trimmed to using higher rates and allowing the users to benefit from them (Wolfe and Schafer, 2014; Zeng, Rebscher, et al., 2008).

This is at first glance understandable considering that better speech understanding correlates with larger DR *across CI-wearers* (Van Der Beek, Briaire, and Frijns, 2015), and that higher stimulation rates lead to larger dynamic ranges *for individual CI-users* (e.g. Bonnet et al., 2012, as well as chapters 3 & 4). Nevertheless, the existing evidence does not necessarily point towards higher rates producing better performance for speech understanding (Arora et al., 2009; Friesen, Shannon, and Cruz, 2005; Holden, Skinner, et al., 2002; Plant et al., 2007). Indeed, the optimal stimulation rate is individual and varies across users, though higher stimulation rates seem to be preferred for listening to music and radio (Vandali et al., 2000).

Interestingly, there is even a discrepancy between users' subjective preferences on one hand, and their actual performance as measured in controlled environments on the other. Many, but not all users prefer the rates with which they obtain best performance scores (see especially the study by Balkany et al., 2007, with 71 subjects). Whether this reflects some intrinsic factor of electrical hearing, or indicates that the methods used do not accurately reflect the normal conditions for the CI-users is an open question. What becomes clear from the data is that an individualized fitting, among others the precise setting of threshold values (Rader et al., 2018), harbours the possibility of more improvement than a blanket recommendation for higher rates.

There is even evidence from an interventional study by Pelosi et al. (2012) indicating that users with auditory neuropathy spectrum disorder (ANSD) and a worse auditory development actually can profit from decreased stimulation rates.

Two of the few studies showing an improved speech understanding with an increased stimulation rate are the works by Dunn et al. (2006) and Koch et al. (2004) (in the range of 3000 to 5600 pps as opposed to $<$1600 pps). However, these results are confounded by other parameters, because the change in stimulation rate was coupled to a general change in coding strategy.

Besides speech understanding, higher stimulation rates have neither been found to improve intensity resolution nor modulation sensitivity (Galvin and Fu, 2009).

2.3.5 Forward masking in CI users

A last point of methodological interest when dealing with CI stimulation in general and specifically with the possible effects of stimulation rate is the topic of (forward) masking, and especially the time course of recovery from masking.

In general, past experiments have shown that the temporal course of forward masking is similar in CI and NH (Chatterjee, 1999; Nelson and Donaldson, 2002; Shannon, 1990).

Nelson and Donaldson (2002) reported masking recovery time constants τ that varied across the subjects, but remained below 95 ms for 18 out of 21 subjects for stimulation at 500 pps. They found no effect of the masker level on τ.

Chatterjee and Kulkarni (2017) measured forward masking in 12 CI users at 1000 pps. In monopolar mode, they estimated a mean time constant of $\tau = 99$ ms. The level did not affect the time-constant of masking recovery (only the amount of masking).

2.4 Motivation

Modern CI coding strategies separate the incoming sound into their frequency components with e.g. filter banks. The outputs of these frequency channels are then used to modulate the amplitude of electrical pulses at the different electrodes in the implanted array, with the

goal that each electrode stimulates a specific population of auditory nerve in an independent manner. However, the conductivity of the liquid in the cochlea is much higher than that of the bone separating the electrode and the ANF processes. Thus, electrical stimulation in the cochlea creates a broad spread of the current that only decays slowly with distance along the volume of the cochlea. A population of nerve fibres near one of the electrodes will therefore also be influenced by neighbouring electrodes, thereby increasing the effective stimulation rate. The question then arises of whether known effects of temporal and rate integration can be extended to these higher, more realistic effective rates.

Despite a recent trend for manufacturers to support the use of higher stimulation rates, not much of the available evidence supports higher rates leading to better speech understanding or general performance in CI wearers. Although some theoretical benefits of higher stimulation rates have been proposed, experimental tests so far have not confirmed them. Instead, the optimal stimulation rate for speech understanding varies strongly across subjects without any obvious explanations.

In addition, the investigation of TI in normal hearing necessarily blends the effects brought forth by the cochlear structures and IHC–ANF synapse with those caused by the nerve fibres proper. Moreover, the investigation of very short acoustic stimuli necessarily leads to a spectral splatter which stimulates broader regions of the cochlea. With CI stimulation, both of these issues are bypassed, since the nerve is stimulated directly, and the spatial spread of stimulation does not change with the duration of stimulation.

The experiments detailed in the following chapters measure psychophysical temporal integration in a detailed manner, from single pulses to trains with long durations of up to 1 s, at different levels of stimulation and at two different stimulation rates, a moderate one similar to normal single-channel rates in common coding strategies, and a very high rate more similar to the maximum effective rate caused by cross-channel interactions. These measurements are novel due to the range of durations utilized, as well as the high stimulation rates investigated.

3

Exploratory measurements of loudness integration in cochlear implant users at high rates

Miguel Obando[1,2], Daniela Schwanda[3], Werner Hemmert[1,2]
[1)] Graduate School of Systemic Neurosciences, Ludwig-Maximilian Universität München
[2)] Bio-inspired Information Processing, Department of Electrical Engineering, Technical University of Munich
[3)] Munich University of Applied Sciences

Abstract

The broad spread of the electric fields in the implanted cochlea can greatly increase the effective rate with which individual neurons are stimulated, because every neuron is exposed to the stimulation from several nearby electrode contacts. In order to assess the effects of these increased effective stimulation rates on loudness perception, we conducted a psychophysical experiment to assess the integration of pulse trains as a function of the number of pulses. Particularly, we compared the behavior of normal and very high stimulation rates. Measurements at these high rates are novel, and can shed light on the temporal behavior of the electrically stimulated auditory nerve.

Four subjects (five ears) with Med-El cochlear implants participated in the experiment. We measured threshold levels (THR) and maximum acceptable levels (MAL). The durations of the trains varied from a single pulse up to 1 s trains. Subjects freely set the stimulation amplitudes (method of adjustment) for a total of three repetitions per parameter combination. The stimulations occurred at either apical or basal electrodes (E3 and E10, respectively – Med-El arrays have 12 electrode contacts), and at pulse rates of either 1200 or 25 000 pulses per second (pps).

Note: Parts of these results were presented in Obando, Schwanda, and Hemmert (2018).

3.1 Introduction

Monopolar electrical stimulation in the implanted cochlea leads to a very broad spread of electric fields along the cochlear fluid, mainly due to the difference of impedances along the cochlear lymph and across the bone (Ifukube and White, 1987; Kral et al., 1998; Tang, Benítez, and Zeng, 2011).

This effect greatly increases the effective rate with which individual neurons are stimulated. In the extremal case, a fibre would be affected by stimulation from all electrodes, with an effective stimulation rate going well beyond 20 000 pulses per second (pps), depending on the coding strategy and the phase width of the electric pulses.

In order to measure how the individual pulses combine over time to evoke a percept and to explore the effects of these increased rates, we conducted an exploratory psychophysical experiment to measure threshold (THR) and maximum acceptable levels (MAL) of pulse trains as a function of the number of pulses.

In particular, we compared the behaviour of normal single electrode rates, typical for the clinical setting, and of very high stimulation rates, near the maximum rate allowed by the stimulation hardware. If the effective rate can really increase due to the broad spread in the cochlea, the very high rate should be more representative of the everyday behaviour of the auditory nerve. From these results, we calculated both the dynamic range (DR) and a measure of the temporal integration (TI) for these rates.

3.1.1 Integration of loudness

Temporal integration (TI) describes the well observed phenomenon of the decrease in the threshold of stimuli with increasing duration, or alternatively, the decrease on level necessary to keep the stimulus at a constant loudness.

A related effect which exists only in electrical hearing is multi-pulse integration (MPI), the decay of electric thresholds with increasing pulse rate (at constant pulse train durations).

Previous MPI and TI experiments Human TI in cochlear implant users has been measured in several past experiments, but thus far it has only been measured up to more moderate rates, similar to those used in clinical settings for single electrodes in the

array (Donaldson, Viemeister, and Nelson, 1997; Eddington et al., 1978; Shannon, 1983; Zhou, Kraft, et al., 2015). Particularly, there is not much information about the location or existence of a critical duration of TI, after which amplitudes stop decreasing.

Most comparable experiments thus far have measured the dependency of the THR on rate rather than on stimulus duration (MPI) (Bonnet et al., 2012; Kreft, Donaldson, and Nelson, 2004; Shannon, 1985; Skinner et al., 2000; Zhou, Xu, and Pfingst, 2012). Rates of up to 6500 pps have been measured, but without taking into account the duration dependency.

It has been found that the slope of MPI increases at higher rates beyond 1000 pps (Zhou, Kraft, et al., 2015). This faster reduction in threshold has been interpreted as an integration of subsequent pulses through sub-threshold facilitation in the auditory nerve fibres (ANFs), which suggests an inherent time constant of integration.

Regarding the rate dependency, McKay and McDermott (1998) claim that TI can be best explained by a summation by a two-sided exponential window with a duration of 7 ms.

Effect on Dynamic Range One of the known effects of increasing the stimulation rate is an increase of the dynamic range. Both THR and MAL tend to fall as a function of rate, but THR falls more steeply, which leads to an overall increase in DR. The mean size of this effect has been reported to be in the range from 1.2 to 1.6 dB per doubling of pulse rate for rates of up to the range of 4000 to 6500 pps (Bonnet et al., 2012; Zhou, Xu, and Pfingst, 2012).

Kreft, Donaldson, and Nelson (2004) also measured the effects of rate on dynamic range up to 6500 pps in human CI-users. In addition to the decline of DR with rate, they also measure a flattening of the slopes beyond 3250 pps in both individual subjects and the mean data, and attribute this to a saturation effect.

Adaptation effects & loss of firing synchrony An additional effect that is present with higher stimulation rates is a loss of firing synchrony. While lower rates (<800 pps) effectively "lock" the firing of the neurons and cause them to fire more or less simultaneously with a high synchrony (Javel and Shepherd, 2000), more moderate rates create an alternating effect due to an interaction with the refractory period of the neurons, which is for example, visible in ECAP measurements (Hughes et al., 2012). At even higher rates, a quasi-stochastic regime appears, varying across subjects and individual electrodes. Neurons fall into a so-called "pseudo-spontaneous activity" with a reduced across-fibre synchrony, after responding preferably to the first pulse in a train (Wilson et al., 1997). Because the input stimulation is much stronger than neural noise in electrical stimulation, the individual fibres will run in limit cycles, and thus this measured quasi-stochasticity grounds on the biophysical properties of the individual ANFs. The temporal structure of the action potentials and the distribution of the firing rates becomes more similar to the acoustic case (e.g. cat ANF measurements by Litvak et al., 2003). It has been proposed that this desynchronization at very high rates should allow an improvement of the hearing quality (Rubinstein et al., 1999), because it should allow more of the temporal-fine-structure information to be represented and conveyed by the auditory nerve.

TI models in normal hearing Three different kinds of explanations have been given for this observed behaviour in normal hearing (NH) (see Verhey, 2010, for a more general overview).

The first one is based on a leaky integration of the sound intensity, not unlike an RC circuit integrates voltages, with a relatively long time constant τ on the order of hundreds of milliseconds (Plomp and Bouman, 1959; Zwislocki, 1969).

A second type of model has much shorter time constants (τ<10 ms), and was primarily motivated by experimental results from the integration of multiple bursts that are less compatible with a slower integrator. In order to explain the known decay of the threshold-duration function, further details have to be added. An example is a compressive non-linearity, which, if *followed* by an integrator, can nevertheless effect an integration at longer times (Penner, 1978).

A more recent type of model uses the concept of "multiple-looks" to combine the short and long time constants of previous models. A fast integrator ($\tau = 3$ ms) is joined with a short-term memory with a time constant around 300 to 500 ms (Viemeister and Wakefield, 1991). The models by Heil and Neubauer (2003) as well as Meddis and Lecluyse (2011) are of a similar probabilistic nature. There, the stimuli increase the probability of detection events, which are then observed over a variable, task-dependent interval.

TI models in CIs In electric hearing, the models that have been created to explain the behaviour of TI (and MPI) are of a *phenomenological* nature. The model by Shannon (1989) is composed of two parts, but for short phase durations (below some 400 µs), it is dominated by a process of envelope extraction. In it, the rectified stimulus is expanded by a power function ($y = x^p$, with $p > 1$), with an exponent that is a free parameter, specific to each subject and electrode. The output of this step is then integrated with a time constant of 1 to 3 ms, and the result is the magnitude of perception.

In the model of Carlyon et al. (2005), the electric waveform is first passed through a low-pass filter that is primarily designed to reflect the responses to sinusoidal electric stimuli. The output of the filtered signal is windowed with overlapping Hann functions (of 10 ms width), and the RMS of each window is calculated. Threshold is reached when the maximum

RMS value surpasses a fixed limit. One of the main differences between the models of Shannon (1989) and Carlyon et al. (2005) is that the former cannot reflect the effects of varying the inter-phase gap due to its rectification step.

The models proposed by McKay and McDermott (1998) (see also McKay, Lim, and Lenarz, 2013) first estimate the excitation caused by each pulse in a train, relative to the first pulse. This is mediated by a recovery from refraction and adaptation effects. These excitations are then integrated in an asymmetric exponential time window with an equivalent rectangular duration of 7 ms. The magnitude of the loudness perception is then identified with the maximum output of this integrator – the threshold is reached when a specific output is exceeded. These models do not take into account sub-threshold facilitation, and are thus unsuitable for rates above ∼2000 pps.

3.2 Methods

3.2.1 Stimuli

In the present experiment, we measured thresholds and maximum acceptable levels of pulse trains from single pulses to durations of 1 s, at rates of 1.2 and 25 kpps. All stimuli consisted of trains of biphasic, rectangular, monopolar pulses, with a phase width of 15 µs and an inter-phase gap of 2.1 µs, with a leading negative phase, presented in a single electrode.

We compared single pulses and pulse trains of two different rates: a clinically relevant stimulation rate of 1200 pps (pulses per second) and a high rate of 25 000 pps.

The duration of the 1200 pps pulse trains was varied between a single pulse and 500 ms, and between single pulse and 500 ms for the pulse trains with a 25 000 pps repetition rate.

Stimuli were presented on electrodes 3 and 10 (out of 12 in MED-EL electrode arrays, lower numbers stand for more apical electrodes).

Due to the difficulty of reaching MALs with very short pulse trains, they were only measured for trains of at least 20 ms.

3.2.2 Subjects

4 adult subjects (5 ears) with a MED-EL implant participated in this experiment, for a total of 10 measured electrodes. Their detailed demographic data can be found in Table 3.1. All subjects gave their informed consent for the participation and received monetary compensation. Measurements were conducted in accordance to the Declaration of Helsinki, and were approved by the medical ethics committee of the Klinikum rechts der Isar (Munich, 2126/08). For S1, data was collected in two sessions for the left and right CI.

3.2.3 Psychophysical Measurements

A Research Interface Box (RIB2, University of Innsbruck) was used to control the CIs. We used custom software written in Python to generate the stimuli and to let subjects adjust the stimulation amplitudes.

For the first part of the experiment, the threshold value (THR) of each pulse train was determined by letting the subjects adjust the amplitude of the repeating stimulus (Method of Adjustment). Subjects were instructed to increase the stimulation level until the signal could be clearly heard, and then decrease it until the stimulus could barely be perceived. Pulse trains were presented continuously once per second. As this did not work for 1 s trains, we decided to add a 500 ms gap in this condition. In this way, it is equal to the condition with the 500 ms stimulus, which makes at least these two conditions comparable.

Subjects were first allowed to familiarize themselves to the presented stimuli and the adjustment controls in a training session.

Users were able to change the current amplitude by either large or small steps. Large steps were of approximately 28.4 CU until the first reversal of the THR measurement and 18.9 CU afterwards. Small steps had an amplitude of 1.2, 2.4, 4.7, or 9.5 CU, as limited by the resolution of the stimulation hardware and depending on the current amplitude. 1 CU is roughly equivalent to 1 µA.

A total of 46 duration/electrode/rate combinations formed a block. Blocks were shuffled and repeated for a total of 3 measurements per parameter combination. All further results refer to the median measurement. This same procedure was then used

	CI side	Age (years)	CI experience (years)	Etiology	Onset	CI type
S1	left	54	11	Genetic	Congenital	Pulsar
	right		9			Sonata
S2	right	57	9	Infection	Postlingual progressive	Pulsar
S3	left	32	11	Radiation	Postlingual progressive	Pulsar
S4	left	53	4	Genetic	Postlingual progressive	Concerto

Table 3.1: Biographical details of participating subjects.
For all, stimuli were presented on electrodes 3 and 10.

to find the maximum acceptable levels (MAL). This was defined as the maximum level of stimulation that is perceived as very loud, but not uncomfortable or painful.

From those measurements, we calculated both the dynamic range from the quotient of MAL and THR, and a measure of TI, namely the slope of the THR-duration curve in logarithmic units, as well as the critical duration after which slopes stop decaying with increasing duration.

3.3 Results

Because the output of the implant is limited to a current amplitude of 1200 CU, some data points could not be measured. This was more often the case for measurements with the lower stimulation rate and those in the more basal electrode, because these factors lead in general to higher amplitudes of THR and MAL.

The phase length used in this experiment (15 µs) is shorter than in standard clinical settings. This was necessary in order to allow for the high stimulation rate, but it leads to less efficient pulses. A reduction of the phase length requires a corresponding increase in the simulation amplitude, which is approximately linear for monopolar stimulation (Bonnet et al., 2012) – this is equivalent to keeping the charge constant.[1]

This allows us to more easily compare them to the rest of the dataset.

3.3.1 All data

Figure 3.1 shows a summary of the measured data. THR and MAL are shown as a function of the number of pulses in a train.

For all subjects and electrodes tested, the pulse-integration curve decays more steeply for the 25 kpps stimulation rate before flattening out – for the same number of pulses, lower rates lead to lower thresholds and maximum acceptable levels. In comparison, there is no definite effect of the electrode in the integration curves.

By analysing the amplitudes as a function of the train duration for the different conditions, the effect of the stimulation rate becomes clearer. This is shown in Figure 3.2. It is not possible to unambiguously assign a train-duration to a stimulation with single pulses. Considering that the change of one to two pulses consists in a doubling of charge, they were positioned as having an equivalent duration equal to one over the stimulation rate. Overlaid is the schematic shape for TI in normal hearing, with a slope of $-20/3$ dB per decade (Heil, Matysiak, and Neubauer, 2017), and a critical duration of 200 ms (Scharf, 1978). The general shape of the THR-duration curve of the CI users is much closer to the NH case for the higher rate, especially due to more similar slopes. The decrease of amplitudes with the length of the pulse trains is clear for durations of up to around 100 ms. After this point, the results vary between different subjects and electrodes.

Thresholds

Threshold values at the maximum train duration of 1000 ms decreased between 1200 and 25 000 pps in all cases. The median change was 12.69 dB, which is equivalent to a decrease of 2.90 dB per doubling of rate. In addition, lower thresholds were generally found on the apical electrode, but the drop in THR between the two stimulation rates did not change significantly between the electrodes.

For some of the subjects and conditions, the slope of the integration curves flattened out with an increasing number of pulses – this behaviour is for example clear for subject S1l. However, this was not always the case, as with subject S3, for whom the THR dropped with more or less constant slopes up to the highest duration (see Figure 3.5).

3.3.2 Dynamic range

The dynamic range could not be estimated for all electrodes, especially at 1200 pps, due to the inefficacy of the short phase lengths. This either caused the MAL to occur beyond the limit of 1200 CU or the necessary output voltage fell beyond the compliance limit of the electrode.

For the rest of the electrodes, the mean dynamic range at 500 and 1000 ms was calculated in dB from the ratio between median THR and MAL, under the assumption that these do not change much beyond 200 ms. It is important to note that the DRs at the lower rate might be underestimated for the group as a whole, because the missing data at the lower rates might actually be caused by larger DR values.

Figure 3.3 shows the dynamic range for the measured electrodes and stimulation rates. As it was expected, increasing the pulse rate from 1200 pps to the higher rate of 25 000 pps increased the dynamic range between MAL and THR. This increase had a median of 9.10 dB for the electrodes where both DRs could be measured. This is equivalent to an increase of 2.08 dB / doubling of rate.

3.3.3 Integration slope

If the THR-duration curves are plotted on a double-logarithmic scale (see Figure 3.1), most of the curves tend to fall on a straight line in a large region, especially for the lower rate. There are, however, systematic deviations from this shape at long durations. The slope often decreases towards the longest pulse-trains, and in some cases, the THR value flattens well before 500 ms.

A second effect that was more rarely seen in the THR-duration curves was a flattening of the curve

[1] This effect is less pronounced in bipolar simulation, so that a smaller increase in amplitude is necessary (Chatterjee, Fu, and Shannon, 2000; Zeng, Galvin III, and Zhang, 1998)

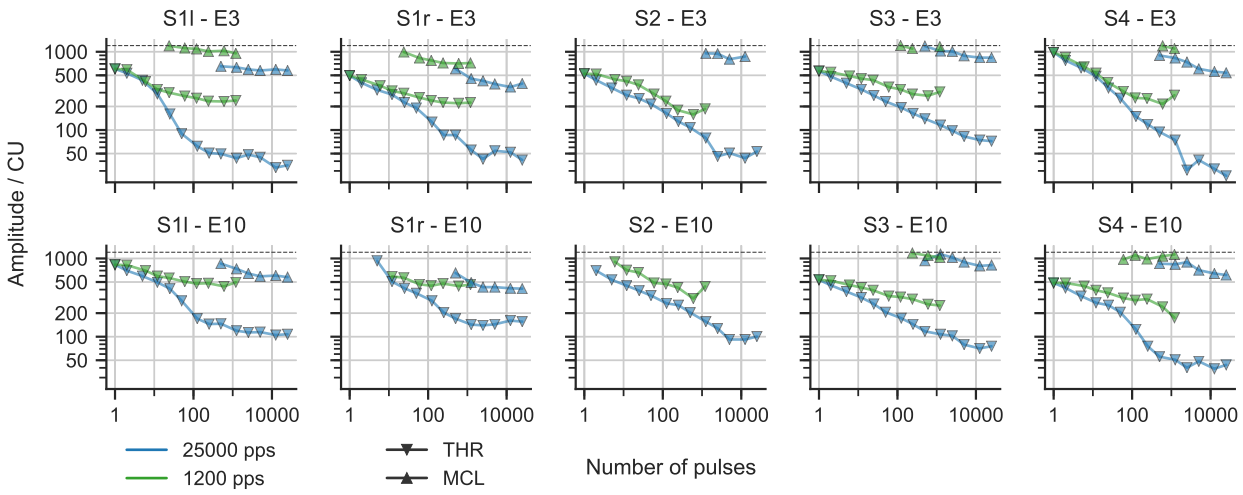


FIGURE 3.1: Median THR as a function of the number of pulses for the different electrode/rate combinations for all subjects, in a double-logarithmic scale.

