Temporal charge interactions in cochlear implant listeners

Guérit, François

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

CONTRIBUTIONS TO
HEARING RESEARCH

Volume 34

François Guérit

Temporal charge interactions in cochlear implant listeners

DTU Electrical Engineering
Department of Electrical Engineering

Temporal charge interactions in cochlear implant listeners

PhD thesis by
François Guérit

Technical University of Denmark

2018

The defense was held on June 26, 2018.
Cover illustration by Anna-Julia Plichta.

This PhD dissertation is the result of a research project carried out at the Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark. Parts of the project were carried out at the MRC Cognition and Brain Sciences Unit, University of Cambridge.

The project was financed by the Oticon Centre of Excellence for Hearing and Speech Sciences (2/3) and by the Technical University of Denmark (1/3).

Supervisors

Assoc. Prof. Jeremy Marozeau
Assoc. Prof. Bastian Epp
Hearing Systems Group
Department of Electrical Engineering
Technical University of Denmark
Kgs. Lyngby, Denmark

External Advisor

Dr. Robert P. Carlyon
MRC Cognition and Brain Sciences Unit
University of Cambridge
Cambridge, United Kingdom

Abstract

More than half a million people with severe to profound hearing loss have a cochlear implant (CI). These are surgically implanted devices, which can restore partial hearing by electrically stimulating the auditory nerve. Many CI users understand speech well in quiet, but even the most successful struggle in noisy situations, with a minority performing poorly even in quiet.

Not much is known about the integration of electrical currents by the auditory nerve. A reason for this stems from the clinical use of charge-balanced biphasic pulses. These consist of two phases with opposite-polarity currents, and create complex patterns of activation and cancellation at the level of the neurons.

This thesis investigated the polarity-specific temporal integration of currents. This was achieved by measuring the loudness, detection thresholds and localization abilities of CI users with pulse pairs, while varying the inter-pulse interval and the polarity of each pulse.

Overall, results showed a variety of inter-pulse and polarity effects, as well as interactions between both factors. These results are not only relevant for clinical applications, such as the estimation of neural survival in the auditory nerve, but also for the development of models of the electrically activated auditory nerve.

Resumé

Mere end en halv million mennesker med alvorligt til svært høretab anvender et Cochlear Implantat (CI). Disse er kirurgisk implanterede apparater, som delvist genskaber hørelsen ved at stimulere hørenerven elektrisk. De fleste CI brugere er gode til at forstå tale i stilhed, men selv de mest succesfulde har problemer i støjfyldte miljøer. En minoritet har dog problemer selv i stilhed.

Meget er stadig ukendt om hvordan hørenerven omsætter elektrisk strøm til nervesignaler. Dette skyldes til dels brugen af bifasiske ladnings-balancerede pulser i kliniske sammenhænge. Disse består af to faser med strøm i omvendt polaritet, som skaber komplekse mønstre af aktivering og deaktivering i nerverne.

Denne afhandling undersøgte effekten af polaritets-specifik tidsmæssige integration af strøm. Dette var gjort ved at måle lydstyrke, høretærskel og lydlokalisering hos CI brugere med puls-par hvor inter-puls interval og polaritet af hver puls var varieret.

Resultaterne viser flere forskellige effekter af inter-puls interval og polaritet, såvel som interaktioner mellem begge faktorer. Disse resultater er relevante i såvel kliniske sammenhænge til f.eks. vurdering af hørenervens helbred, såvel som til udviklingen af modeller for den elektrisk aktiverede hørenerve.

Résumé

Plus d'un demi-million de personnes ayant une perte sévère ou profonde de l'audition ont un implant cochléaire (IC). Cet implant est une prothèse neuronale implantée chirurgicalement qui stimule électriquement le nerf auditif et permet une restauration partielle de l'audition. Une majeure partie des utilisateurs d'ICs comprennent la parole quand ils se trouvent dans un environnement silencieux. Cependant, même les utilsateurs tirant le plus d'avantages de l'IC ont des difficultés de compréhension en présence de bruit ou de plusieurs locuteurs.

Notre connaissance de l'intégration de courants électriques par le nerf auditif reste limitée. Cela est dû en partie par l'utilisation clinique d'impulsions biphasiques à charge balancée. Ces impulsions sont sans danger d'un point de vue electro-chimique, mais elles compliquent l'interprétation des résultats, puisque les deux phases ont une polarité opposée.

Ce projet a pour objectif de mieux comprendre l'intégration temporelle et polarité-spécifique de courants éléctriques par le nerf auditif. Cela a été accompli en mesurant la sonie, les seuils d'audition et la localization de paires d'impulsions pour des utilisateurs d'ICs, tout en variant l'intervalle entre les impulsions et la polarité de chaque impulsion.

Les résultats montrent des effets variés de l'intervalle et de la polarité, et des interactions entre les deux facteurs. Ces résultats ne sont pas seulement importants pour des applications cliniques, telle que l'estimation de la santé du nerf auditif, mais aussi pour le développement des modèles du nerf auditif stimulé électriquement.

Acknowledgments

Back in September 2010, it was with quite some luck that I learned about the programme in “Engineering Acoustics” at DTU. Yet, already after a little research (“Kongens Lyngby? Aaah, Copenhague!”) and reading reports from two previous students (many thanks, Sylvain and Rémi), my mind was already cycling with the wind in Denmark (I quickly realized though, that one had to cycle *against* the wind most of the time).

I like to compare these last years to a (rather long) bike tour: always going forward, with ups and downs, and many different roads leading to the final destination. Such a tour is only possible (and fun!) with the help of many, and I am immensely grateful for the support I received.

Thank you, Bastian and Jeremy, for your dedicated guidance and support during the project. I attempted many small breakaways on unknown roads. Yet, I never heard instructions to make a U-turn, and you always took the time to go to the whiteboard and make sure that there was a sensible plan.

Thank you, Bob, for hosting me during my very first “days”[1] in Cambridge, and for the in-depth support that followed. I feel extremely lucky and grateful for such a fruitful collaboration, which sparked many ideas. I am looking forward to be a part of the bleepers again.

Thank you, Torsten, for the financial support through the Oticon Centre of Ex-

[1] or rather, weeks..

cellence for Hearing and Speech Sciences. And thank you for supporting my early interest in cochlear implant (CI) research. I am continuously impressed by the energy and motivation you put in the Hearing Systems group.

Thank you Sébastien, Iris and Sepp for an exciting first step in the CI world during my master thesis, and during the writing that followed. Thank you Sébastien also for the help in writing the PhD proposal.

Thank you, Suyash, for the very inspiring discussions during the beginning of the PhD, that picked up my curiosity about polarity effects. And thank you so much Alexandre, for forcing me to use version control!

I would like to give my warmest acknowledgements to all the colleagues at the House of Acoustics at DTU, particularly in rooms 111 and 129 of building 352. There were many small conversations, helpful comments, coffees, lunches and parties, which will all be missed. Special thanks to Charlotte, Andreu and Axel for supporting me during the last (rather stressful) months; Gerard for the help in the ASSR study; Gerd for the Danish conversations; and Niclas for the feedback on the PPS toolbox.

I would like to give equally warm acknowledgments to my colleagues at the Medical Research Council in Cambridge (Alan, Alex, Hedwig, John, Stefano, Tobias and Wiebke). Again, many small conversations (and warm beers) that are essential for a PhD. Through the group in Cambridge, I also got to know Olivier, Quentin and Gaston from the CNRS in Marseille, who I would like to thank for the many fruitful discussions.

I am immensely grateful to the many listeners I met while gathering the results presented in this thesis. I must admit that their patience, motivation and life stories helped me a lot in staying motivated.

Thank you, Patricia and Keith, for hosting me in your villa while I was without

an apartment, and for the many dinners and discussions that followed. You are both truly inspiring.

There are many, many, amazing friends I would like to thank, which I met through studies, trips, music, badminton, football, squash or sharing a flat. That includes the “symfuni/plenum” family, the “Miguel” family, the “sweaty 2v2” family, the “Thoday street” family, the “Icelandic” family, the “Corto Maltese” family, the “Saint-Felix” family, the “House Ear Institute” family, and many others. You all made me feel at home wherever I was, and that was essential. A special thanks to Helga for the Danish/French language exchange, Anna-Julia for the cover illustration of this thesis, and the “Bram Kortekaas octet” for cheering me up in the last months.

Finally, all this would not have been possible without the continuous, loving support of my family, back in France (and Germany!). And lastly, thank you so much, Akshyeta, for caring for me during those last, rather intense months.