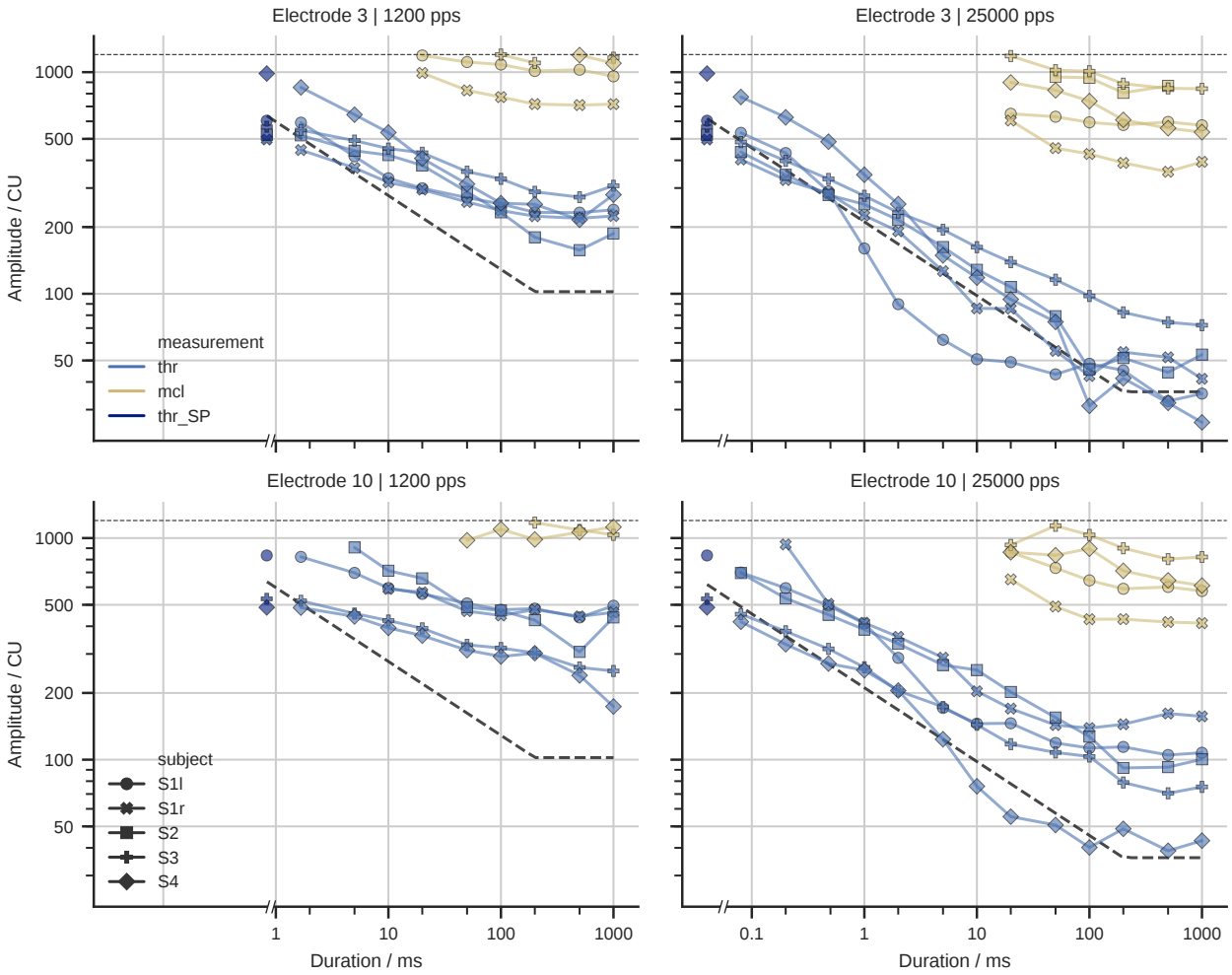


FIGURE 3.2: Median amplitudes for each subject. The black dashed line represents the hardware limit of 1200 CU. In dark blue are single pulse (SP) THR measurements at an equivalent duration of 1/rate. The dark dashed line schematically represents the curve shape for NH, with a slope of −20/3 dB per decade, and a critical duration of 200 ms.

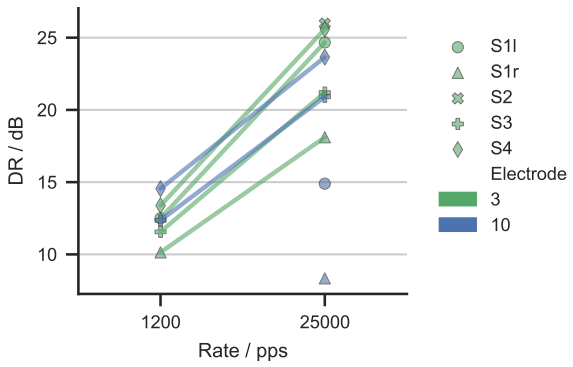


FIGURE 3.3: Mean dynamic range for train durations of 500 and 1000 ms at the two stimulation rates and across electrodes (E3 in green, E10 in blue) and subjects. The median change was 9.15 dB.

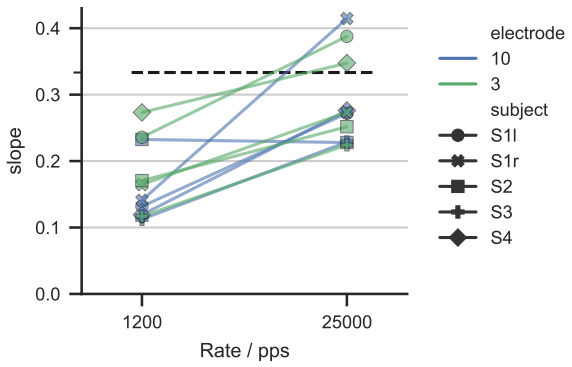


FIGURE 3.4: Fitted slope parameters m. The dashed line represents a slope of 1/3, the usual value found for normal hearing (Heil, Matysiak, and Neubauer, 2017).

towards low numbers of pulses (e.g. S2-E3 or S4-E10). These two effects together gave some of the curves a rather sigmoidal shape. This effect is more pronounced for the high stimulation rate, e.g. in the case of S1l. Since it occurred rarely, the latter effect was ignored in the following analysis.

In order to give a quantitative description of the measurements, the median amplitudes were fitted with a broken-stick curve consisting of a power-law decrease up to a specific critical number of pulses N_C, after which the THR did not decrease further with increasing pulse train length (see examples on Figure 3.5):

$$I(n) = \begin{cases} I_\infty \cdot \left(\frac{n}{N_C}\right)^{-m} & n < N_C \\ I_\infty & n \geq N_C \end{cases} \tag{3.1}$$

Here, I is the threshold amplitude in CU, and n the number of pulses in a train. The thresholds for single pulses were included in the curves for both rates. I_∞ is a multiplicative factor that represents the asymptotic THR for stimulation with long pulse trains. The fits were done using a least-squares method, using the medians of the amplitude values.

The fitted value of the slope parameter m gives a measure of the magnitude of the integration for the different conditions, while $T_C = \text{rate} \cdot N_C$ estimates the critical duration after which the THR-values do not increase. In general, these curves give good fits for the measurements (median $r^2 = 0.977$).

We conducted a two-way ANOVA of the slope of the THR-duration curve, with electrode and rate as fixed factors and subjects as random factors. It revealed a main effect of stimulation rate $F(1, 12) = 18.96$, $p = 0.0009$ (see Figure 3.4).

This means that the THR-curve was significantly steeper for the high than for lower rate: $m_{1200} = 0.153$, compared to $m_{25000} = 0.273$ (medians), indicating a stronger integration of pulses at the high rate. Neither the electrode effect nor the interaction between rate and electrode were significant ($p > .05$).

A similar ANOVA was conducted on the critical duration T_C. Interestingly, there are no significant effects of electrode, stimulation rate or their interaction.

This indicates that the large range of observed values is based on across-subject differences, without systematic changes caused by the stimulation parameters. Figure 3.5 shows two rather extreme examples with very distinct results for T_C. In the first case, the THR values do not fall much beyond 50 ms, and the fitting results in $T_C = 52.8$ ms, whereas in the second case the curve keeps falling almost up to the maximum duration tested, so $T_C = 1000$ ms, and in some cases, no critical duration was found at all. In general, both ears of Subject 1 consistently showed much lower values of T_C than the other three subjects, despite them not showing a particularly different performance.

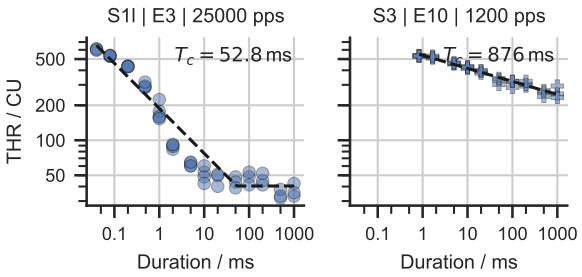


FIGURE 3.5: Exemplary THR data resulting in extremal values of T_C.

Despite the large dispersion in observed values of T_C (see Figure 3.7), there is an internal consistency. There is a significant correlation in the critical durations at the low and high rates ($\rho = 0.755$, $p = 0.007$, Spearman's Rho). This is shown in Figure 3.6 for the fits where a critical duration was found in both rates.

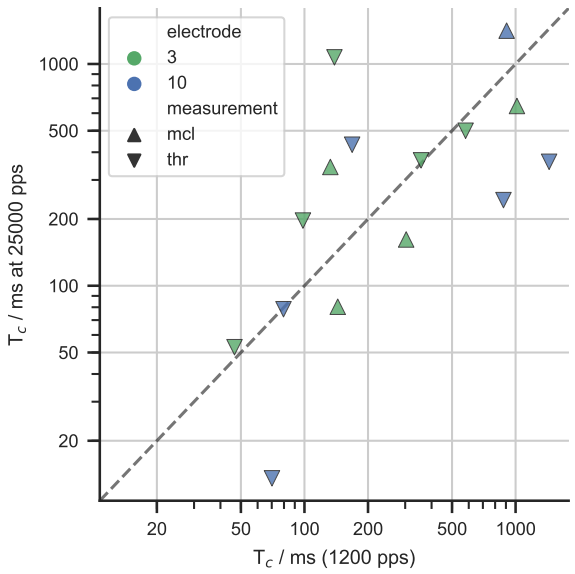


FIGURE 3.6: Correlation between the fitted critical durations at the lower and higher rate for the different electrodes and levels.

3.4 Discussion

In this exploratory analysis of electrical stimulation with very high rates, we investigated the available dynamic range and the slope of temporal integration of threshold. For this, we fitted the amplitude curves to a broken-stick function with a power-law domain and a constant domain.

3.4.1 THR and Dynamic Range

As it was expected, both THR and MAL levels decayed with the increase of the stimulation rate.

The median change in THR was 12.69 dB, which is equivalent to 2.90 dB per doubling of rate. Additionally, the decrease of THR was stronger than of MAL, in those electrodes where it could be measured. This lead to an increase in the dynamic range, equivalent to 2.09 dB/ doubling of rate in the electrodes for which we could gather this data.

Previous experiments have also found increases in DR as a function of stimulation rate (although not to such high pulse rates). However, the increases were not as high as in the present data. In comparison, Bonnet et al. (2012) found a decrease in THR of 2.11 dB per doubling of pulse rate up to 3868 pps. In addition, they found DR increases of 1.30 dB per doubling of pulse rate for ranges up to 3868 pps.

Zhou, Xu, and Pfingst (2012) found an increase in DR of 1.28 dB per doubling of pulse rate for rates up to 5000 pps in monopolar stimulation. Their measurements in humans for rates above 1000 pps show a decrease in THR of around 2.6 dB per doubling of rate (see Fig. 6 in their paper). Our rate of THR-decrease also seems to be comparable to or larger than their result, if extended to 25 000 pps.

Considering the small size of the data set and that MAL measurements were capped at 1200 CU so that we might be ignoring subjects with particularly high DRs at 1200 pps, we cannot easily generalize these data to the general population. Nevertheless, in the cases where data was available, the increase was relatively homogeneous despite the large absolute differences in DR (Figure 3.3). More research with a higher number of subjects and with intermediate pulse rates is necessary to clarify this.

3.4.2 Fitting of amplitude-duration function

In order to quantify the temporal integration, the THR-duration curves were fitted with a function of three parameters. No effects of the stimulating electrode could be found on any of the parameters, and neither did it interact with the stimulation rate. This indicates that no systematic effect of the location of stimulation could be observed in our dataset. If there is one, we could not detect it due to its small size and/or insufficient power.

On the other hand, an increase of the stimulation rate had a clear effect by decreasing the parameter I_∞ on all subjects and electrodes, reflecting the known effect of a higher rate reducing THRs. The rate also had an effect on the slope of integration m: the high rate corresponded to an increase in all subjects and electrodes. Only in one case (S2, E10) the slope was similar. A higher stimulation rate also lead to a decrease of THR and MAL for the same number of pulse trains.

The third parameter, the critical duration T_C after which the THR does not further decrease, showed a large variability. However, in contrast to the values of m and I_∞, it did not change consistently with stimulation rate, and the variability mostly corresponded to the different subjects. The median T_C ranged from 66 ms in the case of S1l to 717 ms in the case of S2, with some values lying beyond 1 s (medians).

Taken together, these results can be interpreted as an indication that the general THR-levels and the magnitude of integration (indicated by the parameters I_∞ and m) reflect lower-level processes in the auditory system and are thus affected similarly by a change in stimulation rate for all subjects.

In comparison, the critical duration T_C might follow a higher-level integration process further from the auditory nerve, thus explaining both the inconsistent effect of the stimulation rate and the high variability across subjects (see Figure 3.7). In the present data, the lowest values were observed in both ears of S1, whose hearing loss started prelingually, as opposed to postlingual onsets for the other subjects, who all had higher of values of T_C spreading over a wider range, sometimes beyond 1 s.

When interpreted based on models with a long integration process, the value of the critical duration closely matches the time constant of the integrator. Smaller values of T_C could indicate a faster detection mechanism, and it could then be expected that these

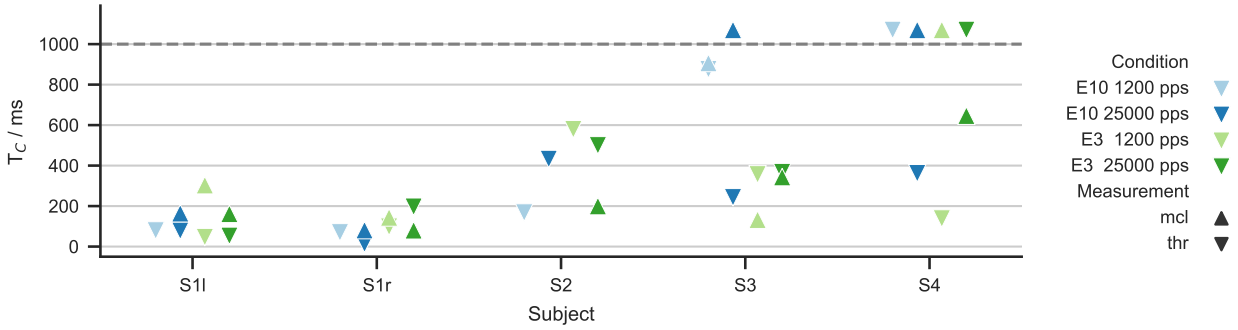


FIGURE 3.7: Distribution of T_C across subjects. Here, fit values over 1000 ms are clipped and shown above the dashed line; these are the cases without a critical duration. Missing datapoints represent conditions without enough measurable MALs.

subjects, for example, would fare better in the detection of fast amplitude modulations. On the other hand, in multiple-looks models with memory, a reduced T_C would instead reflect a decreased capability of the short-term memory process.

3.5 Conclusions

We ran a psychophysical experiment with CI-users and measured THR and MAL as a function of duration (TI) for two different rates. In general, TI is much more shallow for CI users than in normal hearing, especially at moderate rates. Only at very high rates, the slope of integration approached normal values. In addition the increase in rate lead to an increased electrical dynamic range in a manner consistent with previous literature. The critical duration until which integration occurred seems to vary strongly across electrodes and with rate, but in a highly inconsistent manner for the different subjects.

Acknowledgements

We would like to thank the participants for their contributions to this work, and Prof. Armin Giebel for helping with the supervision. Dr. Jörg Encke offered helpful critique regarding methodical questions, and provided a program which built the basis of the software used for data collection. The Institut für Ionenphysik und Angewandte Physik of the University of Innsbruck provided the RIB2 for direct stimulation.

Bibliography

Bonnet, Raymond M; Boermans, Peter-Paul BM; Avenarius, Otto F; Briaire, Jeroen J, and Frijns, Johan HM (2012). "Effects of Pulse Width, Pulse Rate and Paired Electrode Stimulation on Psychophysical Measures of Dynamic Range and Speech Recognition in Cochlear Implants". In: *Ear and hearing* 33.4, pp. 489–496 (cit. on pp. 28, 30, 33).

Carlyon, Robert P.; Van Wieringen, Astrid; Deeks, John M.; Long, Christopher J.; Lyzenga, Johannes, and Wouters, Jan (2005). "Effect of Inter-Phase Gap on the Sensitivity of Cochlear Implant Users to Electrical Stimulation". In: *Hearing research* 205.1-2, pp. 210–224 (cit. on pp. 28 sq.).

Chatterjee, Monita; Fu, Qian-Jie, and Shannon, Robert V. (2000). "Effects of Phase Duration and Electrode Separation on Loudness Growth in Cochlear Implant Listeners". In: *The Journal of the Acoustical Society of America* 107.3, pp. 1637–1644 (cit. on p. 30).

Donaldson, Gail S.; Viemeister, Neal F., and Nelson, David A. (1997). "Psychometric Functions and Temporal Integration in Electric Hearing". In: *Journal of the Acoustical Society of America* 101.6, pp. 3706–3721 (cit. on p. 28).

Eddington, D. K.; Dobelle, Wm H.; Mladejovsky, M. G.; Brackmann, D. E., and Parkin, J. L. (1978). "Auditory Prostheses Research with Multiple Channel Intracochlear Stimulation in Man". In: *Annals of Otology, Rhinology & Laryngology* 87 (6_suppl), pp. 5–39 (cit. on p. 28).

Heil, Peter; Matysiak, Artur, and Neubauer, Heinrich (2017). "A Probabilistic Poisson-Based Model Accounts for an Extensive Set of Absolute Auditory Threshold Measurements". In: *Hearing research* 353, pp. 135–161 (cit. on pp. 30, 32).

Heil, Peter and Neubauer, Heinrich (2003). "A Unifying Basis of Auditory Thresholds Based on Temporal Summation". In: *Proceedings of the National Academy of Sciences* 100.10, pp. 6151–6156 (cit. on p. 28).

Hughes, Michelle L.; Castioni, Erin E.; Goehring, Jenny L., and Baudhuin, Jacquelyn L. (2012). "Temporal Response Properties of the Auditory Nerve: Data from Human Cochlear-Implant Recipients". In: *Hearing Research* 285.1–2, pp. 46–57 (cit. on p. 28).

Ifukube, Tohru and White, Robert L (1987). "Current Distributions Produced inside and Outside the Cochlea from a Scala Tympani Electrode Array". In: *IEEE transactions on biomedical engineering* 34.11, pp. 883–890 (cit. on p. 27).

Javel, Eric and Shepherd, Robert K. (2000). "Electrical Stimulation of the Auditory Nerve: III. Response Initiation Sites and Temporal Fine Structure". In: *Hearing Research* 140.1, pp. 45–76 (cit. on p. 28).

Kral, Andrej; Hartmann, Rainer; Mortazavi, Dariusch, and Klinke, Rainer (1998). "Spatial Resolution of Cochlear Implants: The Electrical Field and Excitation of Auditory Afferents". In: *Hearing research* 121.1-2, pp. 11–28 (cit. on p. 27).

Kreft, Heather A.; Donaldson, Gail S., and Nelson, David A. (2004). "Effects of Pulse Rate on Threshold and Dynamic Range in Clarion Cochlear-Implant Users (L)". In: *The Journal of the Acoustical Society of America* 115.5, pp. 1885–1888 (cit. on p. 28).

Litvak, Leonid M.; Smith, Zachary M.; Delgutte, Bertrand, and Eddington, Donald K. (2003). "Desynchronization of Electrically Evoked Auditory-Nerve Activity by High-Frequency Pulse Trains of Long Duration". In: *The Journal of the Acoustical Society of America* 114.4, pp. 2066–2078 (cit. on p. 28).

McKay, Colette M.; Lim, Hubert H., and Lenarz, Thomas (2013). "Temporal Processing in the Auditory System". In: *Journal of the Association for Research in Otolaryngology* 14.1, pp. 103–124 (cit. on p. 29).

McKay, Colette M. and McDermott, Hugh J. (1998). "Loudness Perception with Pulsatile Electrical Stimulation: The Effect of Interpulse Intervals". In: *The Journal of the Acoustical Society of America* 104.2, pp. 1061–1074 (cit. on pp. 28 sq.).

Meddis, Ray and Lecluyse, Wendy (2011). "The Psychophysics of Absolute Threshold and Signal Duration: A Probabilistic Approach". In: *The Journal of the Acoustical Society of America* 129.5, pp. 3153–3165 (cit. on p. 28).

[p)] Obando, Miguel; Schwanda, Daniela, and Hemmert, Werner (2018). "Pulse Train Integration in CI-Wearers and Effects of High Stimulation Rates". In: *41st ARO Annual MidWinter Meeting*. 41st ARO Annual MidWinter Meeting. Vol. 41. San Diego, California, USA, p. 34 (cit. on p. 27).

Penner, Merrilynn J. (1978). "A Power Law Transformation Resulting in a Class of Short-Term Integrators That Produce Time-Intensity Trades for Noise Bursts". In: *The Journal of the Acoustical Society of America* 63.1, pp. 195–202 (cit. on p. 28).

Plomp, R. and Bouman, M. A. (1959). "Relation between Hearing Threshold and Duration for Tone Pulses". In: *The Journal of the Acoustical Society of America* 31.6, pp. 749–758 (cit. on p. 28).

Rubinstein, J. T; Wilson, B. S; Finley, C. C, and Abbas, P. J (1999). "Pseudospontaneous Activity: Stochastic Independence of Auditory Nerve Fibers with Electrical Stimulation". In: *Hearing Research* 127.1, pp. 108–118 (cit. on p. 28).

Scharf, Bertram (1978). "Loudness". In: *Hearing*. Ed. by E. C. Carterette and M. P. Friedman. Vol. 4. The Handbook of Perception. New York: Academic Press, pp. 187–242 (cit. on p. 30).

Shannon, Robert V. (1983). "Multichannel Electrical Stimulation of the Auditory Nerve in Man. I. Basic Psychophysics". In: *Hearing Research* 11.2, pp. 157–189 (cit. on p. 28).

– (1985). "Threshold and Loudness Functions for Pulsatile Stimulation of Cochlear Implants". In: *Hearing Research* 18.2, pp. 135–143 (cit. on p. 28).

– (1989). "A Model of Threshold for Pulsatile Electrical Stimulation of Cochlear Implants". In: *Hearing Research* 40.3, pp. 197–204 (cit. on pp. 28 sq.).

Skinner, Margaret W.; Holden, Laura K.; Holden, Timothy A., and Demorest, Marilyn E. (2000). "Effect of Stimulation Rate on Cochlear Implant Recipients' Thresholds and Maximum Acceptable Loudness Levels". In: *Journal of the American Academy of Audiology* 11.4, pp. 203–213 (cit. on p. 28).

Tang, Qing; Benítez, Raul, and Zeng, Fan-Gang (2011). "Spatial Channel Interactions in Cochlear Implants". In: *Journal of neural engineering* 8.4, p. 046029 (cit. on p. 27).

Verhey, Jesko L. (2010). "Temporal Resolution and Temporal Integration". In: *The Oxford Handbook of Auditory Science: Hearing*. Vol. 3. The Oxford Handbook of Auditory Science. Oxford: Oxford University Press, pp. 105–121 (cit. on p. 28).

Viemeister, Neal F. and Wakefield, Gregory H. (1991). "Temporal Integration and Multiple Looks". In: *The Journal of the Acoustical Society of America* 90.2, pp. 858–865 (cit. on p. 28).

Wilson, Blake S.; Finley, Charles C.; Lawson, Dewey T., and Zerbi, Mariangeli (1997). "Temporal Representations with Cochlear Implants." In: *Otology & Neurotology* 18.6, S30–S534 (cit. on p. 28).

Zeng, Fan-Gang; Galvin III, John J., and Zhang, Chaoying (1998). "Encoding Loudness by Electric Stimulation of the Auditory Nerve". In: *Neuroreport* 9.8, pp. 1845–1848 (cit. on p. 30).

Zhou, Ning; Kraft, Casey T.; Colesa, Deborah J., and Pfingst, Bryan E. (2015). "Integration of Pulse Trains in Humans and Guinea Pigs with Cochlear Implants". In: *Journal of the Association for Research in Otolaryngology* 16.4, pp. 523–534 (cit. on p. 28).

Zhou, Ning; Xu, Li, and Pfingst, Bryan E. (2012). "Characteristics of Detection Thresholds and Maximum Comfortable Loudness Levels as a Function of Pulse Rate in Human Cochlear Implant Users". In: *Hearing Research* 284.1, pp. 25–32 (cit. on pp. 28, 33).

Zwislocki, Josef J. (1969). "Temporal Summation of Loudness: An Analysis". In: *The Journal of the Acoustical Society of America* 46 (2B), pp. 431–441 (cit. on p. 28).

4 – On the Effect of High Stimulation Rates on Temporal Integration in Cochlear Implant Users

Miguel Obando,[1,2] Anna Dietze,[1] and Werner Hemmert[1,2]
[1)]*Graduate School of Systemic Neurosciences, Ludwig-Maximilian Universität München*
[2)]*Bio-inspired Information Processing, Department of Electrical Engineering, Technical University of Munich*

ABSTRACT

In this study, the effect of high stimulation rates on temporal integration in hearing with cochlear implants was investigated. Threshold amplitudes, maximum acceptable levels and a line of equal loudness in 11 cochlear implant users (all with implants from MED-EL) were obtained. The measurements were done (a) with a clinically used single channel stimulation rate of 1500 pps and (b) at a high stimulation rate of 18 000 pps, both for an apical electrode and a basal electrode. The length of the stimulation pulse trains varied from a single pulse to 300 ms.