François Guérit, 04 April 2018

Related publications

Journal papers

- Guérit, F., Santurette, S., Chalupper, J. and Dau, T. (**2014**). "Investigating interaural frequency-place mismatches via bimodal vowel integration". *Trends in Hearing* **18**.

- Guérit, F., Marozeau, J. and Epp, B. (**2017**). "Linear combination of auditory steady-state responses evoked by co-modulated tones". *Journal of the Acoustical Society of America* **142**, EL395-400.

- Guérit, F., Marozeau, J., Deeks, J.M., Epp, B. and Carlyon, R.P. (**2018**). "Effects of stimulation order of opposite-polarity pulses on loudness for cochlear implant listeners". *Journal of the Acoustical Society of America* **144(5)**, 2751-2763.

- Guérit, F., Marozeau, J., Epp, B. and Carlyon, R.P. (**in preparation**). "Polarity-specific charge summation in cochlear implant listeners".

- Guérit, F., Marozeau, J., Joshi, S.N. and Epp, B. (**in preparation**). "Effects of polarity on the perception of interaural time differences in cochlear implant listeners".

Contents

1

General introduction

The inner ear is a very sensitive and fragile organ where mechanical vibrations from incoming sounds are transformed into neuronal signals (e.g., Plack, 2013, for a review). The inner hair cell is at the core of this transduction, and its malfunction or damage can create irreversible and profound hearing loss. In such cases, a successful treatment is the cochlear implant (CI), a surgically implanted device that stimulates electrically the spiral ganglion neurons (SGNs) innervating the hair cells. CI users typically understand speech well in quiet conditions (with a large variation across listeners, Blamey et al., 2013). However, even the most successful users struggle in noisy situations.

Figure 1.1A shows a schematic CI electrode array, near the spiral ganglion neurons (SGNs) that constitute the auditory nerve. For normal hearing (NH) listeners, the mechanics of the cochlea are such that high-frequency sounds induce a maximum vibration of the basilar membrane (Figure 1.1B) close to the base of the cochlea. Low-frequency sounds induce a maximum vibration towards the apex of the cochlea. Those vibrations are then captured by the hair cells, which are innervated by the SGNs. Accordingly, the CIs mimic this frequency-to-place distribution by having several electrodes distributed along the electrode array (Figure 1.1A).

The CIs create a large spread of current, because of the placement of the electrode array in a conductive fluid (blue, Figure 1.1A). A large population of SGNs will thus be activated by the current field from each electrode. This impairs the perception of complex sounds -such as speech in noise or music- by CI listeners (Friesen et al., 2001; Oxenham and Kreft, 2014; Won et al., 2007). In order to limit the current field interactions between neighbouring channels, CIs typically make use of short (25-100 μs), symmetric, rectangular and biphasic pulses. These allow for quickly interleaving the stimulation at different electrodes, without a direct

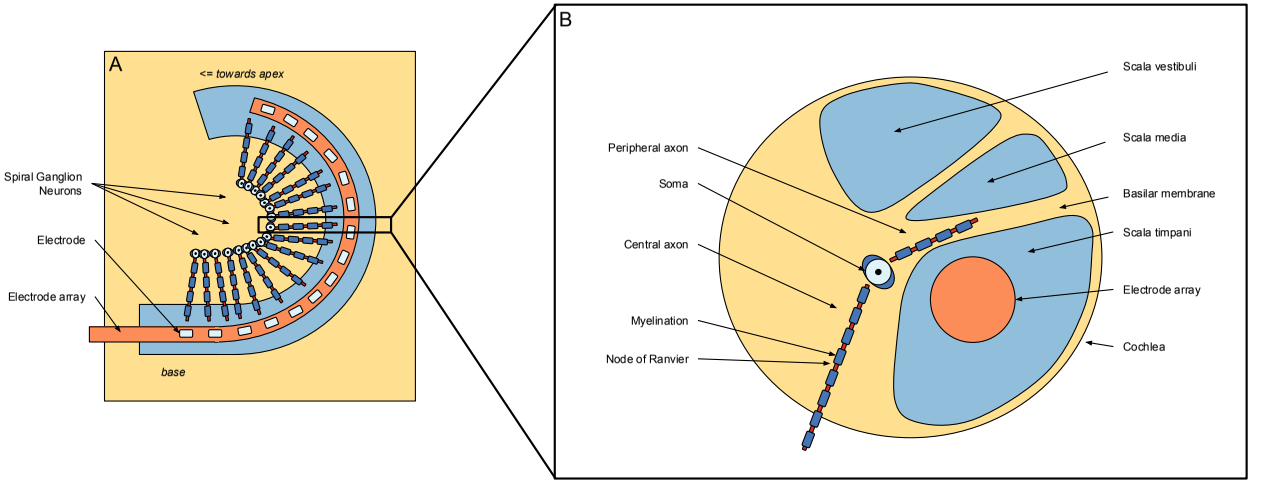

Figure 1.1: **A.** Schematics of a CI electrode array inserted inside the cochlea. Different electrodes (also called contacts) intend to target different populations of spiral ganglion neurons. **B.** Radial view. The electrode array is usually inserted inside the scala timpani, and electrical currents can theoretically depolarize the neurons at various nodes of Ranvier (in red) along its axis.

vector summation of the electrical currents (Wilson et al., 1991). These pulses are also charge-balanced, and avoid any DC currents that create oxidation products (Brummer and Turner, 1977; Lilly et al., 1955).

Not only does the spread of current imply that many neurons are activated at once, but it also implies that several nodes of Ranvier (red, Figure 1.1B) within one neuron might create an action potential (AP). Insights in such behaviour can be gained from looking at an equivalent electrical model of the SGN, such as shown in Figure 1.2 (simplified version inspired from Imennov and Rubinstein, 2009; Smit et al., 2009). Each node of Ranvier is approximated as a leaky integrator, in parallel with voltage-gated ion channels. When presenting electrical stimuli with a CI, the non-uniform distribution of voltage along the SGN will create a current flow in between the nodes. At some nodes of Ranvier, this current flow will eventually trigger an AP. Furthermore, modelling suggests that nodes of Ranvier are more likely to trigger an AP if the voltage distribution along the SGN exhibits a locally positive second derivative.

Before triggering an AP, the current flow at each node will be integrated by the membrane, with a time constant depending on the passive properties, C and Rleak (Lapicque, 1907). The diameter and myelination of the SGNs can vary significantly along the neuron (e.g., Liberman, 1984). This will affect the passive properties of the neural membrane and consequently its time constant of charge integration (Bostock et al., 1983; Colombo and Parkins, 1987; Resnick et al., 2018; Smit et al., 2008). There are also differences across nodes in terms of the quantity and variety of ion channels. They imply the presence of various refractory and facilitation time constants along the neuron (Boulet et al., 2016). Finally, the soma (located between the central and peripheral axon) has a relatively high capacitance. This can delay the travel time of an AP created at the peripheral nodes by a few tens to hundreds of microseconds (Javel and Shepherd, 2000; Miller et al., 1999b; Rattay et al., 2001a). All these factors taken together might affect perception, particularly at inter-pulse intervals (IPIs) below a few hundred microseconds.

The importance of the multiplicity of nodes of Ranvier (and their different properties) has gained attention since it was shown that both anodic and cathodic currents could elicit a neural response, and possibly at different locations along the

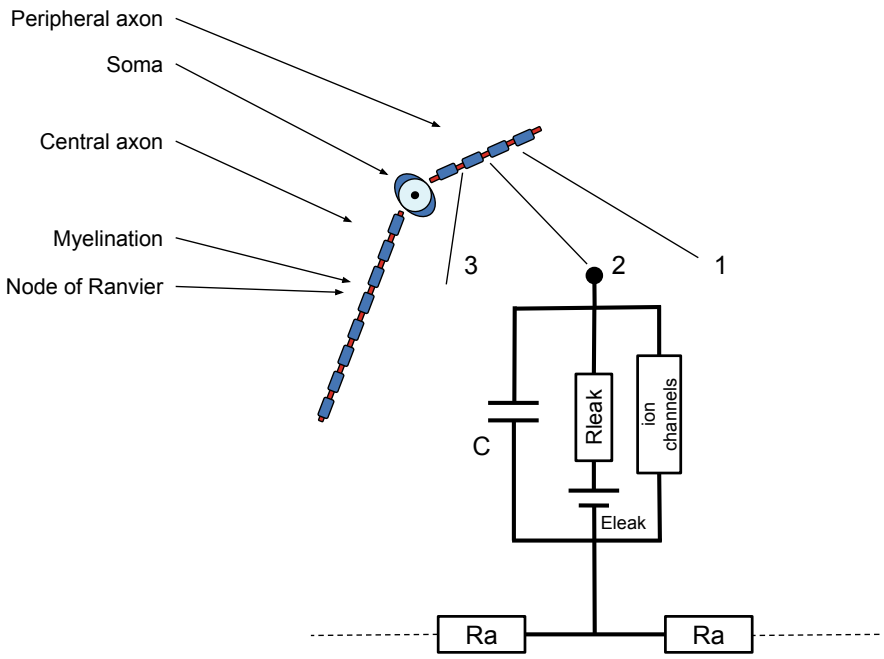

Figure 1.2: Simplified electrical model of one SGN, with C and Rleak the passive capacitive and resistive properties at the node of Ranvier, Eleak the voltage source ensuring a negative resting potential of the transmembrane potential, Ra is the axoplasmic resistance. Voltage-gated ion channels govern the creation of APs, and have been put together in one component for simplification. If nodes 1 and 3 were to have on average an external potential higher than node 2, current would flow outwards node 2 (depolarizing it). Formally, this means that a positive second derivative of the voltage along the neurons is likely to trigger an action potential.

SGNs (Miller et al., 1999b; Rattay et al., 2001a). This is because, as shown in Figure 1.3, both anodic and cathodic stimulation create areas with a positive second derivative of the voltage along the neuron. Cathodic currents might, furthermore, trigger APs at more peripheral nodes (i.e. closer to the electrode) than anodic currents (cf. areas of positive second derivative in Figure 1.3). Single-neuron recordings in cats show that both polarities can trigger APs, with longer latencies and lower thresholds for cathodic stimulation than with anodic stimulation (Miller et al., 1999b). This is consistent with the idea that cathodic currents trigger APs at

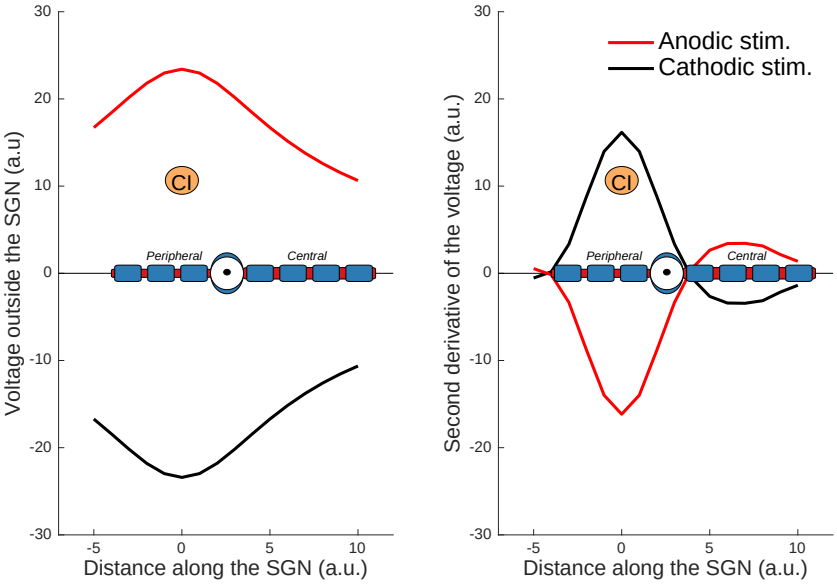

Figure 1.3: **Left.** Simulated voltage distribution along the neuron axis, for an extracellular anodic and cathodic stimulation (proportional to 1/r, with r the distance from the electrode). **Right.** Second derivative of the extracellular voltage along the neuron. This shows an inflection point, with the second derivative being negative close to the electrode for anodic stimulation, and positive for cathodic stimulation.

more peripheral nodes (i.e. closer to the electrode) than anodic currents (cf. areas of positive second derivative in Figure 1.3).

Since monophasic cathodic or anodic stimulation is unsafe (Brummer and Turner, 1977), these pulses cannot be used in human listeners. Polarity effects can, however, be investigated in humans using pseudo-monophasic (Figure 1.4) or quadraphasic pulses. The applicability of these pulses in human CI listeners has been supported by several studies within the last decade (Carlyon et al., 2013; Macherey et al., 2006, 2008, 2010, 2017; Undurraga et al., 2013; Wieringen et al., 2005). For example, using pseudo-monophasic pulses, the auditory brainstem response (ABR) exhibits longer latencies for cathodic than anodic stimulation,

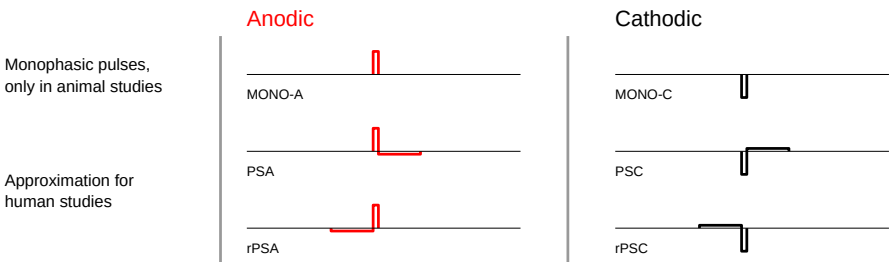

Figure 1.4: Monophasic pulses are unsafe, as they are charge-unbalanced and create oxidation products. In order to study polarity effects in human CI users, one can however use pseudo-monophasic pulses (or also triphasic and quadraphasic pulses, not shown here). Pseudo-monophasic pulses are charge-balanced within a short time window. Because the membrane at the nodes of Ranvier behaves as a leaky integrator, the short-high phase will be more efficient than the long-low phase in eliciting a neural response.

particularly at low levels, consistent with results in animal studies (Undurraga et al., 2013).

If each polarity activates different regions on the SGNs, the temporal properties of the neural response might differ, because of the differences in morphology mentioned above. The reverse approach is equally applicable: knowing the response properties for each polarity might reveal underlying pathologies, such as the shrinkage or demyelination of the SGNs. Furthermore, a recent approach showed that one may account for various effects of pulses shape and pulse rate with biphasic pulses (Joshi et al., 2017), when including polarity-specific response properties in a computational model of electrical stimulation.

This thesis aimed at better characterizing the polarity-specific response of the electrically stimulated auditory nerve in human CI listeners. The integration of pulse pairs at short inter-pulse intervals (IPIs) was studied (Figure 1.5), both monaurally (Chapters 2 and 4) and bilaterally (Chapter 3). In all chapters, the IPI and the polarity of each pulse were varied. It was expected that the effects of IPI would interact with the polarity of each pulse, since the polarity might change the locus of excitation along the SGNs.

In Chapter 2 (Figure 1.5A-B), the charge integration over short IPIs (below 344 µs) was studied with pulse pairs of the same polarity. The goal was to improve our

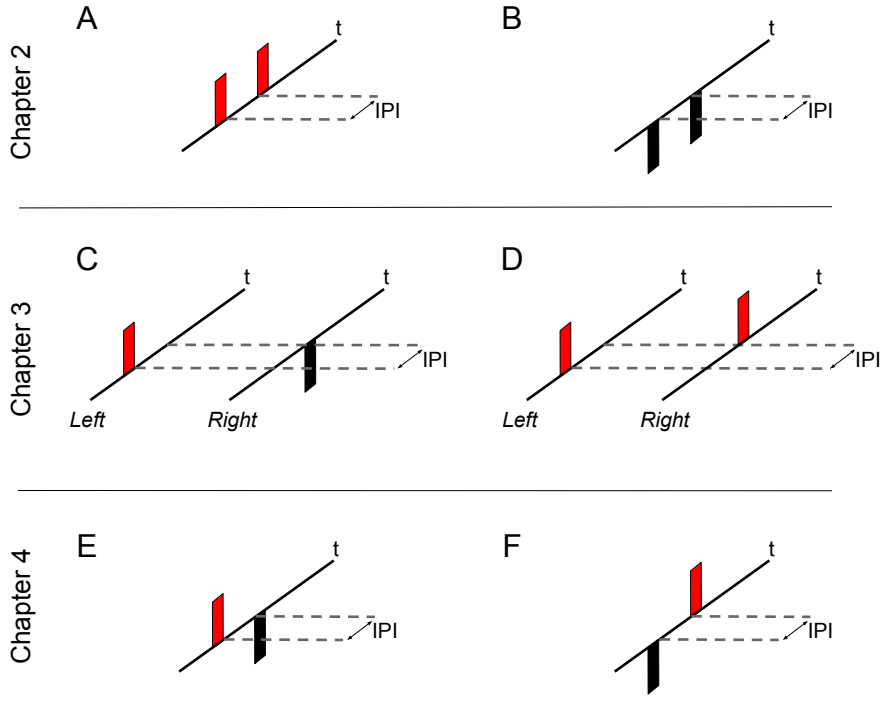

Figure 1.5: Schematics of the paradigms used in the different chapters. The counteracting, charge-balancing phases are not shown here for improved readability. **A-B.** Paradigm used in Chapter 2, investigating monaural interactions at short inter-pulse intervals when each pulse has anodic (A) or cathodic polarity (B). **C-D.** Chapter 3 investigated the perception of interaural time differences (equivalent to IPI in this case) when the polarity of the stimulus in the right ear changed from anodic (C) to cathodic (D). **E-F.** Chapter 4 investigated the order effects of stimulation for opposite-polarity paired pulses. Panel E shows anodic-first stimulus, panel F cathodic-first stimulus.

understanding of the connection between polarity-specific charge integration and the sensitivity or loudness perception in human CI listeners. The results and the proposed paradigm are relevant for improving the clinical fitting and characterizing underlying pathologies.