We found decreased amplitudes for the high rate in all measurements. A power-law like function was used to fit the obtained amplitudes of the three measurements for individual subjects with a high accuracy ($R^2 = 0.92 \pm 0.16$). Threshold slopes of -3.44 dB and -5.43 dB per tenfold increase in duration for the low and high rate respectively were found. The change of the stimulation electrode did not cause any systematic effects regarding threshold amplitudes of the temporal integration (TI) curve, but appeared to alter the slopes of the TI curves. The DR was increased with the high rate by 6.58 ± 3.79 dB.

Overall, with the high stimulation rate decreased amplitudes and steeper TI curves were observed, beyond those values reported by previous studies using lower stimulation rates.

INTRODUCTION

Signal processing in cochlear implants (CIs) assumes a close relation of stimulation amplitude and perceived loudness. Understanding the exact mechanisms and the influences of different parameters will help to find a superior approach to loudness coding. The aim of this study was to investigate in which way temporal integration in CIs is affected by different stimulation rates at different stimulation levels.

Temporal integration (TI) describes the observation that detection thresholds decrease with an increase in stimulus duration, or alternatively, the necessary reduction of suprathreshold stimuli of increasing duration in order to maintain equal loudness. Early studies already investigated this duration-intensity reciprocity in acoustic hearing. According to Stevens and Hall [30], when keeping the stimulation level constant, loudness grows following a power function of duration. They noted that the perceived loudness increases only up to a critical duration of about 150 ms, from where on loudness is independent of duration.

McFadden [19] did measurements on the loudness of stimuli that differed in sound pressure level and duration. He found that, to maintain equal loudness, intensity must decrease by between 3 and 15 dB for each doubling of duration, depending upon the subject.

Gerken et al. [11] reviewed previous studies of temporal integration, and came to the conclusion that the data is best described as following a power-law function, with an exponent that falls with increase in frequency. For this, however, the definition of duration needs to be modified, since the rise and fall of signals played an under-proportional role in TI.

By now, this kind of measurement has been done with many species, including humans, primates, carnivores, birds and even fish (reviewed by Heil et al. [14]). The summary revealed – besides inter-individual differences – striking similarities in the overall shape of these curves. A power-law function describes the relationship of threshold and time very well for acoustic hearing. This relationship has a slope of about -2 dB per doubling of duration or approximately $-20/3$ dB per tenfold increase of duration (decade), again with a specific definition of loudness (from the beginning of the onset to the end of the offset).

While it is assumed that the same central processing takes place in acoustic and electric hearing, only with different neural inputs, TI curves are much shallower with electric stimulation (see e.g. Donaldson et al. [6], Gerken et al. [12]).

In addition, when keeping the duration fixed, an increase in stimulation rate also reduces electric detection thresholds (multi-pulse integration, see Pfingst et al. [23], Zhou et al. [35]).

Nerve cells with processes close to an electrode contact might still register electric pulses by other electrodes (particularly of those in immediate vicinity), especially with monopolar, sequential stimulation. This assumption is supported by studies that show a broad current spread throughout the whole cochlea (e.g. Ifukube and White [16], Malherbe et al. [18]). Therefore, the actual stimulation rate of a single neuron might be up to the repetition frequency of a single electrode multiplied by the number of stimulation contacts in use. The single channel stimulation rate is usually set to approximately 1500 pps in MED-EL implants, in which 12 electrodes are available. This would lead to a maximum global stimulation rate of about 18 000 pps, under the assumption that a stimulation of any electrode contact acts on each cochlear place.

The term temporal integration might be misleading, since it is not clear if some quality of the stimulus is really integrated, as this would require some kind of computation (Viemeister and Wakefield [31]). Also, the slope of the above mentioned functions would have to be steeper (-10 dB per decade) in case of perfect integration of intensity. Consequently, there have been different attempts to describe and explain the amplitude vs. duration relationship, including models by Shannon [27], Carlyon et al. [5], and McKay and McDermott [21]. In

general, it is known that the slopes of TI are shallower than those from acoustic hearing, with high variability across subjects, and in some cases across electrodes in a single subject's cochlea (Donaldson et al.[6], Shannon[28]).

For the present experiment, we were interested in the characteristics of temporal integration in CIs, as different mechanisms are in place than in normal hearing, as well as the effect of stimulation rate with clinically relevant rates. Additionally, we asked to which extent a simple power-law function is suitable to model electric hearing data up to moderate durations, and in particular how TI is affected by different stimulation rates and stimulation levels. To this end, in addition to thresholds (THR), we also measured maximum acceptable levels (MAL) and a line of equal loudness (BAL) between THR and MAL, as a function of stimulus duration. To our knowledge, TI functions have not been obtained for stimulation rates as high as the one used in our experiments.

METHODS

Experimental Procedure

The experiment consisted of three parts. In the first part, current amplitudes corresponding to threshold (THR) were determined for different combinations of electrode, stimulation rate and duration, with four repetitions each. For the second part, subjects adjusted the stimulus levels to the maximum acceptable amplitude, referred to as MAL. The third part consisted of a loudness balancing procedure (BAL). Participants matched the perceived loudness of a probe stimulus to a reference stimulus with a longer duration in all of the four electrode $\times$ rate combinations. A method of adjustment was used, where participants increased and decreased the stimulation amplitude to the desired value by themselves (for more information on this method see Gelfand[10]).

Participants

Eleven subjects ($M = 56$ years, $SD = 14$ years; 8 female) with cochlear implants from MED-EL participated in our study. Two women (subjects S1 and S8) agreed to do the experiments with both ears, for a total of 13 measured ears. Details are presented in Table I. Subject S11 perceived onset and offset artefacts of the stimulation pulses at the higher rate with both tested electrodes, even for pulse amplitudes of zero CU. This participant was excluded from the analyses.

All subjects gave their informed written consent for the participation and received monetary compensation. Measurements were conducted in accordance to the Declaration of Helsinki, and were approved by the medical ethics committee of the Klinikum rechts der Isar (Munich, 2126/08).

Equipment

The core of the experimental setup was a computer equipped with a digital card (Model NI PCIe-6361, National Instruments, Austin, Texas, USA). Pulse trains were created by sending all parameters (phase and gap duration, stimulation rate and duration of the pulse train, as well as the stimulation amplitude) to the Research Interface Box 2 (RIB2, Institute of Ion Physics and Applied Physics, University of Innsbruck). It turns the given information into pulses that are then sent out directly to the implanted parts of a CI. With this we had full control over the stimulation, bypassing the sound processor. Stimuli were checked with a detector box (Institute of Ion Physics and Applied Physics, University of Innsbruck), which emulates a pulsar implant, and was connected to a digital oscilloscope.

Participants responded to the stimuli either by using a computer mouse to click response buttons of a graphical user interface on the computer screen or by pressing the respective buttons on a computer keyboard.

To create the stimulation pulses and to adapt them in real time corresponding to the participants responses, Python (Version 2.7, 32-bit) was used. The scripts were written by members of the research group. For data analysis, MATLAB with the Curve Fitting Toolbox (Version 9.5.0.944444 (R2018b), The MathWorks Inc., Natick, Massachusetts) and R (Version 1.2.747 (2018), R Foundation for Statistical Computing, Vienna, Austria) with various packages, including WRS (Version 0.35 (2018), Wilcox & Schönbrodt) for robust statistics were used.

Stimuli

For all measurements, monopolar stimulation was applied. Biphasic, charge-balanced pulses were used with the cathodic (negative) phase leading. To allow for very high stimulation rates, the phase duration was chosen to be 23.33 µs only, with a minimally available gap of 2.1 µs between the phases. A stimulation amplitude of 1200 CU (62 dB re 1 CU) is set as an upper limit for stimulation by both the RIB2 software and the CI. 1 CU is roughly equivalent to 1 µA.

Stimuli differed in stimulation electrode, rate and the number of pulses.

All stimuli of one measurement were presented in randomised order. This does not apply to any of the preliminary measurements. Randomisation was also limited to a certain amount in the loudness balancing task. In all experiments, the presented pulse trains were separated by a fixed silent gap of 500 ms. Since the longest duration used was 300 ms, the effect of forward masking was presumably reduced to a minimum after the pause. Nelson and Donaldson[22] found average time constants of 54 ms for the exponential decay of masking after a 320 ms long masker with a frequency of 500 Hz. Even with the largest time constant they found (163ms), less than 5 % threshold shift would be observed for a 10 or 30 ms probe pulse train after a pause of 500 ms. Regarding the effect of

Table I. Biographical data and etiology of participants; (HL: Hearing loss).

Subject	Age (years)	Gender	HL onset (years)	Etiology	CI use (months)	CI Model
S1r	65	F	40	unknown	60	Sonata
S1l	65	F	40	unknown	54	Sonata
S2	23	M	1	Meningitis	120	Sonata
S3	53	F	30	hereditary	28	Synchrony
S4	78	M	58	acute HL	148	Pulsar
S5	42	F	1	Cholesteatoma	36	Synchrony
S6	55	M	5	Otitis	72	Concerto
S7	42	F	birth	acute HL	60	Concerto
S8r	64	F	35	Meningitis	90	Concerto
S8l	64	F	27	Meningitis	30	Synchrony
S9	60	F	9	Otitis	72	Concerto
S10	59	F	30	unknown	132	Pulsar
S11	56	F	birth	unknown	144	Pulsar

stimulation rate on forward masking time constants, no studies investigated rates as high as the ones we use in this study. However, Adel et al. [1] showed that masking by pulse trains of high stimulation rate was even less pronounced than masking induced by low-rate pulse train maskers, when presented at the same loudness, as measured through the shift in probe detection thresholds 16 ms after the masker offset.

Electrodes. The same measurements were done at two different electrodes on the array. If not hindered by any reason, electrodes 3 (apical) and 10 (basal) were selected out of 12 electrodes in MED-EL CIs. Otherwise, neighbouring electrodes were chosen.

Stimulation Rates. Two stimulation rates (the inverse of the distance between the starting points of following pulses in a pulse train) were used. The lower stimulation rate (1500 pps) represents a typical stimulation rate present at a single contact of the electrode array in normal CI settings (without channel crosstalk), whereas the higher rate (18 000 pps) was chosen to investigate the effect of an increased overall effective stimulation rate of 12 times the single channel stimulation rate (at 100 % channel crosstalk).

Number of Pulses. Pulse trains with 1 to 5400 pulses were presented at 18 000 pps. At the higher rate, stimuli with the number of pulses ranging from 1 to 5400 were presented. For the lower rate, the stimuli consisted of 1 to 450 pulses. The longest investigated pulses had thus a duration of 300 ms. In total, seven different durations for 18 000 pps and five durations for 1500 pps were used. The single pulse condition was only measured once, and later assigned to both stimulation rates for the analyses.

Thresholds. Prior to threshold measurements, a training phase made sure that the participants understood the task and became familiarised with the setup. After the pure training trials that were not used for the experiment, preliminary threshold estimates were acquired. In this phase, three points (four at the higher rate) in the duration-threshold curve were determined, for each electrode $\times$ rate combination. From these, a first estimation of the threshold was obtained by a linear interpolation in the log-log representation of the duration-threshold pairs. In the training phase, the starting current amplitude was always zero. The resulting estimates were then used in the following measurements with the aim to reduce biases and to save time.

For each parameter combination, the threshold was measured four times. The starting points of the adjustment by the subjects were varied randomly in a range from 80 % to 90 % or 110 % to 120 % of the preliminary estimates mentioned above. Care was taken to ensure that two of the starting points were above and two below the threshold estimate. These variations were unknown to the subject and reduce biases induced by the starting point and by the direction from which the threshold is reached[10].

From the starting points, participants adjusted the perceived loudness by increasing and decreasing the current amplitudes until the stimulus was just barely audible. They were encouraged to use the larger step buttons first to reach the vicinity of the threshold fast, and then use the smaller step changes for fine adjustments. Further, they were asked to bracket their thresholds (i.e. reach them from above and below) before saving the response, which reduces biases as well[10].

In cases where the result was out of range (above 1200 CU) twice, the other two trials of this condition were skipped.

Maximum Acceptable Levels. Again, a short training phase made sure that each participant understood the task. Preliminary MAL measurements were done for the longest pulse train duration (300 ms) for a first estimation of the dynamic range (DR). From just below 50 % of the estimated

DR as starting point, the participants adjusted the perceived loudness by increasing (and decreasing if needed) the current amplitude until the stimulus was very loud, but still acceptable over a longer time. They were encouraged to use the larger steps first to reach the amplitude fast and then use the smaller step changes for fine adjustments to the MAL. If MAL was not reached (before arriving at the limit of 1200 CU) twice, further repetitions of this duration were skipped.

Curve of Equal Loudness. In this part of the experiment, we measured a curve of equal loudness as a function of duration for each of the four electrode × rate conditions. For this experimental part, the graphical interface was slightly extended by two visual displays that flashed in grey and yellow at the time the fixed stimulus and the probe stimulus were presented. This was done to give some indication on which of the two stimuli is fixed (grey) and for which the amplitude can be adjusted by the participant (yellow). Colours were chosen such that they are perceived on the computer screen with approximately the same luminance. In the present experiment, the 300 ms stimuli were used as references for each electrode × rate combination, with the shorter stimuli as probes. Additionally, 300 ms stimuli were balanced against themselves as a measure of consistency. The subjects were instructed to adjust the amplitude of the probe stimulus until its loudness matched the loudness of the fixed stimulus.

Before the comparisons across durations of each electrode × rate condition, all 300 ms-references were first balanced to the loudness of a main reference stimulus presented with 1500 pps, 300 ms at the apical electrode, and which was fixed to 60 % of its DR. The purpose of this balancing step was that all four curves represent an equal loudness, which allows for valid comparisons across different conditions later on. Right afterwards, subjects were presented with all of the four balanced stimuli in a row. As they had been balanced before, all of them should appear at the same loudness. In case this was not true, subjects were able to change the amplitude of individual signals (making them either louder or softer in small step sizes) to equalise them in terms of loudness.

After these steps, the experiment continued with adjusting loudness of two signals of different duration within one electrode × rate condition. In each condition, the perceived loudness of the shorter duration stimuli had to be matched to the previously adjusted 300 ms stimulus. All comparison pairs (stimulus with duration smaller than 300 ms vs. stimulus with duration equal to 300 ms) were compared four times each. In two of the four trials, the reference (the stimulus that is fixed in amplitude) was the 300 ms signal at the previously obtained amplitude. These trials are called non-inverted. In the other two trials, the 300 ms signal served as probe (the stimulus that has to be adjusted) which was compared to the fixed amplitude of the shorter stimuli. These were the so-called inverted trials. This procedure allows a more accurate balancing and has been used by McKay et al. [20] and Adel et al. [1] with

$$L_{\text{probe}} = L_{\text{reference}} + \frac{d_{\text{non-inverted}} + d_{\text{inverted}}}{2} \qquad (1)$$

The formula describes the average of the differences d between amplitudes of short and long stimuli by the two methods added to the amplitude of the long reference stimulus $L_{\text{reference}}$.

In both cases (non-inverted and inverted) the probe stimulus had a starting value which randomly varied above (between +5 to +10 % DR) or below (between −10 to −5 % DR) the estimated level. This procedure was chosen in order to minimize the bias when adjusting to a standard stimulus[25] (chap. 3). In case of the first trials, the best estimation of the probe was to take the same level of the DR that was used for the reference stimulus. In the specific cases where no MALs could be obtained, care was taken to first start from below, so that the second run could start from a louder level.

User Interface

A graphical user interface (GUI) created with PyGObject (Version 3.24.1) was displayed to the participants of our study. Buttons to increase or reduce the stimulation amplitude in large and small steps were visible, as well as a button to save the amplitudes. All responses could also be entered via a computer keyboard. The corresponding keys were colour-matched to those on the screen.

Changing the amplitude in small steps caused an increase or decrease of the current level by 1.18 CU to 9.45 CU, depending on the amplitude. The large steps changed the level by ±18.90 CU, with an exception. For preliminary trials starting at zero amplitude, large steps increased the current amplitude by 28.35 CU, up until the first reversal.

During the training phase of THR and MAL, a representation of chosen amplitudes was visible, which allowed to give users feedback about e.g. bracketing. No feedback regarding the chosen current amplitude was given outside of training. Throughout the experiment pauses were automatically initiated every 20 minutes if participants did not ask for a pause before.

Analyses

Robust statistical methods were chosen for all averages and the statistical analyses of effects on a measure. Instead of the mean of the answers of each participant, the median was used. For averaging over different subjects, trimmed means with 20 % trimming were calculated. A robust version (based on 20 % trimmed means) of a factorial repeated measures ANOVA was applied when calculating the effects of several factors on the results. This amount of trimming is suggested by Rand Wilcox[32] and was confirmed in private correspondence. Concretely, the functions *wwtrim* and *wwwtrim* from the WRS package[33] were used. These functions allow for two- and three-factorial repeated measures analyses (two or three within-subject factors). The outcome of these methods are the test statistic Q and its corresponding p-value. For all statistical tests, an alpha-level of 0.05 was used.

It was not possible to only use robust methods throughout the analyses. Regular standard deviations were calculated to show variability within or between subjects, in order to allow easier comparisons to other publications. In addition, for fitting of THR, MAL and BAL amplitudes to a power-law as a function of the number of pulses, conventional non-linear least squares estimation was used (although there are robust methods available, see Wilcox [32], p. 471-629).

RESULTS

The majority of calculations were done with amplitudes in CU, and only then transformed into dB re 1 CU for visual display and comparisons.

Effects of Duration, Rate and Electrode

Figure 1 shows exemplary results for two participants. It can be seen that the amplitudes decrease with increasing number of pulses. It is also visible that sometimes, for very short durations, no data points for the MAL could be obtained when the stimulation limit of 1200 CU was exceeded. In panel A (S11), the fitted curves of THR and MAL appear to be almost parallel, and the line of equal loudness is shallower than the other two. In contrast, in panel B (S7) MAL and BAL seem to drop similarly for increasing duration, whereas here THR curves differ.

In general, BAL data needs to be analysed with caution as many participants reported to have problems comparing the loudness of short and long stimuli precisely. This becomes visible in a relatively large scatter of BAL data points, especially at the high stimulation rate. It is important to note that with equation 1 two averaged data points are obtained from the four trials of each duration.

Figure 2 provides an overview of all participants' results. The axes above and below the panels display the number of pulses contained in the stimuli, whereas the axes in the middle show the equivalent duration of the stimuli, with duration = pulses/rate.

Each marker represents the median of each participant's amplitudes for each measurement. The solid lines show the trimmed means of all participants' data points. Data points of subject S5 are partly left outside the lower axis limits for clarity, but were still included in the analysis. Just as in the exemplary data (Figure 1), the decrease of amplitudes with increasing duration can be observed across all measurements.

We first ran a robust repeated-measures ANOVA (see Methods) on amplitude with the factors MEASUREMENT, RATE, and ELECTRODE for the duration of 300 ms. There is a significant effect of MEASUREMENT $Q = 12.06, p < 0.001$ (amplitudes for THR < BAL < MAL), as well as significantly lower amplitudes for the higher RATE ($Q = 232.42, p < 0.001$). There is no significant effect of ELECTRODE ($Q = 1.96, p = 0.162$). The interaction term of RATE × MEASUREMENT reached the significance level: $Q = 11.12, p < 0.001$. This effect describes the stronger difference of THR amplitudes to those of the two other measurements at the high rate compared to the low rate, which reflects that the dynamic range increases with stimulation rate.

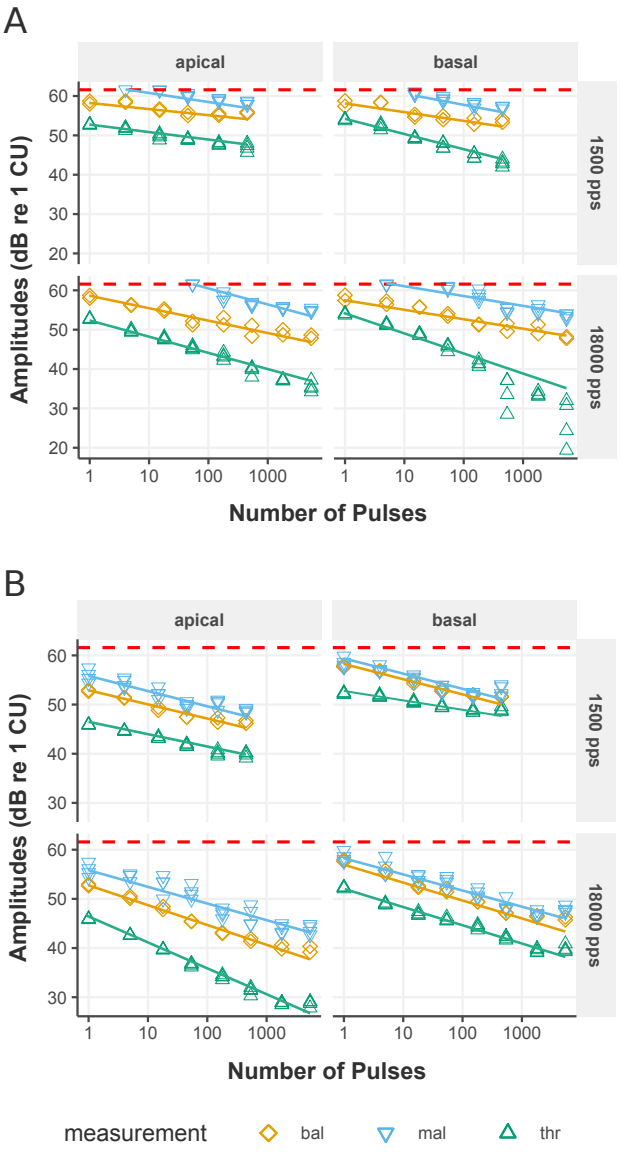


Figure 1. Exemplary data obtained from subject S11 (**A**) and S7 (**B**) for thresholds (green), maximum acceptable levels (blue) and loudness balancing (orange). The solid lines depict power functions fitted to the data points. The dashed red line represents the maximum possible stimulation amplitude of 1200 CU. **Top:** 1500 pps. **Bottom:** 18000 pps. **Left:** Apical electrode. **Right:** Basal electrode.

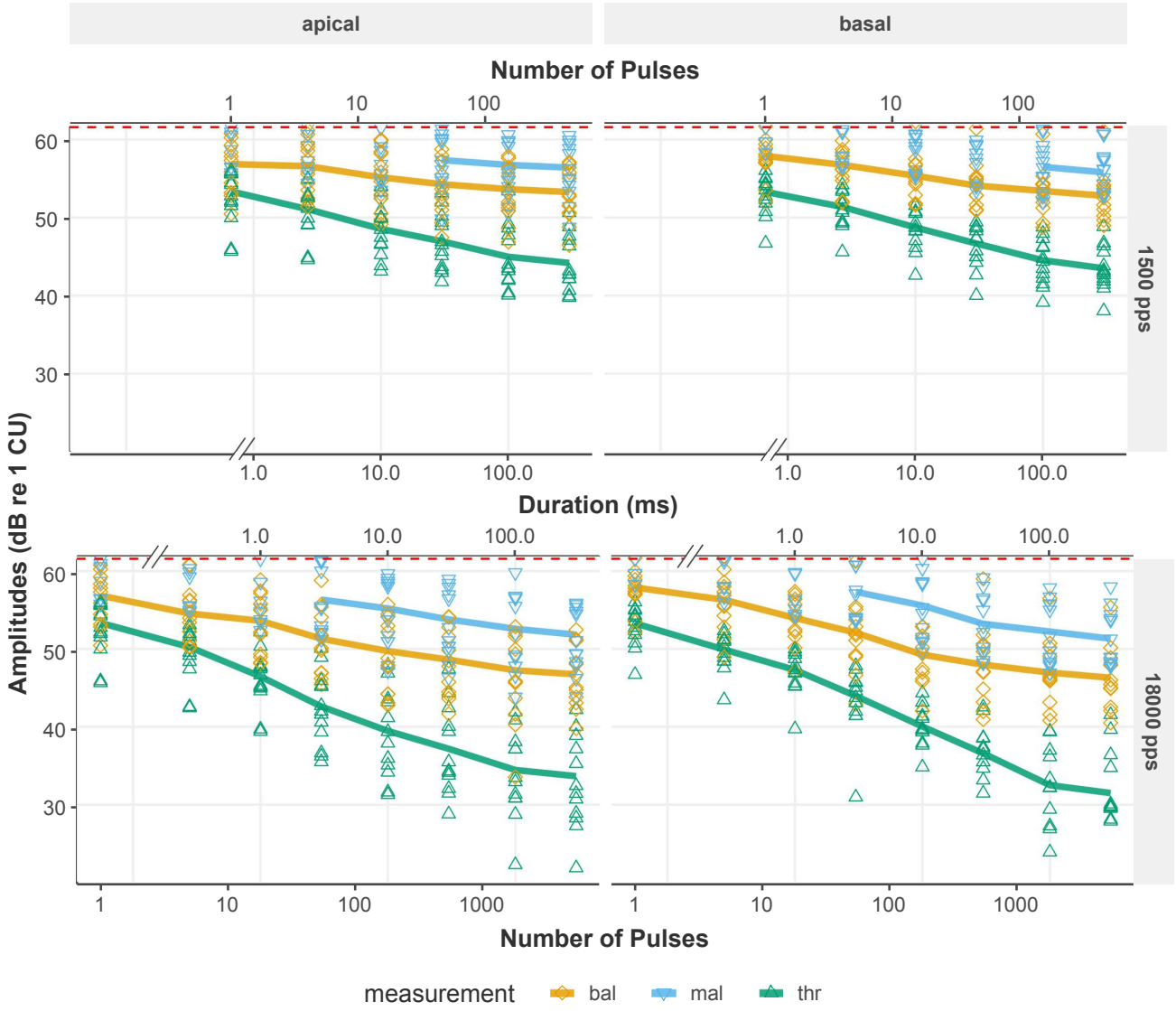


Figure 2. Median amplitudes as a function of duration (middle axes) and number of pulses (axes above and below the graphs) for thresholds (green), maximum acceptable levels (blue) and loudness balancing (orange) for all participants. The solid lines depict the trimmed mean of all subjects' data points. The dashed red line represents the maximum possible stimulation amplitude of 1200 CU. **Top:** 1500 pps. **Bottom:** 18000 pps. **Left:** Apical electrode. **Right:** Basal electrode.