Because the suggested difference in delay between the latency of the neural response for anodic and cathodic stimulation (200 µs, Miller et al., 1999b) is in the range of perceivable ITDs, Chapter 3 investigated whether changing the polarity

of the stimulus in one ear was enough to affect ITD-based localization (Figure 1.3C-D).

Chapter 4 (Figure 1.5E-F) investigated the effects of stimulation order of opposite-polarity pulses on loudness. Order effects might stem from differences in charge integration (Chapter 2) as well as differences in timing and locus of excitation associated with each polarity (Chapter 3). Furthermore, this chapter investigated whether the order effects changed when each pulse had the same loudness or the same level.

Finally, Chapter 5 summarizes the main findings of the thesis, highlighting some of the limitations encountered with the necessary use of asymmetric pulses. This chapter also suggests further applications based on the results and experimental paradigms of this thesis.

Appendices: Towards clinic-friendly methods

The core chapters of this thesis investigate basic mechanisms of charge integration within and across ears. This requires each subject to come for several sessions (8-10 hours per experiment). While achievable for research purposes, this amount of time is not available in clinics. The two appendices of this thesis therefore investigated methods that either limit the participation of subjects to an every-day like task (Appendix A), or do not require the active participation of the subjects at all (Appendix B).

For patients having residual hearing in one ear and a cochlear implant (CI) in the opposite ear, interaural place-pitch mismatches might be partly responsible for the large variability in individual benefit. Although behavioural pitch matching between the two ears could help to individualize the fitting, it is rather tedious when using methods that do not suffer behavioural bias (Cosentino et al., 2017). In Appendix A, an alternative method using two-formant vowels was developed and tested. The results suggest a possible use of such vowel spaces to derive interaural frequency-place mismatches. However, the method remains limited by difficulties in bimodal fusion of the two formants.

Clinics commonly measure electroencephalographic (EEG) potentials in response to sound, as this typically does not require the active participation of the listeners. Appendix B investigates the linearity of one type of EEG measure, the auditory steady-state response (ASSR). As this showed a linear response in normal hearing listeners, this method is promising in tracking the numerous nonlinear interactions at the level of the auditory nerve in CI listeners.

2

Polarity-specific charge summation in cochlear implant listeners[a]

Abstract

Knowledge about the capacitive-resistive properties of spiral ganglion neurons is relevant for the clinical fitting of cochlear implants, as it may reveal underlying pathologies. However, the clinical use of biphasic pulses limits such measures, because currents of opposite polarities interact at the level of the neural membrane. Here, we propose a paradigm to study polarity-specific summation of currents. We used pairs of pseudo-monophasic pulses with inter-pulse intervals ranging from 0 to 345 μs. We assumed that most of the excitation would stem from the short-high phases, which had the same polarity. The inter-pulse interval had a significant effect on the perceived loudness, and this effect was consistent with an underlying leaky integrator. Furthermore, the effect of interval interacted with the polarity of the pulse pairs. At thresholds, there was only an effect of polarity, but no effect of interval nor an interaction between both factors. We discuss possible peripheral origins of these results.

[a] This chapter is based on Guérit F., Marozeau J., Epp B. and Carlyon R.P. (*in preparation*)

2.1 Introduction

Both the myelination and the diameter of the spiral ganglion neurons (SGNs) can decrease following sensorineural hearing loss (Leake and Hradek, 1988; Nadol, 1997). These morphological changes affect how the SGNs integrate the electrical charge delivered by cochlear implants (CIs, Bostock et al., 1983; Colombo and Parkins, 1987; Koles and Rasminsky, 1972; Resnick et al., 2018; Smit et al., 2008). Measures that relate to the capacitive-resistive properties of charge integration could, therefore, reveal underlying pathologies (Miller et al., 1995; Prado-Guitierrez et al., 2006).

In order to characterize such properties, animal single-neuron recordings usually report the current level required to elicit a spike as a function of the duration of a rectangular monophasic pulse (e.g., Honert and Stypulkowski, 1984; Miller et al., 1999b). Such a curve should exhibit a slope of -6 dB per doubling of pulse duration for a perfect integrator, with lower absolute values for a leaky integrator. Alternatively, one can present pairs of monophasic pulses at an equal level, with different inter-pulse intervals (IPIs), and compare the response to that of a single pulse (Cartee et al., 2000). At 0-μs IPI, this is equivalent to doubling the phase duration of a single pulse. When increasing the IPI, the membrane will slowly return to its resting potential between the two pulses, at a rate inversely proportional to the time constant of the underlying leaky integrator. Cartee et al. (2000) investigated this in the SGNs of cats and reported time constants between 100 and 400 μs.

The clinical use of symmetric biphasic pulses complicates such measures in human CI listeners. This is because the two phases are necessarily of opposite polarity and so partially cancel each other at the cell membrane (e.g., Honert and Mortimer, 1979). Increasing the phase duration of biphasic pulses will reduce charge cancellation, because of increasing the time between the start of the first phase and the end of the second. As a result, the slope of reduction in current with increasing phase duration will be a complex combination of adding charge in each phase and reducing charge cancellation between the phases (Ramekers et al., 2014). Furthermore, both anodic and cathodic phases of biphasic pulses can be excitatory, probably by eliciting spikes at different portions of the SGN, as shown

in animals (Miller et al., 1999b) and humans (Macherey et al., 2008; Undurraga et al., 2013). Complex order effects can stem from this, depending on the relative ratio of current and neural excitation from each phase (Guérit et al., *submitted*).

Here we propose and test a paired-pulse paradigm approximating that used with monophasic pulses in animals (Cartee et al., 2000; Miller et al., 1999b). The new paradigm uses asymmetric, pseudo-monophasic, charge-balanced pulses (Figure 2.1), consisting of one long-low and one short-high phase. The underlying assumption is that the short-high phase will be more efficient in eliciting a response than the long-low one because the neural membrane behaves as a leaky integrator (Lapicque, 1907; Miller et al., 2001a; Undurraga et al., 2013). Furthermore, inserting a gap of 2 milliseconds between the short-high and long-low phases should avoid both phases cancelling each other at the level of the membrane.

As in Cartee et al. (2000), we expected that short IPIs between two short-high phases of the same polarity would lead to "summation," i.e., integration of the charge at the level of the membrane. We, therefore, investigated whether short IPIs reduced detection thresholds and loudness, and compared the time course of this effect in conditions where the two short-high phases were anodic or cathodic. We included a 0-μs IPI condition, allowing us to study the effects of doubling the phase duration without the influence of a temporally adjacent equal-amplitude phase of opposite polarity.

The peripheral and central axons of the SGNs might differ in their amount of myelination, and diameter. On each axon, the distance between the nodes of Ranvier, as well as the distance between the nodes and the highly capacitive soma (Liberman and Oliver, 1984) might also vary. All these factors are likely to affect the time constants of charge integration. For example, modelling studies suggest that peripheral axons should exhibit longer time constants of charge integration (Cartee, 2000; Cartee, 2006; Joshi et al., 2017). Since one might be able to target preferentially the peripheral and central axons with cathodic and anodic stimulation, respectively (Miller et al., 1999b; Rattay et al., 2001a), we expected our results to show longer time constants with cathodic currents, when compared to anodic currents.