Loudness Integration

Several phenomenological models for (acoustic) loudness integration have been proposed in the past[9,13,15,24]. Of those, the power function (which appears as a linear function in the double-logarithmic space) is one of the simplest options. The decision to fit this function to our data is not only based on its simplicity, but because it also allows for a direct comparison of the slopes of the fitted functions.

The power function does not allow to compare the critical duration, after which loudness does not vary with duration any more. However, from visual inspection it seems like this saturation point is only reached in few participants, for specific conditions. This possible critical duration will be ignored in the following analyses, as it has not been consistently found in the acoustic literature at shorter durations[8,11]. The current amplitudes (I in CU) as a function of the number of pulses were fitted for each individual subject and condition to a power-law of the form

$$I(N) = I_1 \cdot N^m. \tag{2}$$

The parameter I_1 is the current value for single pulse stimulation, and m represents the slope of the decreasing values for increasing number of pulses.

Fits of the threshold data work quite well, with an average of $R^2 = 0.96 \pm 0.05$. Data obtained in the other mea-

Table II. Results of data fits with equation 2 for the parameters I_1 and m (slopes), ± standard deviations (SD) across subjects.

Meas.	Elec.	Rate/pps	R^2	I_1/siCU	m
THR	apical	1500	0.94 ± 0.09	479 ± 146	-0.15 ± 0.07
THR	apical	18000	0.98 ± 0.02	492 ± 147	-0.27 ± 0.06
THR	basal	1500	0.95 ± 0.04	468 ± 116	-0.19 ± 0.04
THR	basal	18000	0.97 ± 0.01	474 ± 114	-0.27 ± 0.04
MAL	apical	1500	0.75 ± 0.26	1050 ± 305	-0.09 ± 0.05
MAL	apical	18000	0.87 ± 0.14	1286 ± 542	-0.15 ± 0.04
MAL	basal	1500	0.84 ± 0.20	1146 ± 257	-0.11 ± 0.04
MAL	basal	18000	0.89 ± 0.09	1371 ± 584	-0.16 ± 0.03
BAL	apical	1500	0.88 ± 0.26	771 ± 320	-0.09 ± 0.04
BAL	apical	18000	0.94 ± 0.06	758 ± 336	-0.17 ± 0.04
BAL	basal	1500	0.84 ± 0.13	792 ± 265	-0.10 ± 0.04
BAL	basal	18000	0.92 ± 0.06	823 ± 305	-0.17 ± 0.05

surements were also fitted with $R^2 = 0.90 \pm 0.16$ (BAL) and $R^2 = 0.85 \pm 0.19$ (MAL). Detailed numbers and the goodness of fit obtained for fitting the amplitudes to a function of the number of pulses are listed in Table II. Additionally, the fitting parameters I_1 and m are listed.

Exemplary fittings of the power function are included in Figure 1. In this figure, the data points of subject S11 (panel A) and S7 (panel B) are shown together with three lines, corresponding to the fits of the different measurements (THR, BAL, MAL).

Testing the effects of MEASUREMENT, RATE and ELECTRODE on the slopes reveals a significant disparity for the different MEASUREMENTS ($Q = 33.63, p < 0.001$). The slopes at threshold level are steeper than those of BAL and MAL, which entails an increase in dynamic range with increasing duration. In addition, the curves are steeper for the higher RATE ($Q = 353.65, p < 0.001$). The effect of ELECTRODE is also significant with $Q = 8.79, p = 0.003$; slightly larger slopes are observed at the basal electrode. Further, the interaction terms of MEASUREMENT and RATE ($Q = 5.37, p = 0.005$). The difference of the slope between THR and BAL levels is larger at the high rate. The interaction of RATE and ELECTRODE ($Q = 8.03, p = 0.005$) also reached the significance level of 0.05. The slope difference observed for the two cochlear positions (steeper curves basally) is more pronounced at the low stimulation rate.

Dynamic Range

An increase in dynamic range in dB can be seen in Figure 2 for both increasing duration and for the high rate (lower panels). When only looking at the maximum duration of 300 ms, the DR is larger in the high rate than in the low rate by 6.58 ± 3.79 dB (trimmed geometric means). The duration of 300 ms was chosen as this is approximately the duration of a syllable, and also because it is commonly used as pulse train duration in clinical fitting[34].

Statistical testing of dynamic range (in dB) using RATE and ELECTRODE as factors confirmed that the DR at 300 ms is significantly larger for the higher RATE ($Q = 41.49, p < 0.001$). It is not affected by the ELECTRODE ($Q = 2.99, p = 0.084$), nor by the interaction of the two factors.

DISCUSSION

In investigations with CI users, often a large variability between the individual participants' performance is observed, as it was also the case in the present experiment. This can partly be explained by different progression of diseases (including degeneration of the distal parts of SGNs), different levels of training with the CI due to differing time spans since implantation or residual hearing of the non-implanted side, or a dissimilar positioning of the electrode array in the cochlea, among others. For this reason, most analyses were done on an individual level first. Nonetheless, we measured effects of stimulation duration and stimulation rate that are common to all participants, even if the absolute current amplitudes differ.

Effects of Duration, Rate and Electrode

Overall, MAL and BAL curves in Figure 2 look very similar, only shifted vertically, whereas the threshold curve differs in its shape (steeper slope) from the other two. This is reflected in the fitted slope parameters.

The main effect of stimulation rate (lower amplitudes for the higher rate) on THR is in line with previously published results – if not somewhat higher. For pulse trains of 300 ms duration, we found 10.4 dB lower threshold amplitudes when stimulating with 18 000 pps instead of 1500 pps (rate increased by factor 12). This is equivalent to a decrease of 3.1 dB for a doubling of the rate.

Comparable but slightly lower values were found by Carlyon et al.[4], with a decrease of 7.7 dB for 400 ms pulse trains after increasing the rate from 500 pps to 3500 pps, equivalent to a threshold decrease of 2.7 dB per doubling of the rate. Kreft et al.[17] found a reduction in THR of 2.4 dB per doubling of the rate between 200 and 6500 pps, with a reduced effect beyond 3250 pps. Lower amplitudes caused by higher rates might be attributed to the effects of facilitation. With small inter-pulse intervals, several sub-threshold pulses can enable an action potential[3]. In Zhou et al.[35], human MPI above 1000 pps showed a median slope of around 2.6 dB per doubling of rate.

As hypothesized, there was no systematic effect of stimulation electrode found for threshold amplitudes. Nevertheless, some subjects showed markedly lower thresholds for one of the electrodes. An extreme case was S7, for whom thresholds at the basal electrode were at least 8.9 dB higher than at the apical electrode, for both of the tested rates.

In comparison, MAL decreased by 4.88 dB (median) between the two rates, or 1.25 dB per doubling of rate, comparable to the results in Kreft et al.[17] of 1.2 dB/doubling,

and somewhat lower than those in Skinner et al.[29] of around 1.7 dB/doubling between 1200 pps and 2400 pps.

Loudness Integration

In electrical hearing, differences between threshold and supra-threshold curves have been described by McKay and McDermott[21]. They describe that at higher levels, an increase in stimulus amplitude leads to a larger increase in aggregated stimulation than at lower levels. This was investigated by measuring loudness changes evoked by varying the interpulse interval between two pulses in a pair – which was itself repeated at 50 Hz. If it is possible to extend their findings to longer pulse trains, and it is assumed that similar levels of excitation in the nerve lead to similar loudness percepts, then it would be expected that smaller amplitude reductions are necessary at higher levels to compensate for longer pulse trains. This would imply steeper temporal integration curves at threshold than at higher levels, as it was seen in this experiment.

A study with normal hearing subjects revealed nonmonotonical dependencies on level on the slope of TI curves[7]. Both 1 kHz tones and white noise (WN) were investigated at 5, 30 and 200 ms. The strongest TI was found at medium levels, for both tones (around 56 dB SPL) and WN (around 76 dB SPL). Interestingly, they attribute this effect on properties of the loudness function related the basilar membrane mechanics, which should not play any role in our subjects.

In the present measurement, slopes at THR were steeper (median $m_{\mathrm{THR}} = 0.237$) than at moderate ($m_{\mathrm{BAL}} = 0.133$) or maximum loudness ($m_{\mathrm{MAL}} = 0.132$).

As in the results of Gerken et al.[11], who also used a power-law like function for fitting, we do not have to vary the exponent with duration (and in contrast to the results of e.g. Green et al.[13]). At least for threshold amplitudes up to 300 ms, fitting is satisfying with an overall mean of $R^2 = 0.96 \pm 0.05$.

Just like reported in other studies, our data contain large inter-individual variability. Regarding the slopes of the fitted threshold function results at conditions of e.g. basal positions at 1500 pps vary from $m = -0.07$ for subject S7 to $m = -0.27$ for subject S4r.

The mean of $m = -0.17 \pm 0.06$ for the low rate and $m = -0.28 \pm 0.05$ for the high stimulation rate can be translated to -3.44 ± 1.18 dB and -5.43 ± 1.00 dB decrease in threshold per decade, respectively. The slope for the high stimulation rate is closer to the amplitude decrease of $-20/3$ dB per decade reported by Heil et al.[14] for acoustic hearing, whereas the curve for the lower rate stimulation is not as steep. Already Donaldson et al.[6] reported much shallower slopes for CI users when compared to normal hearing listeners, whose slopes are at about 2 or 3 dB per doubling of duration (see also Shannon[26]). It seems that only with very high stimulation rates do the slopes of CI users become more similar to those of normal hearing participants.

Dynamic Range

The dynamic range increased with increasing stimulation rate for all durations, but even more for longer than for shorter stimulus duration. For the 300 ms pulse trains, an increase of the DR of 6.58 ± 3.79 dB was observed. The general increase can be explained by the fact that especially threshold amplitudes are lowered by increasing the rate, but not so much the amplitudes of measurements at the higher stimulation levels. The same has been reported by McKay and McDermott[21].

Bonnet et al.[2] report an increase in DR of 1.3 dB for a doubling of the stimulation rate. Our values suggest 1.84 dB for a doubling of the rate. Since the rates we compared (1500 dB and 18 000 dB) are much higher then theirs (774 dB and 3868 dB), there could be an additional influence on the effect of rate on THR and MAL for rates higher than 3868 dB. The increase of DR in our data is also much stronger than the one found by Zhou et al.[36]. They report an increase in DR of 1.19 dB for a doubling of the stimulation rate, for increasing the rate up to 5000 pps.

The fact that the dynamic range increased even more for long durations is refelected by the differences in slopes for the two rates. The threshold curve for the higher stimulation rate showed a steeper slope, leading to a more pronounced DR increase for long durations.

For CI fitting in the clinics, only single electrode DRs are set. Since we assume that many neurons throughout the cochlea are stimulated by single electrode contacts, the true possible DR might be larger than set in clinical CI fittings.

CONCLUSIONS

In the present work, we measured temporal integration in cochlear implant users as a function of duration at three different levels, at moderate and very high rates and at two positions in the electrode array. In general, an increase in stimulation rate lowered the current amplitudes at all loudness levels, and this effect of rate was stronger at threshold than at suprathreshold stimulation. This lead to a increased dynamic range for the higher rate that seems consistent, if not higher than those found in the literature.

Slopes of integration as a function of duration where also steeper for THR than for suprathreshold stimuli. The increased rate also led to an increase in slope, and the effect of rate itself was also stronger at threshold. Additionally, there were effects of electrode, with the basal electrodes showing steeper TI-curves than apical electrodes; this effect was more prominent at the lower stimulation rates.

Acknowledgments

We are particularly grateful to the study participants for their important contribution, as well as to Dr. Sonja Karg for her valuable feedback on the methods. The Institut für Ionenphysik und Angewandte Physik, University of Innsbruck

provided the RIB2. The ENT at Klinikum Großhadern, the Hörzentrum of the Klinikum rechts der Isar, the Therapiezentrum für Hörgeschädigte (Praxis Hanik, Laim) and the Campus Lauscher at Großhadern all collaborated in recruiting volunteer subjects with cochlear implants.

REFERENCES

References

[1] Adel, Y., Hilkhuysen, G., Noreña, A., Cazals, Y., Roman, S., and Macherey, O. (2017). Forward masking in cochlear implant users: electrophysiological and psychophysical data using pulse train maskers. *Journal of the Association for Research in Otolaryngology*, 18(3):495–512.

[2] Bonnet, R. M., Boermans, P.-P. B., Avenarius, O. F., Briaire, J. J., and Frijns, J. H. (2012). Effects of pulse width, pulse rate and paired electrode stimulation on psychophysical measures of dynamic range and speech recognition in cochlear implants. *Ear and hearing*, 33(4):489–496.

[3] Boulet, J., White, M., and Bruce, I. C. (2016). Temporal considerations for stimulating spiral ganglion neurons with cochlear implants. *Journal of the Association for Research in Otolaryngology*, 17(1):1–17.

[4] Carlyon, R. P., Deeks, J. M., and McKay, C. M. (2015). Effect of pulse rate and polarity on the sensitivity of auditory brainstem and cochlear implant users to electrical stimulation. *Journal of the Association for Research in Otolaryngology*, 16(5):653–668.

[5] Carlyon, R. P., Van Wieringen, A., Deeks, J. M., Long, C. J., Lyzenga, J., and Wouters, J. (2005). Effect of inter-phase gap on the sensitivity of cochlear implant users to electrical stimulation. *Hearing research*, 205(1-2):210–224.

[6] Donaldson, G. S., Viemeister, N. F., and Nelson, D. A. (1997). Psychometric functions and temporal integration in electric hearing. *Journal of the Acoustical Society of America*, 101(6):3706–3721.

[7] Florentine, M., Buus, S., and Poulsen, T. (1996). Temporal integration of loudness as a function of level. *The Journal of the Acoustical Society of America*, 99(3):1633–1644.

[8] Florentine, M., Fastl, H., and Buus, S. (1988). Temporal integration in normal hearing, cochlear impairment, and impairment simulated by masking. *The Journal of the Acoustical Society of America*, 84(1):195–203.

[9] Garner, W. R. and Miller, G. A. (1947). The masked threshold of pure tones as a function of duration. *Journal of Experimental Psychology*, 37(4):293–303.

[10] Gelfand, S. A. (2017). *Hearing: An introduction to psychological and physiological acoustics*. CRC Press.

[11] Gerken, G. M., Bhat, V. K. H., and Hutchison-Clutter, M. (1990). Auditory temporal integration and the power function model. *The Journal of the Acoustical Society of America*, 88(2):767–778.

[12] Gerken, G. M., Solecki, J. M., and Boettcher, F. A. (1991). Temporal integration of electrical stimulation of auditory nuclei in normal-hearing and hearing-impaired cat. *Hearing research*, 53(1):101–112.

[13] Green, D. M., Birdsall, T. G., and Tanner Jr, W. P. (1957). Signal detection as a function of signal intensity and duration. *The Journal of the Acoustical Society of America*, 29(4):523–531.

[14] Heil, P., Matysiak, A., and Neubauer, H. (2017). A probabilistic Poisson-based model accounts for an extensive set of absolute auditory threshold measurements. *Hearing research*, 353:135–161.

[15] Hughes, J. W. (1946). The threshold of audition for short periods of stimulation. *Proceedings of the Royal Society of London B*, 133(873):486–490.

[16] Ifukube, T. and White, R. L. (1987). Current distributions produced inside and outside the cochlea from a scala tympani electrode array. *IEEE transactions on biomedical engineering*, 34(11):883–890.

[17] Kreft, H. A., Donaldson, G. S., and Nelson, D. A. (2004). Effects of pulse rate on threshold and dynamic range in Clarion cochlear-implant users (L). *The Journal of the Acoustical Society of America*, 115(5):1885–1888.

[18] Malherbe, T. K., Hanekom, T., and Hanekom, J. J. (2016). Constructing a three-dimensional electrical model of a living cochlear implant user's cochlea. *International journal for numerical methods in biomedical engineering*, 32(7):e02751.

[19] McFadden, D. (1975). Duration- intensity reciprocity for equal loudness. *The Journal of the Acoustical Society of America*, 57(3):702–704.

[20] McKay, C. M., Henshall, K. R., Farrell, R. J., and McDermott, H. J. (2003). A practical method of predicting the loudness of complex electrical stimuli. *The Journal of the Acoustical Society of America*, 113(4):2054–2063.

[21] McKay, C. M. and McDermott, H. J. (1998). Loudness perception with pulsatile electrical stimulation: The effect of interpulse intervals. *The Journal of the Acoustical Society of America*, 104(2):1061–1074.

[22] Nelson, D. A. and Donaldson, G. S. (2002). Psychophysical recovery from pulse-train forward masking in electric hearing. *The Journal of the Acoustical Society of America*, 112(6):2932–2947.

[23] Pfingst, B. E., Colesa, D. J., Hembrador, S., Kang, S. Y., Middlebrooks, J. C., Raphael, Y., and Su, G. L. (2011). Detection of pulse trains in the electrically stimulated cochlea: Effects of cochlear health. *The Journal of the Acoustical Society of America*, 130(6):3954–3968.

[24] Plomp, R. and Bouman, M. A. (1959). Relation between hearing threshold and duration for tone pulses. *The Journal of the Acoustical Society of America*, 31(6):749–758.

[25] Poulton, E. C. (1989). *Bias in Quantifying Judgements*. Lawrence Erlbaum, Hove.

[26] Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. i. basic psychophysics. *Hearing research*, 11(2):157–189.

[27] Shannon, R. V. (1989). A model of threshold for pulsatile electrical stimulation of cochlear implants. *Hearing Research*, 40(3):197–204.

[28] Shannon, R. V. (1990). A model of temporal integration and forward masking for electrical stimulation of the auditory nerve. In Miller, J. M. and Spelman, F. A., editors, *Cochlear Implants: Models of the Electrically Stimulated Ear*, pages 187–205. Springer, New York, NY.

[29] Skinner, M. W., Holden, L. K., Holden, T. A., and Demorest, M. E. (2000). Effect of stimulation rate on cochlear implant recipients' thresholds and maximum acceptable loudness levels. *Journal of the American Academy of Audiology*, 11(4):203–213.

[30] Stevens, J. C. and Hall, J. W. (1966). Brightness and loudness as functions of stimulus duration. *Perception & Psychophysics*, 1(9):319–327.

[31] Viemeister, N. F. and Wakefield, G. H. (1991). Temporal integration and multiple looks. *The Journal of the Acoustical Society of America*, 90(2):858–865.

[32] Wilcox, R. R. (2011). *Introduction to robust estimation and hypothesis testing*. Academic press.

[33] Wilcox, R. R. and Schönbrodt, F. D. (2018). *The WRS package for robust statistics in R (version 0.35)*.

[34] Wolfe, J. and Schafer, E. (2014). *Programming Cochlear Implants*. Core Clinical Concepts in Audiology. Plural Publishing, Inc., San Diego, CA, second edition edition.

[35] Zhou, N., Kraft, C. T., Colesa, D. J., and Pfingst, B. E. (2015). Integration of pulse trains in humans and guinea pigs with cochlear implants. *Journal of the Association for Research in Otolaryngology*, 16(4):523–534.

[36] Zhou, N., Xu, L., and Pfingst, B. E. (2012). Characteristics of detection thresholds and maximum comfortable loudness levels as a function of pulse rate in human cochlear implant users. *Hearing research*, 284(1-2):25–32.

5

Electrode Impedance and Voltage Spread

Miguel Obando[1,2], Jakob Eberharter[2], Siwei Bai[2] and Werner Hemmert[1,2]
[1)] Graduate School of Systemic Neurosciences, Ludwig-Maximilian Universität München
[2)] Bio-inspired Information Processing, Department of Electrical Engineering, Technical University of Munich

Abstract

We used the telemetry capabilities built in cochlear implants to measure the voltage spread along the electrode array caused by monopolar electrical stimulation. For stimulation at each of the 12 electrodes in the array, the voltage at each of the unstimulated 11 electrodes was measured at the end of the anodic phase of a biphasic pulse of 50 CU amplitude. Ten subjects participated in the present experiment, for a total of 16 implants measured.

These results were compared to simulation data from a volumetric finite-element model developed in our workgroup. The simulation fell within the measured range, and the average decay of voltage along the array was well described. Most of the variance between subjects could be attributed to a subject-dependent voltage offset.

Note: Parts of this work has been published as a part of Bai et al. (2019). The reader is directed to that publication for further details regarding the finite-element model, especially those concerning the parameters of the simulation and the software used.

5.1 Introduction

Detailed geometrical models of the cochlea are a crucial tool for the development of cochlear implants. By splitting the cochlea into small elements, finite-element models (FEM) allow the numerical solving of physical equations and thus the calculation of the current spread in the cochlea and its surrounding structures. Together with biophysically realistic models of the auditory nerve fibres (ANFs) and their activation, they provide a valuable way to better understand and predict the responses to cochlear implant (CI) stimulation. This can then be used, together with physiological and psychophysical experiments, to evaluate and develop new coding strategies for CIs, as well as to understand how to individualize their settings better.

Because the output of biophysical neural models is conditioned by its input, it is important to validate the 3D models with realistic and accurate measurements of the actual voltage along the cochlea. In this work, we present current spread measurements from 16 MED-EL implants, with the goal to allow the evaluation of a volumetric cochlear model developed by Bai et al. (2019) in our workgroup.

5.1.1 3D Finite element models

Hanekom and Hanekom (2016) gives a detailed review on different 3D models in their cochlea and their current and possible applications within the context of cochlear implant development. Two main different types of 3D models with drastically different goals and limitations can be differentiated: generic models on one hand, and on the other user-specific models. The former type attempts to generalize predictions based on a single morphology, with the goal of explaining typical, population-wide behaviour. On the other hand, the latter attempts to specifically model the cochlea of living implantees, in order to explain user-specific behaviours that are, for example, influenced by the specific morphology.

Broadly speaking, the generation of such models requires at least five steps with specific challenges:

1. Image acquisition
2. Image segmentation
3. Geometric refinement and repairing
4. Finite element meshing
5. Model simulation (including physics)

First, the cochlea and the surrounding structures have to be imaged. Cochlear images from living human beings and in general of living mammals can be acquired for example through computer tomography (CT). This process has the limitation that the high energy photons that are needed for high detail and resolution are destructive for the tissue and noxious

to the individual. *Ex vivo* imaging allows for much higher resolutions, but is not useful in the context of individual specific models of living patients.

Then, from the images (e.g. 2D slices in a CT), the 3D model has to be reconstructed. This entails the proper identification of structures in a slice (segmentation), as well as identifying structures in neighbouring slices as part of a single 3D domain, with the result being shell geometries. These shells have to be refined and repaired, so that they accurately reflect the original anatomy. Later, the volumes enclosed in these shell are filled with a finite number of polyhedral elements – usually tetrahedral in most physics applications. At this point, the whole model can be integrated into an encompassingy head model with coarser mesh elements for a better simulation. As a last step, the physical properties of the system must be brought down to equations which can then be numerically solved.

5.1.2 Previous measurements

There are several ways to validate data from such a 3D model. The most direct possibility is to measure directly *in situ*, either inside of e.g. a cadaveric extraction or *in vivo* in animals, as done in publications by Ifukube and White (1987), Jolly, Spelman, and Clopton (1996), and Kral et al. (1998).

This has the disadvantage that the context of the measurement, that is the structures outside of the temporal bone, is lost. Additionally, the extraction of the anatomy is by nature destructive, which might distort measurements. In the specific case of humans, it is hard to come by large numbers of bones, which will necessarily reduce the available sample sizes. Cochleas of experimental animals (e.g. guinea pigs or cats) are structurally very different to human ones, which limits the validity of measurements. This is especially important considering that the results from both animals and computational fibre models are at odds with behavioural and physiological results measured in humans (e.g. Macherey et al., 2006; Undurraga et al., 2010, among others).

Other indirect possibilities for validation consist in measuring the nerve response to stimulus. This can again be done directly with single-fibre recordings in animals, or indirectly with ECAPs (electrically evoked compound action potential), a measure of the compound neural response of the auditory nerve fibres.

The peripheral neural response can be predicted by joining the volume models with nerve models, which leads to a very indirect validation with an increased number of free parameters. ECAPs in humans have the additional problem that the level of survival of the nerve processes is unknown.

A last possibility which represents a middle way is to use the capabilities of the modern CI-hardware to measure with the implanted electrode itself. Besides the normal stimulation of the cochlea, modern cochlear implants also allow back-telemetry, i.e. to send measurement information from the implant to the speech processor. The two main uses of telemetry are to measure the ECAPs and to measure the intracochlear voltages.

Measuring the voltage at the stimulating electrode during stimulation with a set current allows the calculation of an impedance value $Z = U/I$. These impedance values are commonly used clinically to diagnose the integrity of the electrode contacts, especially short-circuits and open contacts. On the other hand, measuring the voltage at a different electrode provides a measure of the voltage spread in the cochlea along the axis of the electrode array. Depending on the manufacturer, this process is called impedance field telemetry (IFT) for MED-EL devices, or electrical field imaging (EFI) in Advanced-Bionic devices – see Mens (2007) for a review of existing telemetry systems. This procedure has the disadvantage of a limited number of measuring points, namely one per electrode. Additionally, there is some uncertainty regarding the exact positioning of the stimulated electrodes, depending on the availability of detailed imaging of the individual implanted cochleas.

Malherbe, Hanekom, and Hanekom (2015) investigated the effects of the resistivity of the surrounding bone structures in the voltage spread of a previous volumetric model, and compared it to values from literature (Tang, Benítez, and Zeng, 2011, 5 subjects). Many other available data from human subjects is restricted to similarly low sample sizes (Vanpoucke et al., 2004, 5 subjects). An exception is the work of Nogueira et al. (2016), who used measurements from 12 subjects to validate a parametric, individualizable model.

More recently, our group developed a generic model (Bai et al., 2019) from a cadaverous extraction of a human temporal bone. Its novelty lies in the high resolution of the images and of the resulting volumetric meshes, which was brought down to 5.9 µm. The results from the simulation are presented here for comparison. These computation was done for a lateral placement of the array electrode – i.e. along the external wall of the cochlea.

5.2 Methods and Equipment

5.2.1 Finite Element Model

The model is based on a high resolution (with an isotropic resolution of 9.6 µm) micro-CT scan of a cadaveric human temporal bone with an inserted dummy-electrode made from pure silicone. The slices of the CT image were then segmented manually, using interpolation between every second or third slice, with either bony labyrinth, cochlear canal or cochlear nerve as possible compartments. This process results into a surface mesh for every compartment.

A separate volumetric model was acquired by segmenting the T1-MRI images of a human head with a resolution of 1 mm, and scalp, skull and brain as com-

partments. The surface mesh of the temporal bone was then embedded into the head model, and after pre-processing, a volumetric mesh was generated.

The resulting composite mesh was then imported into an FE-solving software for the simulation of the electrical stimulation, using a discrete Laplace Equation. Under a quasi-static approximation, the tissue was assumed to be purely resistive during the relatively fast electrode stimulation. The conductivities of the different tissues were were not adapted or fitted, but fixed at values known from literature.

In this electric model, the voltage was measured at the surface every electrode as a response to a stimulation with an electric current of 50 µA at electrode surfaces with a radius of 0.18 mm. The reference electrode was positioned on the skull, superior and posterior to the external acoustic meatus. Figure 5.1(b) shows the simulation results.

5.2.2 Stimuli and subjects

For the experimental validation of the model results, Biphasic pulses of 40 µs, with an inter-phase gap of 2.1 µs and with the cathodic (negative) phase leading, were used as stimuli. The voltage at the measuring electrode was measured by the system at the end of the anodic phase in the stimulating electrode. The pulses had an amplitude of 50 CU (1 CU $\approx$ 50 µA). 16 MED-EL implants from 10 different subjects were measured.

A full voltage spread matrix was measured, meaning that all (active) electrodes were measured against all other electrodes respectively. Some data points are missing due to the electrodes being deactivated or showing clearly outlying, very high impedances, which indicate bad contacts. This was mostly the case for the most basal electrodes, possibly indicating electrodes not completely inside the cochlear fluid.

The measurements presented in this work were conducted previous to different experiments in our workgroup. All subjects gave their informed consent and received monetary compensation for their participation. Measurements were conducted in accordance to the Declaration of Helsinki, and were approved by the medical ethics committee of the Klinikum rechts der Isar (Munich, 2126/08).

5.2.3 RIB2 IFT and Telemetry

A Research Interface Box (RIB2, University of Innsbruck) was used to communicate with the implants. Custom software written in Python was used to generate the stimuli and record the telemtry results. The MED-EL IFT system uses a track-and-hold circuit, which follows the voltage only during anodic phases and holds the voltage at their end. This measured voltage is then output as 2048 bits of adaptive sigma-delta-modulated data (Zierhofer, 2000). From these, a voltage value can be obtained by averaging and multiplying by a factor provided by the manufacturer. Refer to Neustetter (2014) for a more detailed description and characterization of the IFT system.

5.3 Results and Data Analysis

Figure 5.1 shows the aggregated results of voltage spread, as well as the simulation results. Each line represents stimulation presented on one electrode (blue: electrode 1, yellow: electrode 12), and each position in the x-axis represents the electrode where the measurement was made. Measurements in the stimulating electrode can only be made while the stimulation is still ongoing. They were thus left out, since the values do not take into account the differences caused by the electrode-lymph interface. In MED-EL electrode arrays, electrode 1 is the most apical and electrode 12 the most basal. Each point represents the average over all subjects and implants ($N = 16$).

There is, as with any measurement with cochlear implant subjects, a broad spread of data. For visualization, Figure 5.2 shows the measured data for all individual subjects at two exemplary electrodes, basal and apical, together with its mean (blue), and the simulation data (red).

The general shape of the simulated curve is correct, predicting a steeper voltage decrease towards the base than towards the apex. However, the simulation slightly overshoots the mean curve. Since the experimental curves have a similar shape, a possible way to describe the spread between the subjects is by way of a correcting voltage offset, so that they are better described by the stimulation data. In order to properly fit the curves, these offsets have to be adjusted not only on a subject but also for each measuring electrode of a subject, though in a smaller magnitude. Figure 5.3 shows the offset-spreads along the arrays of individual subjects as box-plots. The offsets thus obtained ranged from about −30 to +50 mV, with standard deviation $s = 25$ mV.

It can be seen that implant-specific across-subjects offset already explains most of the variability; controlling for it, the rest of the offset variability ranges from −9 to +9 mV ($s = 3.16$ mV), and does not have an obvious structure.

5.4 Discussion

We validated predictions from a 3D volumetric voltage-spread model of the cochlea with intra-cochlear measurements from 16 implanted electrodes. The simulation fell within the range of measurements, and in general had the same shape as them. This already indicates a good degree of validity for the model, especially considering that the electrical properties used in it were taken from literature and not fitted. In general, it can be seen in both measurements and simulation that the voltage decays more strongly towards the basis than towards the apex. Additionally, this decrement is always of a very small magni-

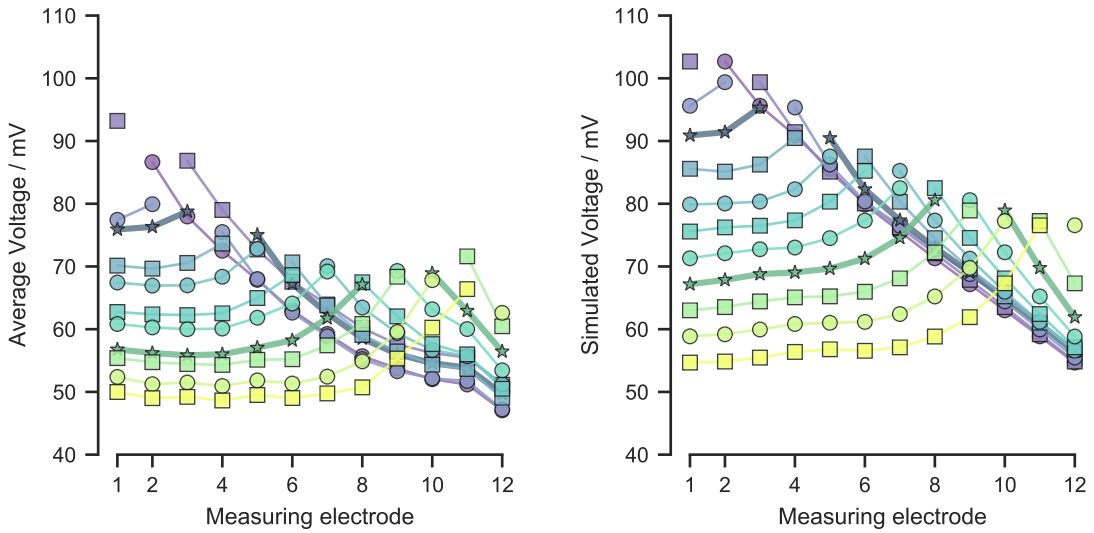


FIGURE 5.1: Left: Averaged measured voltages for each stimulating electrode (blue: electrode 1, yellow: electrode 12), measured across the electrode array (x-axis). Right: simulation results. The thicker, darker lines with stars represent stimulation in electrodes 4 and 9 (see Figure 5.2).

tude. As a matter of example, the largest decrease was seen for stimulation at electrode 1, where the average voltage dropped from around 87 mV at electrode 2 to only 50 mV at electrode 12 (for a maximum drop of 4.8 dB across almost the whole electrode array).

In order to better explain the deviations between measurements and model, a simple voltage offset was added to every implant (subject and side) as well as to every measuring electrode. The subject-dependent offset could explain most of the observed variability. Particularly, such broad offsets should play a negligible role in the activation of neurons, since nerve fibres are activated by the second spatial derivative of potential along their axons (Rattay, 1989).

The fine-structure variations play a more important role. They could for instance stem from actual differences in the cochlear geometry, in the resistivity of different domains (e.g. of bones becoming less dense with age), or different geometries altogether (like thicker skin). This would allow for a fine-tuning of the model in order to more accurately represent subject data.

5.4.1 Limitations of measurements

The present measurement has relevant limitations. For instance, the voltages are measured at a single time point. However, the electrodes are not purely resistive (as is also assumed in the model), and thus the voltage measured by an electrode is not com-

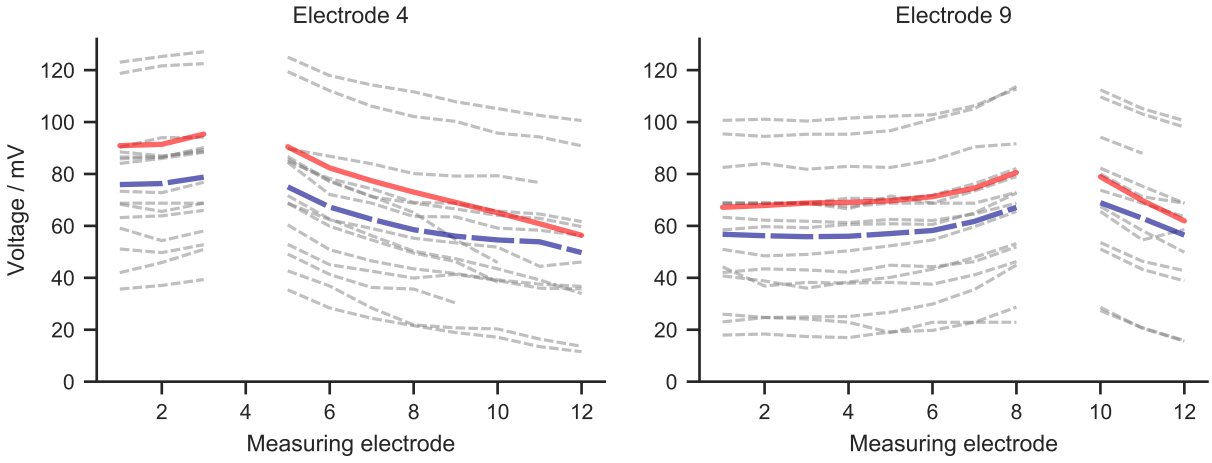


FIGURE 5.2: Individual subject data (dashed lines) for two exemplary electrodes, 4 and 9, and the mean over subjects (blue). The solid red line represents the simulation data.

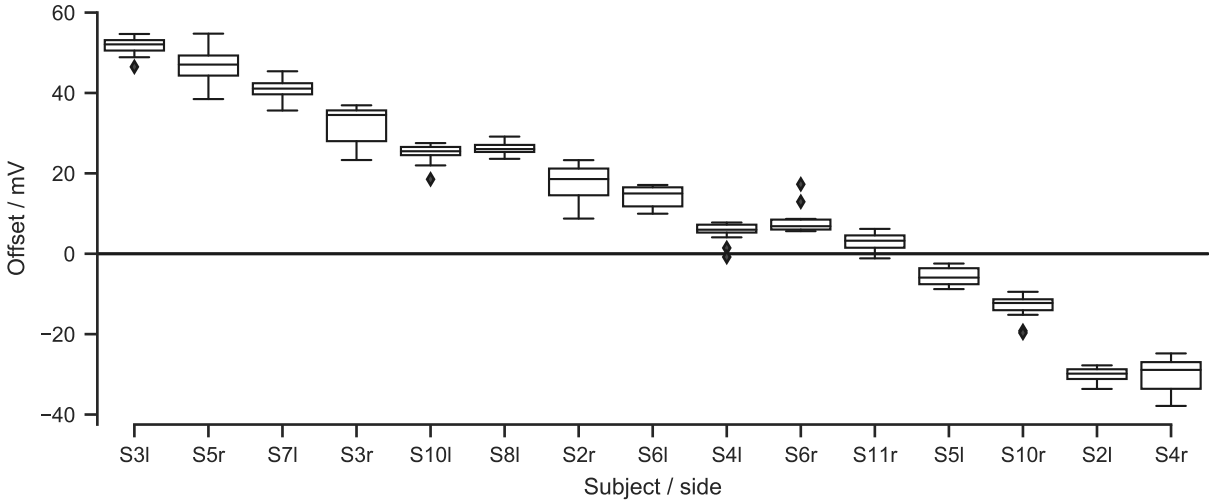


FIGURE 5.3: Distribution of fitted voltage offsets for each subject. Each box-plot represents 12 values, one for each possible measuring electrode. Subjects were ordered by the subject-specific offset.

pletely independent of time. It would be possible to coarsely measure the actual complex impedance of the stimulating electrodes caused by the rectangular current with the telemetry capabilities of the implants that are usually used to measure ECAPs. These measurements are normally dominated by the electrical artefact.

Also, the measurements are not capable of detecting individual disturbances in the arrays, like electrodes having reduced contact due to scaring or tissue growth, or the presence of air bubbles. Lastly, we did not have access to clinical imaging data of the patients, so that we are not able to assess the size of patients' cochleas or the exact placement of the electrode arrays inside of them.

5.4.2 Outlook

For the present experiment, the simulation data was calculated with the geometry of a single measured human cochlea. A wider catalogue of segmented and modelled cochleas would allow to better identify and quantify real variability in the morphology of the human population, even though this is a difficult task.

In addition, it would be worthwhile to find those sensitive parameters in the models that reflect the real variability found in users – either direct physical parameters like individual resistivities, or geometrical parameters related to the shape of the cochlea. Ideally, it should then be possible to fit the individual model to the human data, which could in turn allow for a better understanding of the individual cochlea and thus more personalised predictions about the effect of stimulation parameters on the individual.

With a much larger database of voltage spread measurements, a second possibility arises, namely to carry out a dimensionality analysis of electrode measurements that could help identify key features of variability or characteristics of sub-populations among CI-users.

Acknowledgements

We would like to thank the test subjects for their participation and contributions to this work.

Bibliography

Items marked with † denote (yet) unpublished work.

Bai, Siwei; Encke, Jörg; Obando-Leitón, Miguel; Weiß, Robin; Schäfer, Friederike; Eberharter, Jakob; Böhnke, Frank, and Hemmert, Werner (2019). "Electrical Stimulation in the Human Cochlea: A Computational Study Based on High-Resolution Micro-CT Scans" (cit. on pp. 47 sq.).

Hanekom, Tania and Hanekom, Johan J. (2016). "Three-Dimensional Models of Cochlear Implants: A Review of Their Development and How They Could Support Management and Maintenance of Cochlear Implant Performance". In: *Network: Computation in Neural Systems* 27.2-3, pp. 67–106 (cit. on p. 47).

Ifukube, Tohru and White, Robert L (1987). "Current Distributions Produced inside and Outside the Cochlea from a Scala Tympani Electrode Array". In: *IEEE transactions on biomedical engineering* 34.11, pp. 883–890 (cit. on p. 48).

Jolly, C. N.; Spelman, F. A., and Clopton, B. M. (1996). "Quadrupolar Stimulation for Cochlear Prostheses: Modeling and Experimental Data". In: *IEEE Transactions on Biomedical Engineering* 43.8, pp. 857–865 (cit. on p. 48).

Kral, Andrej; Hartmann, Rainer; Mortazavi, Dariusch, and Klinke, Rainer (1998). "Spatial Resolution of Cochlear Implants: The Electrical Field and Excitation of Auditory Afferents". In: *Hearing research* 121.1-2, pp. 11–28 (cit. on p. 48).

Macherey, Olivier; Wieringen, Astrid van; Carlyon, Robert P.; Deeks, John M., and Wouters, Jan (2006). "Asymmetric Pulses in Cochlear Implants: Effects of Pulse Shape, Polarity, and Rate". In: *Journal of the Association for Research in Otolaryngology* 7.3, pp. 253–266 (cit. on p. 48).

Malherbe, Tiaan Krynauw; Hanekom, Tania, and Hanekom, Johannes Jurgens (2015). "The Effect of the Resistive Properties of Bone on Neural Excitation and Electric Fields in Cochlear Implant Models". In: *Hearing research* 327, pp. 126–135 (cit. on p. 48).

Mens, Lucas H. M. (2007). "Advances in Cochlear Implant Telemetry: Evoked Neural Responses, Electrical Field Imaging, and Technical Integrity". In: *Trends in Amplification* 11.3, pp. 143–159 (cit. on p. 48).

† Neustetter, Christian (2014). "Charakterisierung des Telemetriesystems des MED-EL PulsarCI100 Cochleaimplantates". Dissertation. Innsbruck: Leopold-Franzens-Universität Innsbruck. 106 pp. (cit. on p. 49).

Nogueira, Waldo; Schurzig, Daniel; Büchner, Andreas; Penninger, Richard T., and Würfel, Waldemar (2016). "Validation of a Cochlear Implant Patient-Specific Model of the Voltage Distribution in a Clinical Setting". In: *Frontiers in Bioengineering and Biotechnology* 4 (cit. on p. 48).

Rattay, Frank (1989). "Analysis of Models for Extracellular Fiber Stimulation". In: *IEEE Transactions on Biomedical Engineering* 36.7, pp. 676–682 (cit. on p. 50).

Tang, Qing; Benítez, Raul, and Zeng, Fan-Gang (2011). "Spatial Channel Interactions in Cochlear Implants". In: *Journal of neural engineering* 8.4, p. 046029 (cit. on p. 48).

Undurraga, Jaime A.; Van Wieringen, Astrid; Carlyon, Robert P.; Macherey, Olivier, and Wouters, Jan (2010). "Polarity Effects on Neural Responses of the Electrically Stimulated Auditory Nerve at Different Cochlear Sites". In: *Hearing Research* 269.1-2, pp. 146–161 (cit. on p. 48).

Vanpoucke, Filiep; Zarowski, Andrzej; Casselman, Jan; Frijns, Johan, and Peeters, Stefaan (2004). "The Facial Nerve Canal: An Important Cochlear Conduction Path Revealed by Clarion Electrical Field Imaging". In: *Otology & Neurotology* 25.3, pp. 282–289 (cit. on p. 48).

Zierhofer, C. M. (2000). "Adaptive Sigma-Delta Modulation with One-Bit Quantization". In: *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing* 47.5, pp. 408–415 (cit. on p. 49).

6

Conclusion, outlook and closing remarks

Through the experiments presented in the previous chapters, the temporal integration of electrical stimulation in cochlear implant wearers was investigated. Amplitude duration curves were measured at different levels, electrode positions and rates.

It could be shown that the integration of short pulse trains in CI-wearers follows a very different relation than normal hearing subjects. The slope of integration is much shallower at moderate rates, but at very high rates it increases to values more similar to those found in normal hearing subjects.

For the effect of level, it was found that the slopes of integration are more shallow for suprathreshold stimuli than for threshold stimuli, so that the available dynamic range increases with duration, in a manner consistent with previous literature findings. This effect correlates itself with the rate of stimulation, so that higher rates lead to wider dynamic ranges.

Effects of electrode position were also found. Basal electrodes have steeper amplitude-duration TI curves than apical ones. Additionally, the difference between the electrode positions was higher at the moderate stimulation rates than at the higher ones. Equivalently formulated, the effect of rate (lower amplitudes at higher rates) is stronger at apical electrodes.

The critical duration, i.e. the maximum duration at which integration can be seen, was seen to change dramatically across subjects and electrodes (see also below). Although the order of magnitude seems consistent with values found in the acoustic literature, the variability seems to be even wider that in NH subjects.

The investigations of very high pulse rates in addition to clinically relevant single pulse rates were based on the assumption of voltage spread throughout the cochlear fluid, and that this would lead to increased effective stimulation rates, since neurons would be affected not only by neighbouring electrodes but by those further away. The former point could be confirmed by measurements in actual implanted cochleas of humans through the telemetry hardware made available by the manufacturers. These measurements can be used in the future as a way of validating 3D volumetric models of the electrical currents in the cochlea, as are for example being developed in our work group.

6.1 Loudness integration – other comments

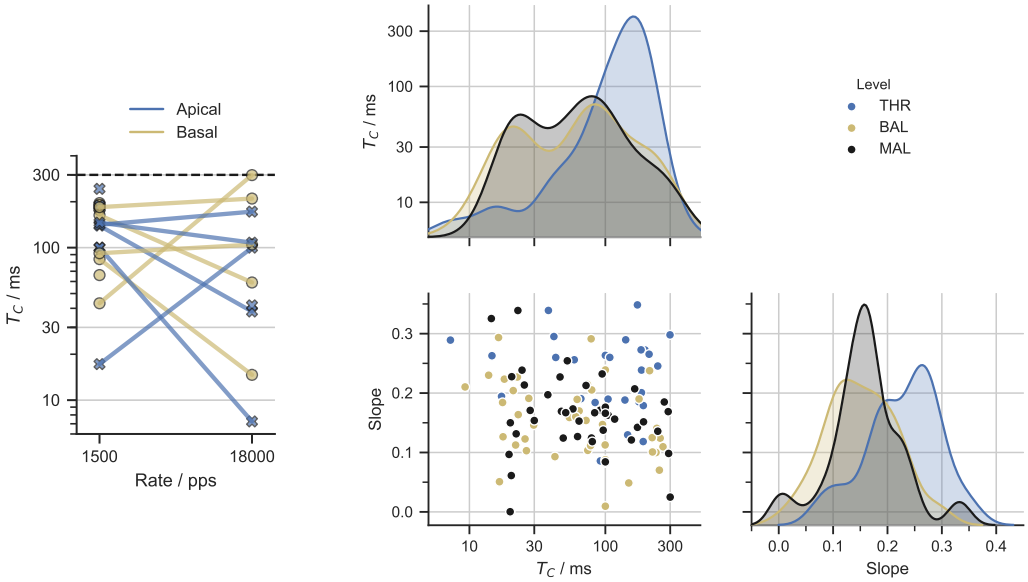
The following content was not included in the previous chapter due to reasons of space.