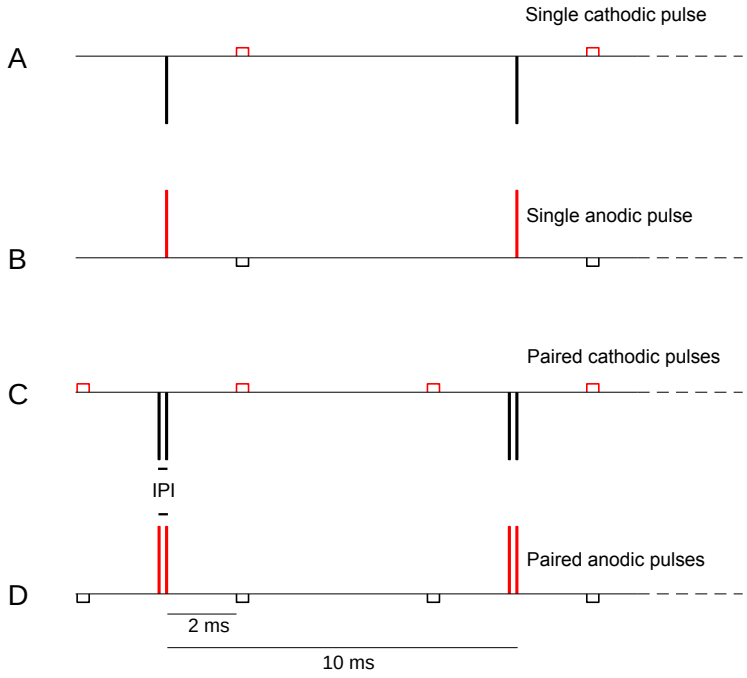

Figure 2.1: Representation of the different stimuli used throughout the study. A-B: “Single” pseudo-monophasic pulses (asymmetry ratio of 8), with the short-high phase being cathodic (A) or cathodic (B). The interphase gap was 2-ms long, and the pulses were repeated at a rate of 100 Hz for 400 ms. C-D: “Paired pulses.” For the paired pulses, the inter-pulse interval (IPI) had values ranging from 0 to 344 µs.

2.2 Methods

2.2.1 Listeners

Six listeners took part, all of whom were recipients of an Advanced Bionics CI (cf. Table 1 for demographics). Listeners were recruited both in Cambridge (UK) and Copenhagen (DK) and the experimental procedure was approved respectively by the National Research Ethics Committee for the East of England (ref. number 00/327) and the Danish Science-Ethics Committee (ref. number H-16036391). All listeners signed a participation agreement before data collection.

2.2.2 Setup and stimuli

We conducted all experiments by means of direct stimulation, i.e. using research hardware (CPI-II clinical interface, PSP speech processor) and software (BEDCS 1.18, PPS toolbox, Matlab 2014a) instead of the clinical speech processor of the listeners.

Stimuli consisted of trains of pseudo-monophasic pulses repeated at 100 pps for 400 ms (Figure 2.1). Each pseudo-monophasic pulse consisted of a short-high and a long-low phase, separated by a gap of 2 ms. The duration of the short-high phase was 43 μs and that of the long-low phase was eight times longer, and with the amplitude reduced by the same factor. In one condition, the short-high phase was cathodic ("Single cathodic", Figure 2.1A) and in another, anodic ("Single anodic", Figure 2.1B).

We also created trains of paired pulses where for the first pulse, the long-low phase preceded the short-high, while it was the opposite for the second pulse (Figure 2.1C and 1D). That way, the two short-high phases (which we assumed would create most of the neural response) were temporally adjacent. Paired-pulse stimuli had inter-pulse intervals (IPI) ranging from 0 to 345 μs[a]. In a similar manner as for the single pulse stimuli, we created a cathodic (Figure 2.1C) and anodic version (Figure 2.1D), with the short-high phases being cathodic and anodic, respectively.

[a] At 0-μs IPI, the design was such that there was no glitch in the amplitude between the two pulses.

Table 2.1: Demographics of the CI listeners. All listeners were implanted with an Advanced Bionics HiRes90k device. 1j is a straight array, ms is curved ("mid-scala").

Subject ID	Age (y)	Duration of implant use (y)	CI Side	Electrode used for testing	Type of electrode array	Etiology
S1 (680)	57	3	Right	9	ms	Menieres
S2 (808)	60	9	Left	9	1j	Pendred Syndrome
S3 (1048)	28	4	Left	8	ms	Unknown, pre-lingual
S4 (AB13)	76	9	Right	9	1j	Unknown
S5 (AB2)	57	9	Left	9	1j	Ototoxicity
S6 (1126)	66	3	Right	9	ms	Rhesus disease

Prior to and throughout the experiments, we checked the stimuli with a test implant (HiRes90k) and a digital storage oscilloscope. Asymmetric, pseudomonophasic pulses are charge-balanced, but only within the limits of compliance of the device (7-8 V, Mesnildrey, 2017): above those, the short-high phase would not reach its assigned amplitude, and charge balancing would rely on the blocking capacitors of the device. We therefore measured impedances at the beginning of each session. Across all listeners and sessions, the maximum voltage we reached was 4.9 V. We also measured impedances at the end of each session, and did not see any significant changes.

2.2.3 Detection thresholds

Detection thresholds for all stimuli were measured with a one-up-three-down two-alternative forced choice procedure (Levitt, 1971). We used eight reversals, two with a step size of 1 dB, followed by six with a step size of 0.25 dB. Because of the minimum step size of the device, not all values were achievable on a logarithmic scale[b]. The procedure hence tracked the desired level of the short-high phases, but we computed the final thresholds from the actual levels of the last six reversals. Each measurement was repeated twice, leading to 24 measurements (for each polarity: single pulse, paired pulses with 0-, 43-, 86-, 172- and 345-us IPGs). We ensured that the starting point of every trial was clearly audible.

2.2.4 Loudness balancing at most comfortable levels (MCLs)

For all stimuli (single and paired pulses at all gaps, for both polarities), we obtained the MCLs using an 11-point loudness scaling chart (number 6 corresponded to the MCL). We then picked a level slightly below the MCL of the single cathodic pulse as a reference for the subsequent loudness balancing. We did not pick the MCL

[b] The HiRes90k device dynamic range is divided in a linear way, from 0 to 2040 μA. The minimum achievable step size depends on the dynamic range used (1 μA between 0-255 μA, 2 μA for 0-510, 4 μA for 0-1020, 8 μA for 0-2040). Asymmetric pulses further limit the minimum step size achievable. With a ratio of 8, this doubles the values (2 μA for 0-255 μA, etc.)

itself as a reference in order to give enough headroom for the loudness balancing procedure without reaching any uncomfortable loudness.

For each loudness balancing, the subject heard two sounds and reported which of the two was the loudest. The experimenter adjusted the level of the second sound until both had the same loudness. The experimenter bracketed several times around this point of subjective equality before going to the next measurement.

We first matched the level of the single anodic pulse to the reference, single cathodic pulse. We ran the loudness procedure four times, twice with the anodic pulse as a reference, twice with the cathodic. The final balanced value was computed from the average of the four adjustments (in dB).

Using only one polarity, the listeners balanced the loudness of the paired-pulse with 0-µs IPI to that of the single pulse; then, the paired-pulse with 43-µs IPI to the paired-pulse with 0-µs IPI; 86- µs IPI to 43- µs IPI and finally 172- µs IPI to 86- µs IPI. The final value was computed from the average of two loudness balancing trials, swapping the reference and adjusted stimuli each time. The paired stimuli with 345-µs IPI were not included to ensure that the loudness balancing for all anodic or cathodic stimuli would fit within one testing session.

2.3 Results

2.3.1 Single pulses

Figure 2.2 shows the individual results of the detection thresholds and loudness balancing measurements for the single pulses. At MCL, the anodic stimuli required less current to achieve the same loudness as cathodic stimuli (+2.50 dB, $t(5) = 7.16$, $p < 0.001$). At threshold however, less current was required for cathodic stimuli than for anodic stimuli (-1.15 dB, $t(5) = 3.41$, $p = 0.019$). In both cases the effect was in the same direction for all listeners.

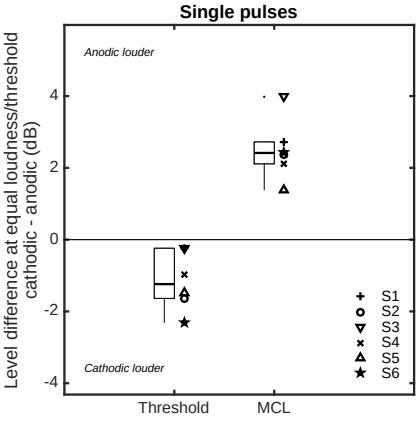

Figure 2.2: Polarity difference in detection thresholds and loudness-balanced MCLs for the single pulses. Lower and Upper limits of the boxes: 25th and 75th percentiles. Horizontal black line (and blue line): median level. Whiskers: 25th (or 75th) percentile minus (or plus) 1.5 the interquartile range. Dots correspond to data points with values outside the range delimited by the whiskers.