6.1.1 Critical duration

Just as critical duration was investigated in chapter 3 with a small population, it is possible to also analyse it with the larger data set of chapter 4. However, this analysis was omitted in the original manuscript due to the maximum pulse-train being much shorter (300 ms instead of 1 s), which did not allow for a good estimation. Indeed, for many subjects and electrodes (13 % at BAL and MAL, 40 % at THR), no critical duration could be estimated by the curve fitting. It is

expected that this is the case when T_C lies above the maximum measured pulse-train duration of 300 ms.

Figure 6.1(a) shows the effects of stimulation rate on T_C on those electrodes where it could be estimated at both rates, and Figure 6.1(b) shows the effect of stimulation level on both T_C and slope, as well as the joint distribution.



(a) Effects of stimulation rate on the critical duration T_C of thresholds.

(b) Effects of stimulation level on the distribution of T_C and of slope, as well as their joint distribution. Blue: Threshold. Yellow: Equal-loudness at comfortable levels. Black: Maximum levels.

FIGURE 6.1: Distribution of fitted critical durations for data in chapter 4.

Here, it can also be seen that an increase in rate is able to alter the critical duration drastically with otherwise identical conditions. However, neither the magnitude nor the direction of change is consistent, across either rates or electrodes. On the other hand, it can also be seen that the level of stimulation has a clear effect on the group-level distributions of both slope and T_C, with higher values for the threshold curves. Interestingly, the distributions for BAL (at 60 % DR) and MAL (maximum acceptable loudness) are almost indistinguishable. A larger integration period for THR than for suprathreshold stimuli makes sense considering the context of the task (detection vs. magnitude evaluation), and point towards a more central process. In contrast, across all subjects slope and DR seem to be affected very uniformly by stimulation rate, which points towards a more peripheral process occuring.

6.1.2 Variability as a function of duration

Another point that was observable in chapter 3 and later confirmed in chapter 4, but left out due to reasons of space, is an increase in the variability of the threshold adjustment with increasing durations. In order to quantify this, the coefficient of variation, given by $\mathrm{CV} = \sigma/\mu$ (i.e. the standard deviation of amplitude measurements divided by its mean) can be calculated for the repeated adjustments at each duration and condition. Despite the small amount of sample points per condition (3 and 4 in the respective chapters) and variability between subjects, the effect is clear. This is summarized in Figure 6.2. For each of the possible electrode-rate combinations, the CV of THR adjustments is plotted against duration in logarithmic axes. In general, it can be seen to increase with duration. Only THR was included for clarity, since values at the other levels are missing.

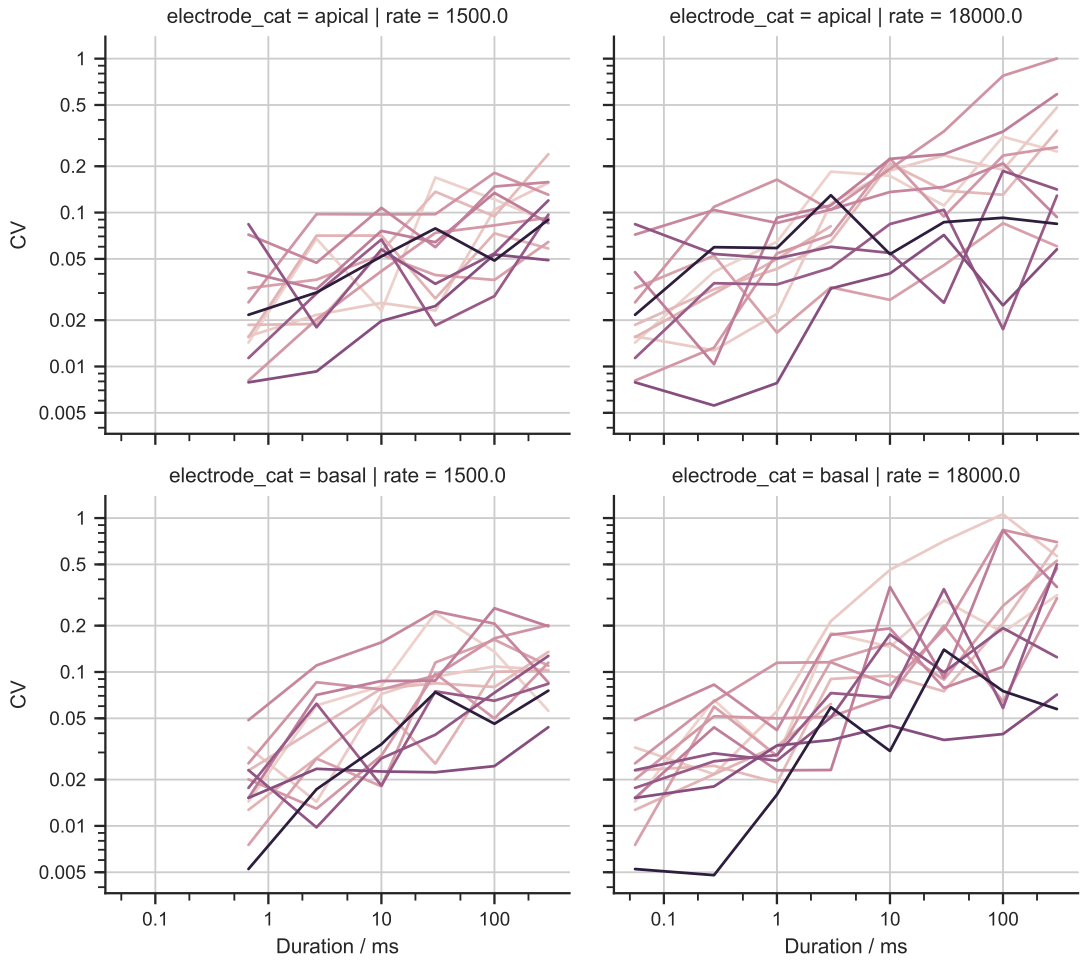


FIGURE 6.2: Coefficient of variation (CV $= \sigma/\mu$) of repeated THR adjustments as a function of duration for the different electrode/rate conditions. Different lines represent different subjects.

At these high rates, with the individual nerve fibres unable to fire with each pulse, the longer durations seem to enhance the effects of refractoriness that differ among the population.

A similar effect that could be seen in the data (not shown), is that the coefficients of variation reliably decreased from THR to MCL for most subjects and conditions.

6.1.3 Number of discriminable steps

Despite the increased dynamic range measured for higher stimulation rates in the previous chapters, this might not lead to possibly expected improvements in speech understanding. There is evidence indicating that the resulting higher dynamic range does not lead to more discriminable intensity steps, or in general to a better intensity discrimination. Azadpour, McKay, and Svirsky (2018) measured an increase in loudness JND with increasing rates between 500 to 3000 pps. As a direct consequence, there are fewer discriminable amplitude steps at higher rates. They suggest that this occurs because high rates lead to less synchronous firing of the ANFs (Miller, Hu, et al., 2008; Zhang et al., 2007), resulting in higher noise in the neural representation. This drawback cannot be corrected by changing the mapping law settings in the speech processor. Also Galvin and Fu (2009) showed that higher rates do not increase intensity

resolution when measured as a Weber fraction[1].

In the present experiments, the standard deviations (SDs) of repeated adjustments can be interpreted as a measure of intensity resolution, not unlike intensity JNDs. There, it was indeed found that the SDs of adjustment increased with rate. However if the SDs are normalised by dividing them by the respective dynamic ranges (in CU), the differences between the rates disappear for THR and MAL. Only for the balancing task there is a significant effect of stimulation rate left on the normalised SDs. This suggests that the number of loudness steps is similar for the two stimulation rates. This was the finding of Kreft, Donaldson, and Nelson (2004a), who measured a similar number of discriminable steps at 200, 1625 and 6500 pps – increases in DR were compensated by higher discriminable intensity difference (measured as Weber fractions), when rate was changed for the individual subjects. Across subjects, however, smaller dynamic ranges correlated with larger Weber fractions.

Lastly, there is an effect of rate on intensity discriminability. In the paper by Kreft, Donaldson, and Nelson (2004a), the Weber fraction was also seen to decrease with increasing level (i.e. more sensitivity to intensity changes), which can be attributed to recruitment of nerve processes in the modiolus (McKay, Henshall, et al., 2003). As mentioned above (subsection 6.1.2), a similar result could be seen in the data of Chapter 4 (not shown): for most subjects and conditions, the coefficient of variation decreased from THR to MCL. It thus seems that the SD of repeated adjustments can be a good proxy for JND measurements.

These results are in conflict with those from acoustic hearing. There, it has been measured that level-JNDs are larger at moderate than at low or high levels (Buus, Florentine, and Poulsen, 1997; Florentine, Buus, and Poulsen, 1996) – in the same way as how temporal integration depends on level. The authors speculate that this similarity between TI and level JNDs result from a common underlying factor, possibly the basilar membrane mechanics, which of course is bypassed with cochlear implant stimulation.

6.1.4 Loudness growth

In the present experiments, amplitude-duration curves were measured at threshold (THR), at maximum acceptable levels (MAL), and at a moderate equal loudness, corresponding to 60 % of dynamic range at a specific set of conditions (basal electrode, lower rate, long duration). During these experiments, it was assumed that both THR and MAL curves also represent equal loudness, and care was taken that BAL curves at different conditions were measured at the same loudness. The experiments were thus limited to only three levels.

In the BAL data of Chapter 4, amplitudes for the higher 18 kpps rate were balanced by the subjects at a significantly lower percentage of the dynamic range than at 1500 pps, for both apical and basal electrodes ($Q = 12.36$; $p < 0.001$). On average, stimuli with the lower rate were set at (59 ± 9) % DR and at (50 ± 12) % DR for the high rate.

From this it seems that, in addition to the results mentioned in the previous chapters, the stimulation rate also changes the shape of the loudness growth as a function of % DR. At the same loudness, the higher rate led to a lower value inside of the dynamic range. Thus, the loudness growth function – i.e. loudness as a function of stimulation amplitude – seems to be more concave at higher rates. This can be compared to the results by Galvin and Fu (2009), who also show differences in the loudness growth at different rates, especially at low stimulation levels.

However, because of the differences in perceptual quality, subjects had difficulty during the loudness balance tasks at different electrodes or different rates. For this reason, it is difficult to make any definite statements about how the positions of the BAL curves inside of the dynamic range compare with each other. A logical extension of these results would be thus to directly measure the loudness growth function in CI-wearers at different rates. Different psychophysical methods could be used to investigate these differences, like categorical loudness

[1]The Weber fraction is given by $(\Delta I/I)$, or in dB: $Wf_{\mathrm{dB}} = 10\log_{10}(\Delta I/I)$, where ΔI represents the JND of a stimulus with the amplitude I. It is not equivalent, but can be compared to the coefficient of variation as discussed above in subsection 6.1.2 and Figure 6.2.

scaling (see ISO, 2006, for a standard on the method; and Brand and Hohmann, 2002, for a possible implementation), magnitude estimation or production, loudness balancing between the different rates, or loudness balancing between an implanted ear and a normal hearing ear at different rates. The latter should give the most accurate and extendible results, although it requires uncommon test subjects.

This has already been measured in previous works for certain combinations of parameters, but the exact relationship is not yet clear. Some researchers propose an exponential relationship (e.g. Chatterjee, Fu, and Shannon, 2000), while others suggest the relationship is linear or follows a power-law (Shannon, 1983), so more data is necessary. Results of loudness growth at high rates could then be compared to existing literature results for lower rates (cf. Zeng and Shannon, 1994, from 100 to 1000 pps; Gallego et al., 1999, from 75 to 300 pps; or Fu, 2005, from 100 to 1000 pps). For these rates, it is known that the loudness growth is highly dependent on rate, with Zeng and Shannon (1994) arguing that the shape is a power law at frequencies below 300 pps, and exponential at higher frequencies. However, in the data by Gallego et al. (1999), loudness was seen to grow much more similarly after controlling for the different dynamic ranges at different rates.

6.1.5 Temporal integration - Slopes

From the results of different acoustic experiments in the literature, it has been assumed that the place of integration has to be central to the auditory nerve (Gerken, Bhat, and Hutchison-Clutter, 1990; Zwislocki, 1960), and maybe even higher than the inferior colliculus (Gerken, Solecki, and Boettcher, 1991).

On the other hand, subjects with hearing impairment show a reduced amount of temporal integration – or rather shallower slopes of TI –, just as CI-users do. With impairment, thresholds of all durations decrease, but they decrease more for longer than for short tones (Florentine, Fastl, and Buus, 1988; Plack and Skeels, 2007). Plack and Skeels additionally conclude from their measurements that the differences in integration between healthy and impaired ears cannot be only ascribed to different compression in the basilar membrane.

As a consequence, the process cannot simply be explained as an integrator of energy with a time constant (variable for each subject), as it would be unclear why a hearing impairment in the cochlea would reduce the time constant of a central integrator. In addition, it would be expected to always lead to the same slope and critical duration for different stimulation conditions. In the results presented here however, it was shown that the slope increases with electrical stimulation rate, and that the critical duration varies in an inconsistent way with rate.

An additional crucial difference observed in the slopes of integration is the effect of stimulation at different places in the cochlea. In acoustic stimulation, the slope of temporal integration has been shown to steadily fall with increasing stimulus frequency (Gerken, Bhat, and Hutchison-Clutter, 1990). This does not fit with the CI results as reported in Chapter 4, where stimulation in the basal electrodes, corresponding to positions in the cochlea that code higher frequency content, led to higher TI-slopes.

6.1.6 Temporal integration - Critical duration

In the data presented in Chapter 3 there is some indication, at least in some subjects, of an increase of threshold at higher durations. In 70 % of the conditions, the median THR amplitude at 500 ms was lower than at 1000 ms. Such an increase of THR at higher durations could indicate a later effect of adaptation with a timescale of some hundreds of milliseconds beyond that of TI. However, the small number of subjects and repetitions per condition in that dataset does not permit too fast generalisations, especially considering the observed variability in estimates of T_C.

The review of the literature on temporal integration (see Chapter 2) shows that despite very commonly cited rule-of-thumb values (e.g. 200 ms for the critical duration), the behaviour is not so simple. Different experimenters have come to very different conclusions about its magnitude and interpretations about its source, e.g. as the time constant of a leaky integrator.

In Chapter 3, the critical duration THR integration of CI wearers was estimated by curve-fitting, assuming that TI could be simply described by two straight-lines in a log-log plot as a good first approximation. This method is of course limited because there is no theoretical underpinning of the exact shape of this curve, and for that reason a simple ad-hoc shape had to be chosen for the parameter estimation. Other, more complex curves might give better fits, but at the cost of reduced parsimony and robustness. For instance, an extension of the exponential model proposed by Plomp and Bouman (1959), with an additional slope parameter m was also tested (results not shown):

$$I(t)/I_\infty = 1/\left(1 - \mathrm{e}^{-(t/\tau)^m}\right). \tag{6.1}$$

This, however, led to very unrobust fitting results, where the parameters changed too much for small variations in the datapoints. The fit with two lines proved to be a good compromise between robustness and description, with few free parameters. A more elegant way of estimating the critical duration in individuals could be achieved by extending the method described in Clauset, Shalizi, and Newman (2009, section 3.3). For plausible values of $x_{\min}$, the authors select the one that result in the best fit to the data $x \geq x_{\min}$, as estimated with Kolmogorov-Smirnov goodness-of-fit statistic.

In the review by Gerken, Bhat, and Hutchison-Clutter (1990), the existence of the critical duration is itself questioned (see e.g. Penner, 1978; Florentine, Fastl, and Buus, 1988, where no such effect was measured). Just as later done by Heil, Matysiak, and Neubauer (2017), they took a detailed look at the definition of duration. Because shorter tone bursts lead to a wider spread in the frequency domain, temporal windows have to be used. However, this slower rising and falling of the envelopes confound with the issue of duration. It might be possible that experiments with cochlear implants like the ones presented in this work shed some light on this, because electrical pulse trains do not suffer from (additional) frequency splatter for short durations, and in general, the duration of the pulse train is quite clear as a windowing is not as necessary as with acoustic stimuli.

6.2 Outlook

6.2.1 Effects of the speech processor

The present experiments were all conducted with direct stimulation, which allows the experimenter to ignore the effects of the speech processors. These include among others filtering, a non-linear compressive amplitude mapping, and dynamic automatic gain control (AGC). Especially the latter effect might play an important role in the detection of short sounds of increasing lengths, comparable to the stimuli used in the experiments. In order to clarify whether the speech processor actually increases or ameliorates the observed differences between CI wearers and normal hearing listeners, it would make sense to measure the response of speech processors to short tones with increasing envelope lengths. The RIB Detector Box would allow an analogue measurement of these responses, which could then be digitized for further analyses, especially of the amplitude envelope of the resulting pulse trains. Since the AGC systems are designed to reduce the levels of impulsive, loud stimuli, it could be expected that they would flatten the slopes at higher levels even further (BAL, MAL). In a similar way, the psychophysical experiments conducted here could be extended to acoustic stimulation in CI wearers, instead of direct stimulation. This would allow to investigate temporal integration including the context of the speech processor.

6.2.2 High stimulation rates and biophysical models

Interestingly, most existing biophysical models of human ANF are unable to respond properly at stimulation rates beyond around 5 kpps (Bachmaier et al., 2019), and even phenomenological models also break down beyond 10 kpps (van Gendt et al., 2016).

Figure 6.3 compares predictions by the biophysical human auditory nerve models from Briaire and Frijns (2005), Rattay, Lutter, and Felix (2001), and from Smit et al. (2010), which

all included a soma in their models, as well as the axon model of Imennov and Rubinstein (2009). The firing rate (as measured in the axon) is plotted as a function of stimulation rate for a 100 ms pulse train. Different lines represent different amplitudes, expressed as multiples of the pulse-train threshold.

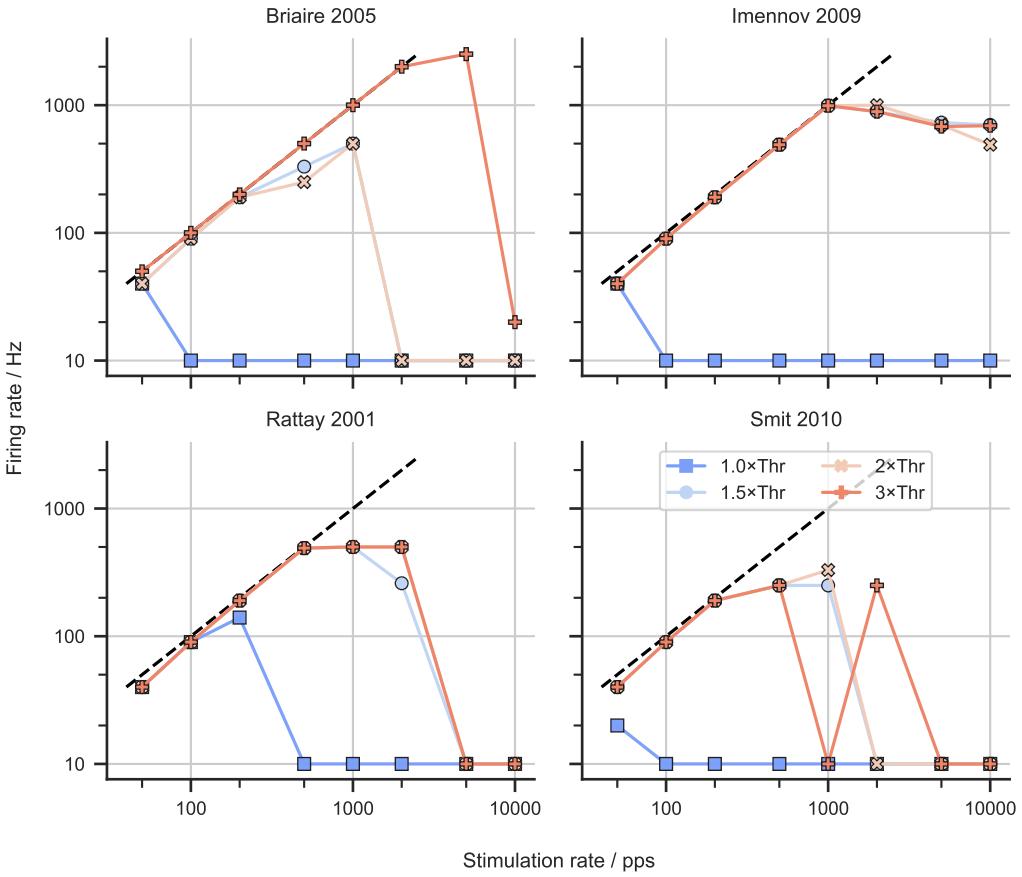


FIGURE 6.3: Firing rate of different ANF models as a function of stimulation rate at different amplitudes. Stimuli were biphasic pulse trains of 20 µs phase, negative phase leading. The black dashed line represents one action potential per stimulation pulse (100 % efficiency). Data was provided by Richard Bachmeier in private correspondence.

It can be seen that all three models with a soma fail to respond appropriately at high stimulation rates. There, the neuron fires once at stimulus onset but then blocks despite an ongoing suprathreshold stimulation. Only the model by Imennov and Rubinstein (2009) is capable of continuously responding to pulse rates of up to 10 kpps.

The ANF models do not necessarily have to reflect the process of integration for long durations in the order of hundreds of milliseconds, since this might well reflect a central process. However, they should ideally reflect the multipulse integration occurring for high rates at shorter time scales. That they do not respond after the stimulus onset at high rates conflicts strongly with the results presented in this work, where THR values dropped as a function for the highest measured rates of above 20 kpps. This is a severe limitation of the models discussed that necessarily have to be overcome by future biohphysical models of the human ANF.

6.2.3 Temporal integration - Sigmoid function & Behaviour at low durations

A peculiarity that was only sometimes observed in the temporal integration data, especially in the results of Chapter 3, and was thus not discussed in the previous chapters, was a sigmoidal-like shape to the TI function. Besides the decrease of the integration slope for higher durations – i.e. the emergence of a critical duration –, some subjects showed a similar slope decrease for short durations, near the single-pulse condition. That this was more commonly seen in the THR curves of the first experiment makes sense, since it had more duration data-points that included shorter pulse trains with 2 to 6 pulses. Figure 6.4 shows such an example curve[2].

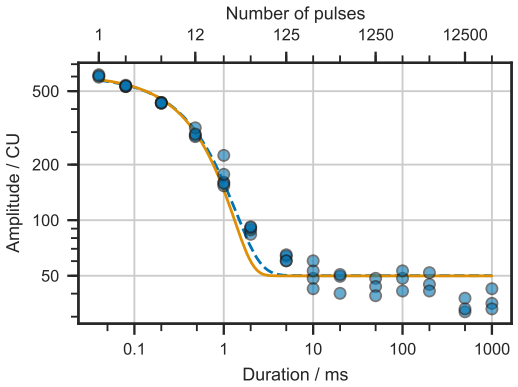


FIGURE 6.4: Example threshold data with a sigmoid shape. The blue dashed curve is a fit with Equation 6.2, the green solid with Equation 6.4.

In a preliminary report, the amplitude (I) curves were fitted with the following function:

$$I(t) = I_\infty + (I_0 - I_\infty) \cdot \exp\left(-\frac{t}{\tau}\right), \qquad (6.2)$$

where t is the duration of the pulse trains, I_∞ the asymptotic amplitude for long durations and τ is a characteristic time constant. The only new parameter is I_0, the asymptotic amplitude for decreasing durations in the limit $t \to 0$.

This fit was able to replicate the sigmoidal shape sometimes seen in the data. However, it had the shortcoming that the slope of the curve between the two asymptotes was fixed based on the I_0 and I_∞ parameters. In order to account for differing slopes, yet another parameter would need to be added to the equation, making the formula even less parsimonious:

$$I(t) = I_\infty + (I_0 - I_\infty) \cdot \exp\left[\left(-\frac{t}{\tau}\right)^m\right] \qquad (6.3)$$

There is also no theoretical underpinning for such a curve shape, while the piecewise defined function used can be seen as a linear approximation (in the double-logarithmic axes) of the smoother curve produced by e.g. a leaky integrator with a non-linear compression.