2.3.2 Paired pulses

Figure 2.3 shows the individual detection thresholds with paired pulses. Levels are normalized to the level of the threshold for the single pulse. As seen from the individual lines, there was a rather large within- and across-subject variability. A repeated measures ANOVA showed an effect of polarity ($F(1, 5) = 11.5$, $p = 0.0195$), but no effect of IPI ($F(4, 20) = 0.77$, $p = 0.56$) nor an interaction between polarity and IPI ($F(4, 20) = 1.1$, $p = 0.40$). When pooled across IPIs, the anodic pulses (red line and symbols) required on average 1.80 dB less current than single pulses to reach threshold. For cathodic pulses (black line and symbols), this reduction was significantly larger ($t(5) = 3.39$, $p = 0.0194$) and amounted 3.76 dB.

Figure 2.4 shows the individual results of the loudness balancing for the paired pulses. Unlike the case for thresholds, all subjects show a consistent and monotonic increase in MCL with increasing IPI. A repeated-measures ANOVA on the normalized levels showed significant effects of polarity ($F(1, 5) = 12.7$, $p = 0.0162$), IPI ($F(3, 15) = 326$, $p < 0.001$) and a significant interaction between polarity and

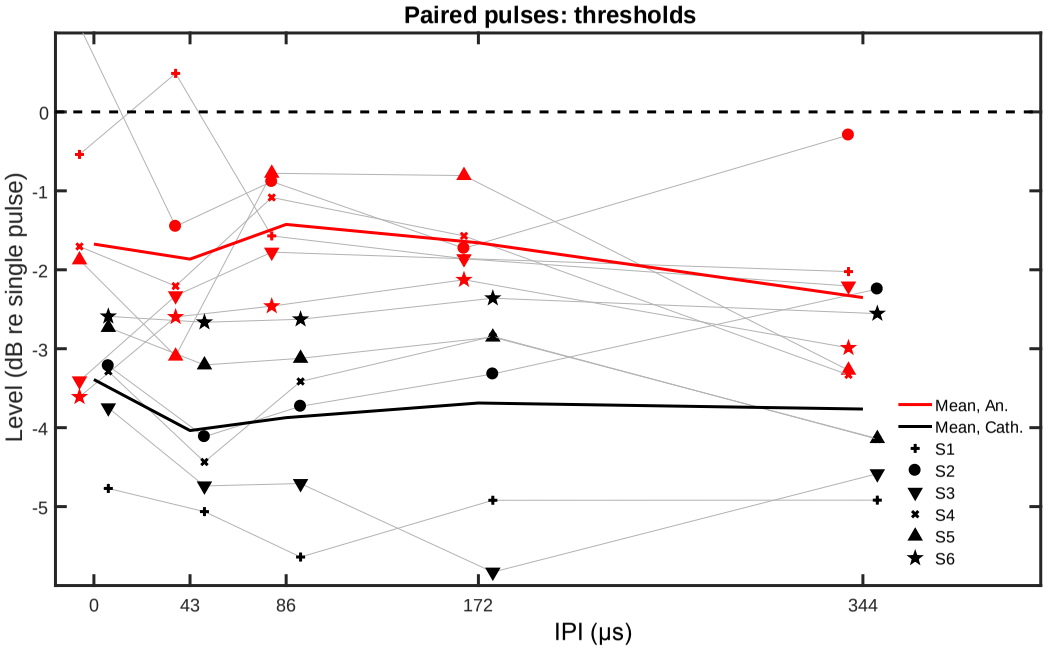

Figure 2.3: Detection thresholds for the anodic (red) and cathodic (black) paired pulses. Values are normalized to the threshold of the single pulse with corresponding polarity (dashed line). Thick lines indicate the mean value across subjects. For visibility, the individual anodic and cathodic results are slightly shifted from their actual IPI value.

IPI ($F(3, 15) = 26.7$, $p < 0.001$).

Post-hoc t-tests with Bonferroni corrections showed that the difference between each polarity was only significantly different at IPIs of 86 and 172 µs (0 µs: $t(5) = 1.3$, $p = 0.25$; 43 µs: $t(5) = 1.97$, $p = 0.106$; 86 µs: $t(5) = 4.22$, $p = 0.0083$; 172 µs: $t(5) = 5.20$, $p = 0.0035$). At 0 µs (equivalent to doubling the phase duration of the single pulse), an average level reduction of 3.96 dB was needed to achieve the same loudness as the single pulses (both polarities pooled together, 4.04 and 3.86 respectively for anodic and cathodic pulses).

2.4 Discussion

With the paired-pulses paradigm presented here, there was a clear effect of the IPI on the loudness of the paired pulses. Paired pulses required less current to elicit the same loudness than single pulses, and the effect was more pronounced at the shortest IPIs. Furthermore, this effect depended on the polarity of the stimulus. For anodic stimulation, the level of the paired pulses was only 0.9 dB lower than that of the single pulse at the largest IPI tested (172 µs). For cathodic stimulation, this was significantly more and amounted to 2.2 dB. No clear effect of IPI occurred at threshold. In the following, we discuss mechanisms that could explain those results, including interactions at the level of the neural membrane up to central loudness integration.

2.4.1 Discrepancy between the results for paired pulses at threshold and MCL

With pseudo-monophasic pulses, the main assumption is that the short-high phase creates most of the neural response. eABR recordings support this assumption in humans, as they show a synchronized response to the short-high phase (Undurraga et al., 2013, with similar parameters to this study). Our results with paired pulses at MCL also support this assumption, as the loudness interacted strongly with the IPI between the two short-high phases. On the other hand, detection thresholds for the paired pulses showed no effect of IPI nor an interaction between IPI and

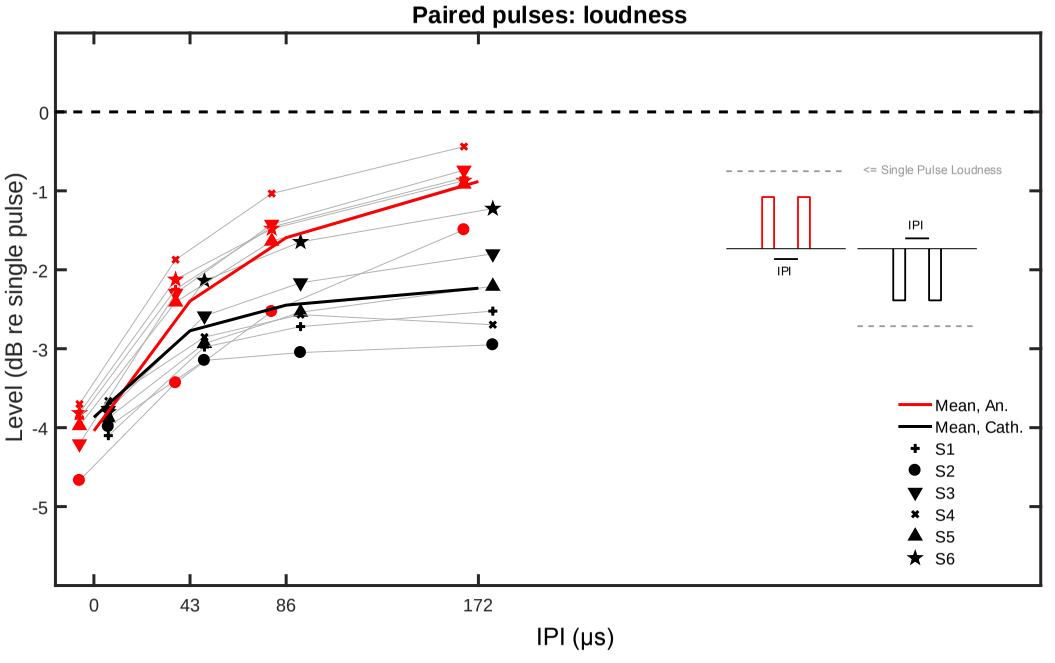

Figure 2.4: Loudness-balanced levels for the anodic (red) and cathodic (black) paired pulses. Values are normalized to the loudness of the single pulse with corresponding polarity (dashed line). Thick lines indicate the mean value across subjects. For visibility, the individual anodic and cathodic results are slightly shifted from their actual IPI value.

polarity. Detection thresholds only exhibited an overall decrease for paired pulses, when compared to singe pulses. This decrease was significantly larger for cathodic (-3.8 dB) than anodic paired pulses (-1.8 dB).