A surprisingly similar curve to Equation 6.2 with the same number of parameters can be instead obtained by using the cumulative distribution function for the normal distribution (Φ):

$$I(t) = \frac{I_0}{\Phi\left(t/\tau\right)} + (I_\infty - I_0) \qquad (6.4)$$

Here, it would be possible to interpret $1/I$ as a measure of sensitivity of the systen, in line with Crozier (1940) (see subsection 2.1.6). In contrast, here the time does not have to be logarithmised – applying the logarithm of time leads instead to a curve without a flattening. The problem with this kind of curve is that it presupposes a finite THR value for stimuli with durations tending towards 0, which seems implausible. It might well be possible that the curve does not flatten completely, but that the slope only flattens somewhat for short durations.

[2] In chapter 4, it was for simplicity assumed that a power-law function – i.e. a linear relationship between log-duration and log-amplitudes – provides a good enough explanation of the TI-functions. Neter, Wasserman, and Kutner (1985, pp 109-133), provide a possible statistical framework way to test this, as cited in Gerken, Bhat, and Hutchison-Clutter (1990). There, this analysis was used in the context of acoustic TI to "evaluate the linearity of the threshold data when plotted in log-log coordinates", which could also be applicable to the data presented here.

Bibliography

Items marked with † denote (yet) unpublished work.

Advanced Bionics Corporation (2003). *New Methodology for Fitting Cochlear Implants*. Technical report. Valencia, USA: Advanced Bionics Corporation, p. 5 (cit. on p. 4).

Algom, Daniel and Babkoff, Harvey (1984). "Auditory Temporal Integration at Threshold: Theories and Some Implications of Current Research". In: *Contributions to Sensory Physiology*. Ed. by William D. Neff. Vol. 8. 8 vols. Contributions to Sensory Physiology. New York: Acad. Press, pp. 131–159 (cit. on pp. 13–15, 17).

Arora, Komal; Dawson, Pam; Dowell, Richard, and Vandali, Andrew (2009). "Electrical Stimulation Rate Effects on Speech Perception in Cochlear Implants". In: *International Journal of Audiology* 48.8, pp. 561–567 (cit. on p. 24).

Azadpour, Mahan; McKay, Colette M., and Svirsky, Mario A. (2018). "Effect of Pulse Rate on Loudness Discrimination in Cochlear Implant Users". In: *Journal of the Association for Research in Otolaryngology* 19.3, pp. 287–299 (cit. on pp. 4, 55).

Bachmaier, Richard; Encke, Jörg; Obando-Leitón, Miguel; Hemmert, Werner, and Bai, Siwei (2019). "Comparison of multicompartment cable models of human auditory nerve fibers". In: *Frontiers in Neuroscience* 13, p. 1173 (cit. on p. 58).

Bai, Siwei; Encke, Jörg; Obando-Leitón, Miguel; Weiß, Robin; Schäfer, Friederike; Eberharter, Jakob; Böhnke, Frank, and Hemmert, Werner (2019). "Electrical Stimulation in the Human Cochlea: A Computational Study Based on High-Resolution Micro-CT Scans" (cit. on p. xiv).

Balkany, Thomas; Hodges, Annelle; Menapace, Christine; Hazard, Linda; Driscoll, Colin; Gantz, Bruce; Kelsall, David; Luxford, William; McMenomy, Sean; Neely, J. Gail; Peters, Brian; Pillsbury, Harold; Roberson, Joseph; Schramm, David; Telian, Steven; Waltzman, Susan; Westerberg, Brian, and Payne, Stacy (2007). "Nucleus Freedom North American Clinical Trial". In: *Otolaryngology–Head and Neck Surgery* 136.5, pp. 757–762 (cit. on p. 24).

Békésy, Georg v. (1947). "A New Audiometer". In: *Acta Oto-Laryngologica* 35.5-6, pp. 411–422 (cit. on p. 9).

Békésy, Georg von (1929). "Zur Theorie des Hörens; über die Bestimmung des einem reinen Tonempfinden entsprechenden Erregungsgebietes der Basilarmembran vermittelst Ermüdungserscheinungen". In: *Physikalische Zeitschrift* 30, pp. 115–125 (cit. on p. 13).

Bento, Ricardo Ferreira; De Brito Neto, Rubens Vuono; Castilho, Arthur Menino; Goffi Gomez, M Valéria Schmidt; Giorgi Sant'anna, Sandra Barreto; Guedes, Mariana Cardoso, and De Ornelas Peralta, Cristina Gomes (2005). "Psychoacoustic Dynamic Range and Cochlear Implant Speech-Perception Performance in Nucleus 22 Users". In: *Cochlear implants international* 6 (sup1), pp. 31–34 (cit. on p. 3).

Bierer, J. A.; Bierer, S. M.; Kreft, H. A., and Oxenham, A. J. (2015). "A Fast Method for Measuring Psychophysical Thresholds across the Cochlear Implant Array". In: *Trends in Hearing* 19 (cit. on pp. 5, 10).

Bonnet, Raymond M; Boermans, Peter-Paul BM; Avenarius, Otto F; Briaire, Jeroen J, and Frijns, Johan HM (2012). "Effects of Pulse Width, Pulse Rate and Paired Electrode Stimulation on Psychophysical Measures of Dynamic Range and Speech Recognition in Cochlear Implants". In: *Ear and hearing* 33.4, pp. 489–496 (cit. on pp. 4, 23 sq.).

Boulet, Jason; White, Mark, and Bruce, Ian C. (2016). "Temporal Considerations for Stimulating Spiral Ganglion Neurons with Cochlear Implants". In: *Journal of the Association for Research in Otolaryngology* 17.1, pp. 1–17 (cit. on pp. 7, 22).

Boullet, Isabelle (2004). "Comparison between Four Methods of Loudness Estimation of Stationary and Nonstationary Sounds". In: *Fortschritte Der Akustik*. 7ème Congrès Francais d'Acoustique (CFA) et 30ème Congrès de La Société Allemande d'Acoustique (DAGA). Strasbourg, France: Société Française d'Acoustique – Deutsche Gesellschaft für Akustik, pp. 257–258 (cit. on p. 10).

Brand, Thomas and Hohmann, Volker (2002). "An Adaptive Procedure for Categorical Loudness Scaling". In: *The Journal of the Acoustical Society of America* 112.4, pp. 1597–1604 (cit. on p. 57).

Briaire, Jeroen J. and Frijns, Johan H.M. (2005). "Unraveling the Electrically Evoked Compound Action Potential". In: *Hearing Research* 205.1-2, pp. 143–156 (cit. on p. 58).

Buus, Søren; Florentine, Mary, and Poulsen, Torben (1997). "Temporal Integration of Loudness, Loudness Discrimination, and the Form of the Loudness Function". In: *The journal of the acoustical society of America* 101.2, pp. 669–680 (cit. on pp. 19, 56).

Carlyon, Robert P.; Deeks, John M., and McKay, Colette M. (2015). "Effect of Pulse Rate and Polarity on the Sensitivity of Auditory Brainstem and Cochlear Implant Users to Electrical Stimulation". In: *Journal of the Association for Research in Otolaryngology* 16.5, pp. 653–668 (cit. on pp. 19 sq.).

Carlyon, Robert P.; Van Wieringen, Astrid; Deeks, John M.; Long, Christopher J.; Lyzenga, Johannes, and Wouters, Jan (2005). "Effect of Inter-Phase Gap on the Sensitivity of Cochlear Implant Users to Electrical Stimulation". In: *Hearing research* 205.1-2, pp. 210–224 (cit. on p. 21).

Chatterjee, Monita (1999). "Temporal Mechanisms Underlying Recovery from Forward Masking in Multielectrode-Implant Listeners". In: *The Journal of the Acoustical Society of America* 105.3, pp. 1853–1863 (cit. on p. 24).

Chatterjee, Monita; Fu, Qian-Jie, and Shannon, Robert V. (2000). "Effects of Phase Duration and Electrode Separation on Loudness Growth in Cochlear Implant Listeners". In: *The Journal of the Acoustical Society of America* 107.3, pp. 1637–1644 (cit. on pp. 4, 57).

Chatterjee, Monita and Kulkarni, Aditya M. (2017). "Recovery from Forward Masking in Cochlear Implant Listeners Depends on Stimulation Mode, Level, and Electrode Location". In: *The Journal of the Acoustical Society of America* 141.5, pp. 3190–3202 (cit. on p. 24).

Clauset, Aaron; Shalizi, C., and Newman, M. (2009). "Power-Law Distributions in Empirical Data". In: *SIAM Review* 51.4, pp. 661–703 (cit. on p. 58).

Colletti, Liliana; Mandalà, Marco; Shannon, Robert V., and Colletti, Vittorio (2011). "Estimated Net Saving to Society from Cochlear Implantation in Infants: A Preliminary Analysis". In: *The Laryngoscope* 121.11, pp. 2455–2460 (cit. on p. 1).

Cosendai, Grégoire and Pelizzone, Marco (2001). "Effects of the Acoustical Dynamic Range on Speech Recognition with Cochlear Implants". In: *Audiology* 40.5, pp. 272–281 (cit. on p. 3).

Crozier, William John (1940). "The Theory of the Visual Threshold: I. Time and Intensity". In: *Proceedings of the National Academy of Sciences* 26.1, pp. 54–60 (cit. on pp. 15, 17, 60).

Drennan, Ward R. and Pfingst, Bryan E. (2006). "Current-Level Discrimination in the Context of Interleaved, Multichannel Stimulation in Cochlear Implants: Effects of Number of Stimulated Electrodes, Pulse Rate, and Electrode Separation". In: *JARO: Journal of the Association for Research in Otolaryngology* 7.3, pp. 308–316 (cit. on p. 4).

Dunn, Camille C.; Tyler, Richard S.; Witt, Shelley A., and Gantz, Bruce J. (2006). "Effects of Converting Bilateral Cochlear Implant Subjects to a Strategy with Increased Rate and Number

of Channels". In: *Annals of Otology, Rhinology & Laryngology* 115.6, pp. 425–432 (cit. on p. 24).

Dynes, Scott Budd Chapman (1996). "Discharge Characteristics of Auditory Nerve Fibers for Pulsatile Electrical Stimuli". PhD Thesis. Boston: Massachusetts Institute of Technology. 125 pp. (cit. on p. 8).

Eddins, David A. and Green, David M. (1995). "Temporal Integration and Temporal Resolution". In: *Hearing*. Ed. by Brian C. J. Moore. Red. by Edward C. Carterette and Morton P. Friedman. 2nd edition. Handbook of Perception and Cognition. San Diego: Academic Press, Inc. Chap. 6, pp. 207–242 (cit. on p. 16).

Ekman, Gösta; Berglund, Birgitta, and Berglund, Ulf (1966). "Loudness as a Function of the Duration of Auditory Stimulation". In: *Scandinavian Journal of Psychology* 7.1, pp. 201–208 (cit. on p. 11).

Feldtkeller, R. and Oetinger, R. (1956). "Die Hörbarkeitsgrenzen von Impulsen Verschiedener Dauer". In: *Acta Acustica united with Acustica* 6.5, pp. 489–493 (cit. on p. 15).

Florentine, Mary; Buus, Søren, and Poulsen, Torben (1996). "Temporal Integration of Loudness as a Function of Level". In: *The Journal of the Acoustical Society of America* 99.3, pp. 1633–1644 (cit. on pp. 18, 56).

Florentine, Mary; Fastl, Hugo, and Buus, Søren (1988). "Temporal Integration in Normal Hearing, Cochlear Impairment, and Impairment Simulated by Masking". In: *The Journal of the Acoustical Society of America* 84.1, pp. 195–203 (cit. on pp. 13, 17, 19, 57 sq.).

Friesen, Lendra M.; Shannon, Robert V.; Baskent, Deniz, and Wang, Xiaosong (2001). "Speech Recognition in Noise as a Function of the Number of Spectral Channels: Comparison of Acoustic Hearing and Cochlear Implants". In: *The Journal of the Acoustical Society of America* 110.2, pp. 1150–1163 (cit. on p. 5).

Friesen, Lendra M; Shannon, Robert V, and Cruz, Rachel J (2005). "Effects of Stimulation Rate on Speech Recognition with Cochlear Implants". In: *Audiology and Neurotology* 10.3, pp. 169–184 (cit. on p. 24).

Fu, Qian-Jie (2005). "Loudness Growth in Cochlear Implants: Effect of Stimulation Rate and Electrode Configuration". In: *Hearing Research* 202.1, pp. 55–62 (cit. on p. 57).

Fu, Qian-Jie and Shannon, Robert V (2000). "Effects of Dynamic Range and Amplitude Mapping on Phoneme Recognition in Nucleus-22 Cochlear Implant Users." In: *Ear and hearing* 21.3, pp. 227–235 (cit. on p. 3).

Gallego, S.; Garnier, S.; Micheyl, C.; Truy, E.; Morgon, A., and Collet, L. (1999). "Loudness Growth Functions and EABR Characteristics in Digisonic Cochlear Implantees". In: *Acta oto-laryngologica* 119.2, pp. 234–238 (cit. on p. 57).

Galvin, John J and Fu, Qian-Jie (2009). "Influence of Stimulation Rate and Loudness Growth on Modulation Detection and Intensity Discrimination in Cochlear Implant Users". In: *Hearing research* 250.1-2, pp. 46–54 (cit. on pp. 24, 55 sq.).

Garner, W. R. and Miller, G. A. (1947). "The Masked Threshold of Pure Tones as a Function of Duration". In: *Journal of Experimental Psychology* 37.4, pp. 293–303 (cit. on pp. 14 sq., 17 sq.).

Geers, Ann; Brenner, Chris, and Davidson, Lisa (2003). "Factors Associated with Development of Speech Perception Skills in Children Implanted by Age Five". In: *Ear and Hearing* 24.1, 24S (cit. on p. 2).

Gelfand, Stanley A (2017). *Hearing: An Introduction to Psychological and Physiological Acoustics*. CRC Press (cit. on p. 9).

Gerken, George M.; Bhat, Vishwa K. H., and Hutchison-Clutter, Margaret (1990). "Auditory Temporal Integration and the Power Function Model". In: *The Journal of the Acoustical Society of America* 88.2, pp. 767–778 (cit. on pp. 13, 15, 17–19, 57 sq., 60).

Gerken, George M.; Solecki, Janet M., and Boettcher, Flint A. (1991). "Temporal Integration of Electrical Stimulation of Auditory Nuclei in Normal-Hearing and Hearing-Impaired Cat". In: *Hearing research* 53.1, pp. 101–112 (cit. on p. 57).

Gescheider, George A. (2013). *Psychophysics: The Fundamentals*. 3rd. New York: Psychology Press. 446 pp. (cit. on p. 9).

Hardie, Natalie A and Shepherd, Robert K (1999). "Sensorineural Hearing Loss during Development: Morphological and Physiological Response of the Cochlea and Auditory Brainstem". In: *Hearing Research* 128.1, pp. 147–165 (cit. on p. 6).

Heffer, L. F.; Sly, D. J.; Fallon, J. B.; White, M. W.; Shepherd, R. K., and O'Leary, S. J. (2010). "Examining the Auditory Nerve Fiber Response to High Rate Cochlear Implant Stimulation: Chronic Sensorineural Hearing Loss and Facilitation". In: *Journal of Neurophysiology* 104.6, pp. 3124–3135 (cit. on pp. 5, 8).

Heil, Peter; Matysiak, Artur, and Neubauer, Heinrich (2017). "A Probabilistic Poisson-Based Model Accounts for an Extensive Set of Absolute Auditory Threshold Measurements". In: *Hearing research* 353, pp. 135–161 (cit. on pp. 13, 18, 58).

Heil, Peter and Neubauer, Heinrich (2003). "A Unifying Basis of Auditory Thresholds Based on Temporal Summation". In: *Proceedings of the National Academy of Sciences* 100.10, pp. 6151–6156 (cit. on p. 18).

Holden, Laura K.; Reeder, Ruth M.; Firszt, Jill B., and Finley, Charles C. (2011). "Optimizing the Perception of Soft Speech and Speech in Noise with the Advanced Bionics Cochlear Implant System". In: *International Journal of Audiology* 50.4, pp. 255–269 (cit. on pp. 3 sq.).

Holden, Laura; Skinner, Margaret; Holden, Timothy, and Demorest, Marilyn (2002). "Effects of Stimulation Rate with the Nucleus 24 ACE Speech Coding Strategy". In: *Ear and Hearing* 23.5, pp. 463–476 (cit. on p. 24).

Hong, Robert S and Rubinstein, Jay T (2006). "Conditioning Pulse Trains in Cochlear Implants: Effects on Loudness Growth:" in: *Otology & Neurotology* 27.1, pp. 50–56 (cit. on p. 23).

Hughes, J. W. (1946). "The Threshold of Audition for Short Periods of Stimulation". In: *Proceedings of the Royal Society of London B* 133.873, pp. 486–490 (cit. on p. 14).

Hughes, Michelle L.; Castioni, Erin E.; Goehring, Jenny L., and Baudhuin, Jacquelyn L. (2012). "Temporal Response Properties of the Auditory Nerve: Data from Human Cochlear-Implant Recipients". In: *Hearing Research* 285.1–2, pp. 46–57 (cit. on pp. 8, 23).

Hughes, Michelle L.; Goehring, Jenny L.; Baudhuin, Jacquelyn L., and Schmid, Kendra K. (2016). "Effects of Stimulus Level and Rate on Psychophysical Thresholds for Interleaved Pulse Trains in Cochlear Implants". In: *The Journal of the Acoustical Society of America* 140.4, pp. 2297–2311 (cit. on pp. 3, 5).

Ifukube, Tohru and White, Robert L (1987). "Current Distributions Produced inside and Outside the Cochlea from a Scala Tympani Electrode Array". In: *IEEE transactions on biomedical engineering* 34.11, pp. 883–890 (cit. on p. 6).

Imennov, N.S. and Rubinstein, J.T. (2009). "Stochastic Population Model for Electrical Stimulation of the Auditory Nerve". In: *IEEE Transactions on Biomedical Engineering* 56.10, pp. 2493–2501 (cit. on p. 59).

ISO (2006). *Acoustics – Loudness Scaling by Means of Categories*. Standard 16832. Geneva, CH: International Organization for Standardization (cit. on p. 57).

Javel, Eric (1990). "Acoustic and Electrical Encoding of Temporal Information". In: *Cochlear Implants*. Springer, pp. 247–295 (cit. on p. 6).

Javel, Eric and Shepherd, Robert K. (2000). "Electrical Stimulation of the Auditory Nerve: III. Response Initiation Sites and Temporal Fine Structure". In: *Hearing Research* 140.1, pp. 45–76 (cit. on pp. 7, 22).

Javel, Eric and Viemeister, Neal F. (2000). "Stochastic Properties of Cat Auditory Nerve Responses to Electric and Acoustic Stimuli and Application to Intensity Discrimination". In: *The Journal of the Acoustical Society of America* 107.2, pp. 908–921 (cit. on p. 6).

Karg, Sonja A.; Lackner, Christina, and Hemmert, Werner (2013). "Temporal Interaction in Electrical Hearing Elucidates Auditory Nerve Dynamics in Humans". In: *Hearing Research* 299, pp. 10–18 (cit. on p. 21).

Koch, Dawn Burton; Osberger, Mary Joe; Segel, Phil, and Kessler, Dorcas (2004). "HiResolutionTM and Conventional Sound Processing in the HiResolutionTM Bionic Ear: Using Appropriate Outcome Measures to Assess Speech Recognition Ability". In: *Audiology and Neurotology* 9.4, pp. 214–223 (cit. on p. 24).

Kreft, Heather A.; Donaldson, Gail S., and Nelson, David A. (2004a). "Effects of Pulse Rate and Electrode Array Design on Intensity Discrimination in Cochlear Implant Users". In: *The Journal of the Acoustical Society of America* 116.4, pp. 2258–2268 (cit. on p. 56).

– (2004b). "Effects of Pulse Rate on Threshold and Dynamic Range in Clarion Cochlear-Implant Users (L)". In: *The Journal of the Acoustical Society of America* 115.5, pp. 1885–1888 (cit. on p. 23).

Krumbholz, Katrin and Wiegrebe, Lutz (1998). "Detection Thresholds for Brief Sounds – Are They a Measure of Auditory Intensity Integration?" In: *Hearing Research* 124.1, pp. 155–169 (cit. on p. 17).

Kucharski, P. (1927). "I. Recherches sur l'excitabilité auditive en fonction du temps". In: *L'Année psychologique* 28.1, pp. 1–74 (cit. on p. 13).

Landsberger, David and Galvin, John J. (2011). "Discrimination between Sequential and Simultaneous Virtual Channels with Electrical Hearing". In: *The Journal of the Acoustical Society of America* 130.3, pp. 1559–1566 (cit. on p. 4).

Levitt, HCCH (1971). "Transformed Up-down Methods in Psychoacoustics". In: *The Journal of the Acoustical society of America* 49 (2B), pp. 467–477 (cit. on p. 9).

Liberman, M. Charles (1978). "Auditory-nerve Response from Cats Raised in a Low-noise Chamber". In: *The Journal of the Acoustical Society of America* 63.2, pp. 442–455 (cit. on p. 6).

Litovsky, Ruth Y.; Goupell, Matthew J.; Kan, Alan, and Landsberger, David M. (2017). "Use of Research Interfaces for Psychophysical Studies with Cochlear-Implant Users". In: *Trends in Hearing* 21, p. 2331216517736464 (cit. on p. 11).

Litvak, Leonid M.; Smith, Zachary M.; Delgutte, Bertrand, and Eddington, Donald K. (2003). "Desynchronization of Electrically Evoked Auditory-Nerve Activity by High-Frequency Pulse Trains of Long Duration". In: *The Journal of the Acoustical Society of America* 114.4, pp. 2066–2078 (cit. on p. 8).

Litvak, Leonid; Delgutte, Bertrand, and Eddington, Donald (2001). "Auditory Nerve Fiber Responses to Electric Stimulation: Modulated and Unmodulated Pulse Trains". In: *The Journal of the Acoustical Society of America* 110.1, pp. 368–379 (cit. on pp. 8, 22 sq.).

Luo, Xin; Fu, Qian-Jie, and Galvin III, John J. (2007). "Vocal Emotion Recognition by Normal-Hearing Listeners and Cochlear Implant Users". In: *Trends in Amplification* 11.4, pp. 301–315 (cit. on p. 2).

Macherey, Olivier; Wieringen, Astrid van; Carlyon, Robert P.; Deeks, John M., and Wouters, Jan (2006). "Asymmetric Pulses in Cochlear Implants: Effects of Pulse Shape, Polarity, and Rate". In: *Journal of the Association for Research in Otolaryngology* 7.3, pp. 253–266 (cit. on p. 21).

Malherbe, T. K.; Hanekom, T., and Hanekom, J. J. (2016). "Constructing a Three-Dimensional Electrical Model of a Living Cochlear Implant User's Cochlea". In: *International journal for numerical methods in biomedical engineering* 32.7, e02751 (cit. on p. 6).

Marks, Lawrence E. and Florentine, Mary (2011). "Measurement of Loudness, Part I: Methods, Problems, and Pitfalls". In: *Loudness*. Ed. by Mary Florentine; Arthur N. Popper, and Richard R. Fay. Red. by Richard R. Fay and Arthur N. Popper. Springer Handbook of Auditory Research. Springer, New York, NY, pp. 17–56 (cit. on p. 10).

McKay, Colette M.; Henshall, Katherine R.; Farrell, Rebecca J., and McDermott, Hugh J. (2003). "A Practical Method of Predicting the Loudness of Complex Electrical Stimuli". In: *The Journal of the Acoustical Society of America* 113.4, pp. 2054–2063 (cit. on pp. 6, 21 sq., 56).

McKay, Colette M.; Lim, Hubert H., and Lenarz, Thomas (2013). "Temporal Processing in the Auditory System". In: *Journal of the Association for Research in Otolaryngology* 14.1, pp. 103–124 (cit. on pp. 20 sq.).

McKay, Colette M. and McDermott, Hugh J. (1998). "Loudness Perception with Pulsatile Electrical Stimulation: The Effect of Interpulse Intervals". In: *The Journal of the Acoustical Society of America* 104.2, pp. 1061–1074 (cit. on pp. 7, 19, 21).

McKay, Colette M.; Remine, Maria D., and McDermott, Hugh J. (2001). "Loudness Summation for Pulsatile Electrical Stimulation of the Cochlea: Effects of Rate, Electrode Separation, Level, and Mode of Stimulation". In: *The Journal of the Acoustical Society of America* 110.3, pp. 1514–1524 (cit. on pp. 21 sq.).