Test-retest variability of the detection thresholds could in theory explain this discrepancy. Indeed, the difference between run 1 and 2 was significantly larger for the detection thresholds than for the loudness balancing (1.12 dB vs 0.27 dB, respectively, $t(5) = 5.91$, $p = 0.002$). This likely reflects a shallower underlying psychometric function at threshold than at MCL. However, although a test-retest variability of 1.12 dB could hide a small effect of IPI at threshold, this would not hide an effect as large as seen at MCL (more than 3 dB for some listeners).

Carlyon et al. (2005) presented listeners with pairs of same-polarity monophasic pulses (in bipolar mode). These alternated in polarity at every pair presentation: each pair of anodic pulses was followed by a pair of cathodic pulses, at a rate of 100 Hz. The alternation of polarity and the use of a bipolar mode do not allow for any polarity-specific interpretation of their results. However, similar to our results at threshold, they showed no effect of varying the IPI on detection thresholds, for intervals ranging from 0 to 4900 μs. That study did not report any MCLs, and it is unknown whether there was a strong effect of IPI at MCL, as shown in our results.

The discrepancy between our results at threshold and MCL might not only stem from the absence of an interaction between the short-high phases, but also from a significant contribution from the long-low phases. If so, for the single-pulse thresholds, both the long-low and short-high phases could have elicited a response, as they were separated by a 2-ms gap. For the paired pulses, the first long-low, the two short-high and the second long-low (Figure 2.1) might have contributed, giving one extra chance to hear a sound (multiple-look advantage, Viemeister and Wakefield, 1991). This would account for both the finding that paired pulses yielded lower thresholds than single pulses, and that the IPI had no consistent effect. Furthermore, the larger decrease in threshold with the cathodic pulse pair (-3.8 dB vs -1.8 dB for the anodic pulse pair), might be explained by a difference in the contribution of the long-low phases across the two conditions.

A contribution of the long-low phases at threshold is consistent with the results of Macherey et al. (2006). In one condition, they decomposed their alternating-

polarity, delayed pseudo-monophasic pulses (ALT-DPS) in a pulse train with the long-low phases or short-high phases alone. The ALT-DPS stimulus had lower detection thresholds than both the trains of short-high and long-low phases. At MCL however, the loudness of the ALT-DPS pulses was similar to that of the short-high phases only. Interestingly, Macherey et al. (2007) proposed a model of the auditory nerve that could account for the effects of decomposing the ALT-DPS stimulus, both at threshold and MCL. They did not provide an explanation as to why the model could account for both results, but the model differed in the determination of thresholds and MCLs. Thresholds were derived as the level for which the stimulus had a 70.7% chance of being correctly detected in a two-interval forced-choice task. The MCLs were however derived as the probability of eliciting 100 or 1000 spikes within their central integration of 20 milliseconds. At threshold, any component (short-high and long-low phases) might contribute to increase the spike probability above that of the noise floor. At MCL however, the number of spikes might be primarily driven by the short-high phases, that are more efficient. If so, loudness perception might be a better proxy than detection thresholds for characterizing the capacitive-resistive properties of the SGNs.

The model from Macherey et al. (2007) does not account for polarity effects, but a recent phenomenological model that could account for polarity effects in the cat has been proposed (Joshi et al., 2017). Figure 2.5 shows the probability of spiking given by this model for monophasic cathodic pulses of 43 and 344 us[c]. With an underlying leaky integrator, more charge is logically required to reach threshold (i.e. 50% probability of spiking) with the longer phase duration. Interestingly, this difference in charge (for a given probability of spiking) is level-dependent, and smaller at the lowest probabilities of spiking. This could explain qualitatively a contribution of the long-low phases in our detection thresholds: the neurons might be at an overall low probability of spiking, hence without much difference between the contribution of the short-high and long-low phases.

[c] Results are similar for anodic stimulation, only with overall higher thresholds. This is because the model is fitted to cat data, that show higher anodic thresholds.

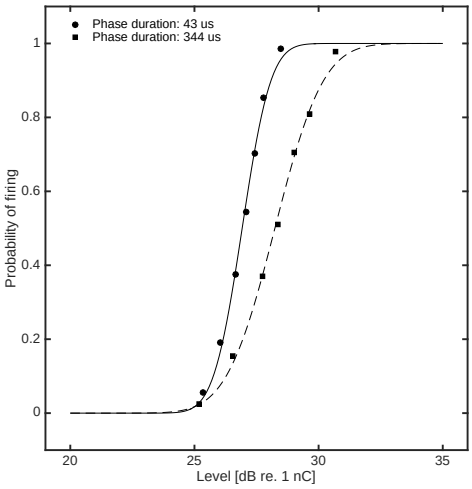

Figure 2.5: Output from the model of Joshi et al. (2017) for monophasic cathodic pulses with durations corresponding to the short-high and long-low phases in our study.

2.4.2 Underlying mechanisms at MCL

Results with paired pulses at MCL suggest a main contribution from the short-high phases (Figure 2.4). For both polarities, paired pulses required less current than a single pulse in order to elicit the same loudness. The difference was largest at 0-µs IPI (equivalent to doubling the phase duration) and decreased with increasing gap. At the longest IPI tested here (172 µs), the loudness of the paired pulses was larger than that of the single pulses, and more so for cathodic than anodic stimulation (2.2 dB vs 0.9 dB, respectively).

Figure 2.6 shows a simple conceptual model that accounts for both the effect of IPI and the non-zero value at the longest IPIs. In panel A of Figure 2.6, the probability of firing of a set of neurons is shown, as a function of the stimulus level [d]. When presenting a single pulse (red vertical line of Figure 2.6), a certain amount

[d] Distributions from Miller et al. (1999a) suggest a normally distributed function of thresholds, on a log-axis

of neurons will spike (coloured in grey in Figure 2.6B). Importantly, some neurons will have a probability of firing in the mid-range (around 0.5) at this level, but might not spike for this first pulse. Thus, even without any facilitation, presenting a second pulse at the same level will give a second chance for these neurons to spike (panel C in Figure 2.6). This explains qualitatively the non-zero value we observed at the largest IPIs, and has been suggested already by McKay and McDermott (1998) in their loudness model. In order to explain the main effect of IPI however, adding a summation term (a.k.a. facilitation, Boulet et al., 2016) at short gaps is required. This is shown in blue in the panel D of Figure 2.6.

The capacitive-resistive properties of the SGNs might be reflected in the summation term of Figure 2.6 (Cartee, 2000; Cartee et al., 2000). However, an estimation of those properties is not as straight-forward as fitting an exponential curve to the results at MCL, nor calculating the reduction in level when doubling the phase duration. This is because the neural contribution given by the second pulse, without summation (Figure 2.6C), will offset the results with pulse pairs at all IPIs. This offset will be directly dependent on the underlying distribution of firing probabilities (Figure 2.6A). That is, this offset might be affected by the number of neurons, their threshold distribution, relative spread (i.e. the slope of curves in Figure2.6A) and jitter. Such properties have been, furthermore, shown to interact with the polarity of stimulation in the cat (Miller et al., 1999b). This could explain the significant effect of polarity, and significant interaction between IPI and polarity in our results. The detailed evaluation of such properties is, however, beyond the scope of this study, and would likely require the use of a geometrical model.

The conceptual model of Figure 2.6 assumes no release of refractoriness with increasing IPI. That is, none one the neurons spiking in response to the first pulse are able to do so for the second pulse. Such assumption is based on reports suggesting that the absolute refractory period of spiral ganglion neurons is around 400 µs (Boulet et al., 2016). Furthermore, modelling with pulse pairs from McKay and McDermott (1998) suggests that it takes two to three milliseconds for a significant proportion of neurons to be active again, due to relative refractoriness. Hence, it seems unlikely that our results are driven by refractoriness at those very short IPIs.