Meddis, Ray and Lecluyse, Wendy (2011). "The Psychophysics of Absolute Threshold and Signal Duration: A Probabilistic Approach". In: *The Journal of the Acoustical Society of America* 129.5, pp. 3153–3165 (cit. on pp. 14, 18).

Middlebrooks, John C. (2004). "Effects of Cochlear-Implant Pulse Rate and Inter-Channel Timing on Channel Interactions and Thresholds". In: *The Journal of the Acoustical Society of America* 116.1, p. 452 (cit. on p. 20).

Miller, Charles A.; Abbas, Paul J., and Robinson, Barbara K. (2001). "Response Properties of the Refractory Auditory Nerve Fiber". In: *JARO: Journal of the Association for Research in Otolaryngology* 2.3, pp. 216–232 (cit. on p. 7).

Miller, Charles A.; Hu, Ning; Zhang, Fawen; Robinson, Barbara K., and Abbas, Paul J. (2008). "Changes Across Time in the Temporal Responses of Auditory Nerve Fibers Stimulated by Electric Pulse Trains". In: *Journal of the Association for Research in Otolaryngology* 9.1, pp. 122–137 (cit. on pp. 22, 55).

Miller, Charles A.; Woo, Jihwan; Abbas, Paul J.; Hu, Ning, and Robinson, Barbara K. (2011). "Neural Masking by sub-Threshold Electric Stimuli: Animal and Computer Model Results". In: *Journal of the Association for Research in Otolaryngology* 12.2, pp. 219–232 (cit. on p. 8).

Moore, Brian C. J.; Glasberg, Brian R.; Plack, C. J., and Biswas, A. K. (1988). "The Shape of the Ear's Temporal Window". In: *The Journal of the Acoustical Society of America* 83.3, pp. 1102–1116 (cit. on p. 16).

Moxon, E. C. (1967). "Electric Stimulation of the Cat's Cochlea: A Study of Discharge Rates in Single Auditory Nerve Fibers". PhD Thesis. Cambridge, MA: MIT (cit. on p. 22).

Mudry, Albert and Mills, Mara (2013). "The Early History of the Cochlear Implant: A Retrospective". In: *JAMA Otolaryngology–Head & Neck Surgery* 139.5, pp. 446–453 (cit. on p. 2).

Munson, W. A. (1947). "The Growth of Auditory Sensation". In: *The Journal of the Acoustical Society of America* 19.4, pp. 584–591 (cit. on p. 15).

National Institute on Deafness and Other Communication Disorders (2013). *Cochlear Implants*. Fact Sheet 11–4798. United States: National Institutes of Health, p. 2 (cit. on p. 2).

– (2016). *Cochlear Implants*. Fact Sheet 00-4798. United States: National Institutes of Health, p. 2 (cit. on p. 2).

Nelson, David A. and Donaldson, Gail S. (2002). "Psychophysical Recovery from Pulse-Train Forward Masking in Electric Hearing". In: *The Journal of the Acoustical Society of America* 112.6, pp. 2932–2947 (cit. on p. 24).

Neter, John; Wasserman, William, and Kutner, Michael H. (1985). *Applied Linear Statistical Models: Regression, Analysis of Variance, and Experimental Design*. 2nd. Homewood, IL: Richard D. Irwin. 1127 pp. (cit. on p. 60).

Niese, Herbert (1959). "Die Trägheit der Lautstärkebildung in Abhängigkeit vom Schallpegel". In: *Hochfrequenztechnik und Elektroakustik* 68.5, pp. 143–152 (cit. on p. 14).

Oxenham, Andrew J. (2001). "Forward Masking: Adaptation or Integration?" In: *The Journal of the Acoustical Society of America* 109.2, pp. 732–741 (cit. on p. 21).

Oxenham, Andrew J. and Moore, Brian C. J. (1994). "Modeling the Additivity of Nonsimultaneous Masking". In: *Hearing Research* 80.1, pp. 105–118 (cit. on p. 21).

Oxenham, Andrew J.; Moore, Brian CJ, and Vickers, Deborah A. (1997). "Short-Term Temporal Integration: Evidence for the Influence of Peripheral Compression". In: *The Journal of the Acoustical Society of America* 101.6, pp. 3676–3687 (cit. on p. 19).

Pelosi, Stanley; Rivas, Alejandro; Haynes, David S.; Bennett, Marc L.; Labadie, Robert F.; Hedley-Williams, Andrea, and Wanna, George B. (2012). "Stimulation Rate Reduction and Auditory Development in Poorly Performing Cochlear Implant Users with Auditory Neuropathy". In: *Otology & Neurotology* 33.9, pp. 1502–1506 (cit. on p. 24).

Penner, Merrilynn J. (1978). "A Power Law Transformation Resulting in a Class of Short-Term Integrators That Produce Time-Intensity Trades for Noise Bursts". In: *The Journal of the Acoustical Society of America* 63.1, pp. 195–202 (cit. on pp. 15–17, 58).

Pfingst, Bryan E.; Bowling, Sara A.; Colesa, Deborah J.; Garadat, Soha N.; Raphael, Yehoash; Shibata, Seiji B.; Strahl, Stefan B.; Su, Gina L., and Zhou, Ning (2011). "Cochlear Infrastructure for Electrical Hearing". In: *Hearing research* 281.1-2, pp. 65–73 (cit. on p. 6).

Pfingst, Bryan E.; Colesa, Deborah J.; Hembrador, Sheena; Kang, Stephen Y.; Middlebrooks, John C.; Raphael, Yehoash, and Su, Gina L. (2011). "Detection of Pulse Trains in the Electrically Stimulated Cochlea: Effects of Cochlear Health". In: *The Journal of the Acoustical Society of America* 130.6, pp. 3954–3968 (cit. on pp. 19 sq.).

Pfingst, Bryan E. and Xu, Li (2004). "Across-Site Variation in Detection Thresholds and Maximum Comfortable Loudness Levels for Cochlear Implants". In: *Journal of the Association for Research in Otolaryngology* 5.1, pp. 11–24 (cit. on p. 10).

Pickles, James O. (2012). *An Introduction to the Physiology of Hearing*. 4th edition. Bingley: Emerald Book Serials and Monographs. 460 pp. (cit. on pp. 5–7).

Plack, C. J.; Oxenham, A. J., and Drga, V. (2002). "Linear and Nonlinear Processes in Temporal Masking". In: *Acta Acustica united with Acustica* 88.3, pp. 348–358 (cit. on p. 21).

Plack, Christopher J. and Skeels, Vicki (2007). "Temporal Integration and Compression near Absolute Threshold in Normal and Impaired Ears". In: *The Journal of the Acoustical Society of America* 122.4, p. 2236 (cit. on pp. 13, 18 sq., 57).

Plant, Kerrie; Holden, Laura; Skinner, Margo; Arcaroli, Jennifer; Whitford, Lesley; Law, Mary-Ann, and Nel, Esti (2007). "Clinical Evaluation of Higher Stimulation Rates in the Nucleus Research Platform 8 System". In: *Ear and hearing* 28.3, pp. 381–393 (cit. on p. 24).

Plomp, R. and Bouman, M. A. (1959). "Relation between Hearing Threshold and Duration for Tone Pulses". In: *The Journal of the Acoustical Society of America* 31.6, pp. 749–758 (cit. on pp. 15, 17, 58).

Poulton, E. C. (1989). *Bias in Quantifying Judgements*. Hove: Lawrence Erlbaum. 304 pp. (cit. on p. 11).

Rader, Tobias; Doms, Philipp; Adel, Youssef; Weissgerber, Tobias; Strieth, Sebastian, and Baumann, Uwe (2018). "A Method for Determining Precise Electrical Hearing Thresholds in Cochlear Implant Users". In: *International journal of audiology*, pp. 1–8 (cit. on pp. 3, 10, 24).

Rattay, Frank; Lutter, Petra, and Felix, Heidi (2001). "A Model of the Electrically Excited Human Cochlear Neuron I. Contribution of Neural Substructures to the Generation and Propagation of Spikes". In: *Hearing Research*, p. 21 (cit. on p. 58).

Reichardt, W. (1970). "Subjective and Objective Measurement of the Loudness Level of Single and Repeated Impulses". In: *The Journal of the Acoustical Society of America* 47 (6B), pp. 1557–1562 (cit. on p. 14).

Robinson, Charles E. (1974). "Simple Form of the Auditory Running-average Hypothesis: Application to the Temporal Summation of Loudness and to the Delayed Perception of the Offset of Brief Stimuli". In: *The Journal of the Acoustical Society of America* 55.3, pp. 645–648 (cit. on p. 15).

Rubinstein, J. T; Wilson, B. S; Finley, C. C, and Abbas, P. J (1999). "Pseudospontaneous Activity: Stochastic Independence of Auditory Nerve Fibers with Electrical Stimulation". In: *Hearing Research* 127.1, pp. 108–118 (cit. on p. 23).

Scharf, Bertram (1961). "Loudness Summation under Masking". In: *The Journal of the Acoustical Society of America* 33.4, pp. 503–511 (cit. on p. 11).

Scharf, Bertram and Houtsma, Adrianus J. M. (1986). "Audition II: Loudness, Pitch, Localization, Aural Distortion, Pathology". In: *Handbook of Perception and Human Performance*. Ed. by Kenneth R. Boff; Lloyd Kaufman, and James P. Thomas. Vol. 1: Sensory Processes and Performance. 2 vols. New York: Wiley-Interscience, pp. 15-1–15-60 (cit. on p. 14).

Shah, Saurabh B.; Chung, Jeannie H., and Jackler, Robert K. (1997). "Lodestones, Quackery, and Science: Electrical Stimulation of the Ear Before Cochlear Implants". In: *Otology & Neurotology* 18.5, p. 665 (cit. on p. 2).

† Shannon, R. V. (2011). "Psychophysics of Electric Stimulation of the Human Cochlea, Brainstem, and Midbrain". In: Association for Research in Otolaryngology, 34th Midwinter Research Meeting. Vol. 34. Baltimore, Maryland, USA, pp. 99–100 (cit. on p. 20).

Shannon, Robert V. (1983). "Multichannel Electrical Stimulation of the Auditory Nerve in Man. I. Basic Psychophysics". In: *Hearing Research* 11.2, pp. 157–189 (cit. on pp. 4, 57).

– (1985). "Threshold and Loudness Functions for Pulsatile Stimulation of Cochlear Implants". In: *Hearing Research* 18.2, pp. 135–143 (cit. on p. 19).

– (1989). "A Model of Threshold for Pulsatile Electrical Stimulation of Cochlear Implants". In: *Hearing Research* 40.3, pp. 197–204 (cit. on p. 20).

– (1990). "Forward Masking in Patients with Cochlear Implants". In: *The Journal of the Acoustical Society of America* 88.2, pp. 741–744 (cit. on p. 24).

Shepherd, Robert K. and Hardie, Natalie A. (2001). "Deafness-Induced Changes in the Auditory Pathway: Implications for

Cochlear Implants". In: *Audiology and Neurotology* 6.6, pp. 305–318 (cit. on p. 6).

Smit, Jacoba E.; Hanekom, Tania; van Wieringen, Astrid; Wouters, Jan, and Hanekom, Johan J. (2010). "Threshold Predictions of Different Pulse Shapes Using a Human Auditory Nerve Fibre Model Containing Persistent Sodium and Slow Potassium Currents". In: *Hearing Research* 269.1-2, pp. 12–22 (cit. on p. 58).

Snyder, Russell L; Middlebrooks, John C, and Bonham, Ben H (2008). "Cochlear Implant Electrode Configuration Effects on Activation Threshold and Tonotopic Selectivity". In: *Hearing research* 235.1, pp. 23–38 (cit. on p. 5).

Stephens, S. D. G. (1974). "Methodological Factors Influencing Loudness of Short Duration Sounds". In: *Journal of Sound and Vibration* 37.2, pp. 235–246 (cit. on p. 11).

Stevens, Joseph C and Hall, James W (1966). "Brightness and Loudness as Functions of Stimulus Duration". In: *Perception & Psychophysics* 1.9, pp. 319–327 (cit. on p. 14).

Stevens, Stanley S. (1955). "The Measurement of Loudness". In: *The Journal of the Acoustical Society of America* 27.5, pp. 815–829 (cit. on p. 16).

Turner, C.W.; Horwitz, A.R., and Souza, P.E. (1994). "Forward- and Backward-Masked Intensity Discrimination Measured Using Forced-Choice and Adjustment Procedures". In: *Journal of the Acoustical Society of America* 96.4, pp. 2121–2126 (cit. on p. 10).

Vaerenberg, Bart; Smits, Cas; De Ceulaer, Geert; Zir, Elie; Harman, Sally; Jaspers, Nadine; Tam, Y.; Dillon, Margaret; Wesarg, Thomas, and Martin-Bonniot, D. (2014). "Cochlear Implant Programming: A Global Survey on the State of the Art". In: *The Scientific World Journal* 2014 (cit. on pp. 3, 10).

Valencia, David M.; Rimell, Frank L.; Friedman, Barbara J.; Oblander, Melisa R., and Helmbrecht, Josephine (2008). "Cochlear Implantation in Infants Less than 12 Months of Age". In: *International journal of pediatric otorhinolaryngology* 72.6, pp. 767–773 (cit. on p. 1).

Van den Honert, C. and Stypulkowski, P. H. (1984). "Physiological Properties of the Electrically Stimulated Auditory Nerve. II. Single Fiber Recordings". In: *Hearing Research* 14.3, pp. 225–243 (cit. on pp. 6 sq., 22).

– (1986). "Characterization of the Electrically Evoked Auditory Brainstem Response (ABR) in Cats and Humans". In: *Hearing Research* 21.2, pp. 109–126 (cit. on pp. 6 sq.).

Van Der Beek, Feddo B.; Briaire, Jeroen J., and Frijns, Johan HM (2015). "Population-Based Prediction of Fitting Levels for Individual Cochlear Implant Recipients". In: *Audiology and Neurotology* 20.1, pp. 1–16 (cit. on pp. 3, 24).

Van Gendt, M.J.; Briaire, J.J.; Kalkman, R.K., and Frijns, J.H.M. (2016). "A Fast, Stochastic, and Adaptive Model of Auditory Nerve Responses to Cochlear Implant Stimulation". In: *Hearing Research* 341, pp. 130–143 (cit. on pp. 8, 58).

Van Wieringen, Astrid and Wouters, Jan (2001). "Comparison of Procedures to Determine Electrical Stimulation Thresholds in Cochlear Implant Users". In: *Ear and Hearing* 22.6, p. 528 (cit. on p. 10).

Vandali, Andrew E; Whitford, Lesley A; Plant, Kerrie L, and Clark, Graeme M (2000). "Speech Perception as a Function of Electrical Stimulation Rate: Using the Nucleus 24 Cochlear Implant System". In: *Ear and hearing* 21.6, pp. 608–624 (cit. on p. 24).

Verhey, Jesko L. and Uhlemann, Michael (2008). "Spectral Loudness Summation for Sequences of Short Noise Bursts". In: *The Journal of the Acoustical Society of America* 123.2, pp. 925–934 (cit. on p. 14).

Viemeister, Neal F. and Wakefield, Gregory H. (1991). "Temporal Integration and Multiple Looks". In: *The Journal of the Acoustical Society of America* 90.2, pp. 858–865 (cit. on pp. 16 sq.).

Volta, Alessandro (1800). "XVII. On the Electricity Excited by the Mere Contact of Conducting Substances of Different Kinds. In a Letter from Mr. Alexander Volta, F.R.S. Professor of Natural Philosophy in the University of Pavia, to the Rt. Hon. Sir Joseph Banks, Bart. K.B. P.R.S." In: *Philosophical transactions of the Royal Society of London* 90, pp. 403–431 (cit. on p. 2).

Vongphoe, Michael and Zeng, Fan-Gang (2005). "Speaker Recognition with Temporal Cues in Acoustic and Electric Hearing". In: *The Journal of the Acoustical Society of America* 118.2, pp. 1055–1061 (cit. on p. 2).

Webster, Molly and Webster, Douglas B. (1981). "Spiral Ganglion Neuron Loss Following Organ of Corti Loss: A Quantitative Study". In: *Brain Research* 212.1, pp. 17–30 (cit. on p. 6).

Wier, Craig C.; Jesteadt, Walt, and Green, David M. (1976). "A Comparison of Method-of-Adjustment and Forced-Choice Procedures in Frequency Discrimination". In: *Perception & Psychophysics* 19.1, pp. 75–79 (cit. on p. 10).

Wilson, Blake S.; Finley, Charles C.; Lawson, Dewey T.; Wolford, Robert D.; Eddington, Donald K., and Rabinowitz, William M. (1991). "Better Speech Recognition with Cochlear Implants". In: *Nature* 352.6332, pp. 236–238 (cit. on p. 2).

Wilson, Blake S.; Finley, Charles C.; Lawson, Dewey T., and Zerbi, Mariangeli (1997). "Temporal Representations with Cochlear Implants." In: *Otology & Neurotology* 18.6, S30–S534 (cit. on p. 23).

Wolfe, Jace and Schafer, Erin (2014). *Programming Cochlear Implants*. Second Edition. Core Clinical Concepts in Audiology. San Diego, CA: Plural Publishing, Inc. 409 pp. (cit. on pp. 1, 3, 5, 24).

Zeng, Fan-Gang; Galvin III, John J., and Zhang, Chaoying (1998). "Encoding Loudness by Electric Stimulation of the Auditory Nerve". In: *Neuroreport* 9.8, pp. 1845–1848 (cit. on p. 4).

Zeng, Fan-Gang; Rebscher, Stephen; Harrison, William V.; Sun, Xiaoan, and Feng, Haihong (2008). "Cochlear Implants: System Design, Integration, and Evaluation". In: *IEEE reviews in biomedical engineering* 1, pp. 115–142 (cit. on pp. 3, 24).

Zeng, Fan-Gang and Shannon, Robert V. (1994). "Loudness-Coding Mechanisms Inferred from Electric Stimulation of the Human Auditory System". In: *Science* 264.5158, pp. 564–566 (cit. on p. 57).

Zhang, Fawen; Miller, Charles A; Robinson, Barbara K; Abbas, Paul J, and Hu, Ning (2007). "Changes across Time in Spike Rate and Spike Amplitude of Auditory Nerve Fibers Stimulated by Electric Pulse Trains". In: *Journal of the Association for Research in Otolaryngology* 8.3, pp. 356–372 (cit. on pp. 7 sq., 22 sq., 55).

Zhou, Ning; Kraft, Casey T.; Colesa, Deborah J., and Pfingst, Bryan E. (2015). "Integration of Pulse Trains in Humans and Guinea Pigs with Cochlear Implants". In: *Journal of the Association for Research in Otolaryngology* 16.4, pp. 523–534 (cit. on pp. 8, 10, 19).

Zhou, Ning and Pfingst, Bryan E. (2016a). "Evaluating Multipulse Integration as a Neural-Health Correlate in Human Cochlear-Implant Users: Relationship to Forward-Masking Recovery". In: *The Journal of the Acoustical Society of America* 139.3, EL70–EL75 (cit. on pp. 10 sq.).

– (2016b). "Evaluating Multipulse Integration as a Neural-Health Correlate in Human Cochlear-Implant Users: Relationship to Spatial Selectivity". In: *The Journal of the Acoustical Society of America* 140.3, pp. 1537–1547 (cit. on p. 11).

Zhou, Ning; Xu, Li, and Pfingst, Bryan E. (2012). "Characteristics of Detection Thresholds and Maximum Comfortable Loudness Levels as a Function of Pulse Rate in Human Cochlear Implant Users". In: *Hearing Research* 284.1, pp. 25–32 (cit. on p. 23).

Zwicker, E. and Wright, H. N. (1963). "Temporal Summation for Tones in Narrow-Band Noise". In: *The Journal of the Acoustical Society of America* 35.5, pp. 691–699 (cit. on p. 15).

Zwicker, Eberhard (1958). "Über Psychologische Und Methodische Grundlagen Der Lautheit". In: *Acustica* 8.4, pp. 237–258 (cit. on p. 11).

Zwicker, Eberhard and Fastl, Hugo (1999). *Psychoacoustics*. Ed. by Thomas S. Huang; Teuvo Kohonen; Manfred R. Schroeder, and Helmut K. V. Lotsch. Vol. 22. Springer Series in Information Sciences. Berlin, Heidelberg: Springer Berlin Heidelberg (cit. on p. 9).

Zwislocki, J. (1960). "Theory of Temporal Auditory Summation". In: *The Journal of the Acoustical Society of America* 32.8, pp. 1046–1060 (cit. on pp. 15–17, 57).

Zwislocki, Josef J. (1969). "Temporal Summation of Loudness: An Analysis". In: *The Journal of the Acoustical Society of America* 46 (2B), pp. 431–441 (cit. on pp. 14 sq., 17).

Zwolan, Teresa A.; O'Sullivan, Mary Beth; Fink, Nancy E.; Niparko, John K., and Team, The CDACI Investigative (2008). "Electric Charge Requirements of Pediatric Cochlear Implant Recipients Enrolled in the Childhood Development After Cochlear Implantation Study". In: *Otology & Neurotology* 29.2, p. 143 (cit. on p. 4).

List of publications

Items marked with † denote submitted manuscripts, [P)] denotes poster presentations.

Bachmaier, Richard; Encke, Jörg; Obando-Leitón, Miguel; Hemmert, Werner, and Bai, Siwei (2019). "Comparison of multi-compartment cable models of human auditory nerve fibers". In: *Frontiers in Neuroscience* 13, p. 1173.

Bai, Siwei; Encke, Jörg; Obando-Leitón, Miguel; Weiß, Robin; Schäfer, Friederike; Eberharter, Jakob; Böhnke, Frank, and Hemmert, Werner (2019). "Electrical Stimulation in the Human Cochlea: A Computational Study Based on High-Resolution Micro-CT Scans".

† Obando, Miguel; Dietze, Anna, and Hemmert, Werner (2019). "On the Effect of High Stimulation Rates on Temporal Integration in Cochlear Implant Users". Manuscript in review. München.

[P)] Obando, Miguel; Schwanda, Daniela, and Hemmert, Werner (2018). "Pulse Train Integration in CI-Wearers and Effects of High Stimulation Rates". In: *41st ARO Annual MidWinter Meeting*. 41st ARO Annual MidWinter Meeting. Vol. 41. San Diego, California, USA, p. 34.

† – (2019). "Exploratory Measurements of Loudness Integration in Cochlear Implant Users at High Rates". Manuscript in review. München.

Wykowska, Agnieszka; Kajopoulos, Jasmin; Obando-Leitón, Miguel; Chauhan, Sushil Singh; Cabibihan, John-John, and Cheng, Gordon (2015). "Humans Are Well Tuned to Detecting Agents among Non-Agents: Examining the Sensitivity of Human Perception to Behavioral Characteristics of Intentional Systems". In: *International Journal of Social Robotics* 7.5, pp. 767–781.

Affidavit

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation *Temporal integration in cochlear implants and the effect of high pulse rates* selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation *Temporal integration in cochlear implants and the effect of high pulse rates* is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

München, den 12. Dezember 2019 Miguel Eduardo Obando Leiton

Author contributions

Chapter 3 – *Exploratory measurements of loudness integration in cochlear implant users at high rates*

Miguel Obando designed the computational framework for the experiment, planned the experiments, analysed the data, and wrote the manuscript with support of D. S.

Daniela Schwanda carried out the measurements, ran first analyses, aided in literature review, and acquired subjects.

Werner Hemmert conceived the study, was in charge of overall direction and funding acquisition, supervised the study, and commented on the manuscript.

Chapter 4 – On the Effect of High Stimulation Rates on Temporal Integration in Cochlear Implant Users

Miguel Obando planned the experiments, wrote the software for acquiring data, acquired subjects, carried out part of the experiments, and reviewed and edited the manuscript.

Anna Dietze carried out part of the experiments, processed the experimental data, performed the statistical analyses, drafted the manuscript, and designed the figures.

Werner Hemmert conceived the experiments, aided in interpreting the results and in acquiring subjects, was in charge of overall direction and funding acquisition, supervised the study, and commented on the manuscript.

Chapter 5 – *Electrode Impedance and Voltage Spread*

Miguel Obando planned the clinical experiment, wrote the software for data acquisition, wrote the manuscript and acquired subjects.

Jakob Eberharter performed the measurements and ran the first analyses.

Siwei Bai directed the development of the model, provided the data from the model simulation and commented on the manuscript.

Werner Hemmert supervised the project.

Miguel Eduardo Obando Leitón

Prof. Dr.-Ing. Werner Hemmert