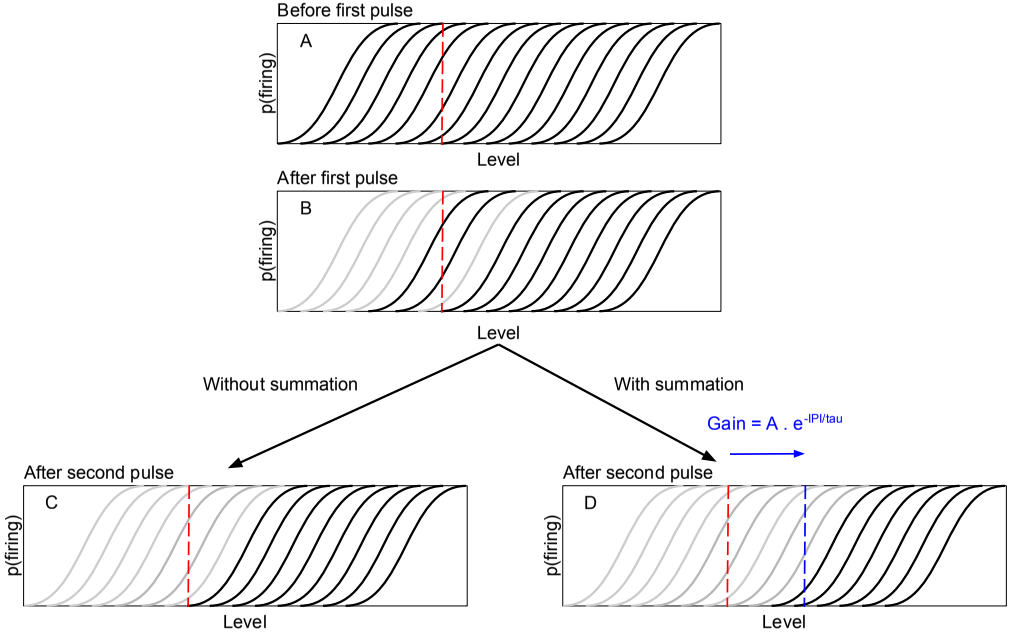

Figure 2.6: Conceptual model for explaining the effects of IPI at MCL. Thick black lines denote the spike probabilities for a set of neurons, as a function of level. If some neurons spike in response to a pulse (at a level indicated by the red vertical line), they go into refractoriness (coloured in grey). In order to explain the effect of IPI, the addition of a summation term at the shortest IPIs is necessary, as indicated by the blue line. This summation term is usually modelled by an exponential decay (Cartee et al., 2000) with time constant τ.

2.4.3 Polarity effects on the single pulses

Changing the polarity of the single pulses had a significant effect on both MCLs and thresholds for single and paired pulses.

At MCLs, the single anodic pulse required on average 2.50 dB less current to achieve the same loudness as the cathodic pulse. This is consistent with most previous studies using pseudo-monophasic pulses, triphasic or quadraphasic pulses (Carlyon et al., 2013; Macherey et al., 2008, 2017; Undurraga et al., 2013; Wieringen et al., 2008). However, in two studies that used respectively 4.7-ms and 4.4-ms inter-phase gaps between the short-high and long-low phases of pseudo-monophasic pulses, (Macherey et al., 2006, 2010), there was no difference in MCL between opposite polarities. Those studies used a longer phase duration than we did (97 μs versus 43 μs), for which the difference in efficiency between the long-low and short-high phases might have been different.

At threshold, polarity had the opposite effect. Cathodic pulses required on average 1.15 dB less current than anodic pulses to reach threshold, and the direction of the effect was the same for all listeners. Other studies have usually reported no effect of polarity on detection thresholds (Macherey et al., 2006; Undurraga et al., 2013). Polarity effects at threshold can however occur on an individual, electrode-to-electrode basis (Mesnildrey, 2017). It might be that with another combination of electrodes and listeners we would see different effects in our results at threshold. It might also be that our results are due to the rather long gap between the short-high and long-low phases (2 ms). Indeed, as discussed in the previous sections, this might allow for both phases to contribute at threshold.

2.4.4 Clinical applications

Knowing the capacitive-resistive properties of the SGNs is relevant in clinics, as it may reveal an underlying pathology. Our paradigm seems promising in measuring such properties (particularly at MCL). One could therefore investigate the results given by such paradigm across listeners and electrodes, and see how it correlates with other proposed predictors of neural health (Bierer, 2010; Pfingst et al., 2015).

Since the results at MCL were monotonic, the duration of the procedure could

be largely reduced by keeping only the conditions with 0- and 172-μs IPI. Furthermore, one could measure the eCAP or eABR in response to the single and paired pulses. This might not only be of interest for patients that cannot perform psychophysical tasks, but this would also allow for comparison with animal models (such as the guinea pig and the cat), where different pathologies can be induced artificially.

2.5 Conclusion

We proposed a paradigm consisting of pairs of pseudo-monophasic pulses where two short-high pulses with the same polarity followed each other. The aim was to study the temporal integration of currents in a polarity-specific manner. For both anodic and cathodic currents, changing the inter-pulse interval had a strong effect on the loudness, consistent with the hypothesis that the short-high phases dominated the neural response. Furthermore, this effect interacted with the polarity of the short-high phases. A simple conceptual model suggests that this might reflect differences in the capacitive-resistive properties of the neural processes depolarized by each polarity. Results at threshold showed no effect of interval, nor an interaction with polarity, which might partly stem from the contribution of the long-low phases.

Acknowledgements

The authors would like to thank the CI listeners for their time and dedication. The authors would like to thank John Deeks and Wiebke Lamping for their help in organizing the testing sessions in Cambridge, Suyash N. Joshi and Olivier Macherey for help in the use and interpretation of their models, and Hedwig Gockel for feedback on an earlier version of the manuscript.

3

Effects of polarity on the perception of interaural time differences in cochlear implant listeners[a]

Abstract

Animal physiological recordings suggest that for a fixed firing efficiency, the neural response to cathodic stimulation has a longer latency than with anodic stimulation (Miller et al. 1999). Since the delay they reported (approx. 200 μs) is in the range of perceivable interaural time difference (ITDs), we investigated whether changing the polarity of the stimulus was enough to affect ITD-based localization. Six subjects balanced in their right ear the loudness of quadraphasic pulses that had either a cathodic or anodic central phase. The binaural stimuli were built with the loudness-balanced anodic or cathodic quadraphasic pulse at the right ear, and a quadraphasic anodic pulse at the left ear, with ITDs varying from -1200 to +1200 μs in 400 μs steps. The subjects subsequently ranked all stimuli on a left-right axis using a mid-point comparison procedure. Results showed only idiosyncratic effects of polarity on the perception of ITD, and no effect at a group level. Furthermore, for some subjects, changing the loudness balancing method had a significant effect on the lateralization results. This suggests that small loudness cues might shadow effects of the polarity on the latency of the neural response.

[a] This chapter is based on Guérit et al, submitted.

3.1 Introduction

While the application of hearing prostheses such as cochlear implants (CIs) leads to substantial improvements for hearing-impaired listeners regarding speech intelligibility in quiet, communication in complex environments remains challenging (Friesen et al., 2001; Nelson et al., 2003; Stickney et al., 2004). Moreover, despite numerous attempts during the last decades to improve the signal processing in CIs, only small effects regarding speech intelligibility have been reported, and spatial hearing remains difficult (Chang et al., 2010; Kan et al., 2015; Kerber and Seeber, 2012; Loizou et al., 2009). One significant contribution to improve performance would be to provide reliable information about interaural time differences (ITDs) for bilateral recipients, i.e., giving a clear cue of which ear the sound arrived at first. Normal hearing (NH) listeners make use of ITDs to separate a target from a spatially separated interferer (e.g., Ihlefeld et al., 2012). CI listeners benefit much less from ITDs with current stimulation strategies. Hence, improving the ability to separate a target from an interferer using ITDs might help CI listeners in communicating in complex environments.

In principle, most CI listeners can locate sounds based on ITDs (for recent reviews, see Kan et al., 2015; Laback et al., 2015). However, compared to those of NH listeners, just noticeable differences (JNDs) are still much higher in CI listeners, (median of 144 µs and 11.5 µs for CI and NH listeners, respectively, Figure 3.2 in Laback et al., 2015). It is reasonable to assume that this discrepancy, at least partly, stems from differences in the auditory nerve responses between the two groups. For example, with electrical stimulation, the auditory nerve response exhibits stronger phase locking than with acoustical stimulation (Dynes and Delgutte, 1992; Hartmann et al., 1984). This unnatural response might be a detrimental input for ITD-specific neurons and hence lead to worse performance in localization (Colburn et al., 2008; Ihlefeld et al., 2015) and in other tasks where ITD cues can be used, for example when speaker and interferers are spatially separated.

Stimuli-specific factors in electrical stimulation can affect the temporal properties of the auditory nerve response, and hence perception of ITDs. For example, single-neuron recordings in cats showed that at low stimulation levels, cathodic

currents elicited a neural response delayed by approximately 200 μs relative to the response evoked by anodic currents (Miller et al., 1999b). This delay decreased with increasing level and was associated with a larger jitter for cathodic polarity compared to anodic polarity. It has been proposed that this delay and difference in jitter result from activating different sites (i.e., nodes of Ranvier) within the neurons for the different polarities and levels (Briaire and Frijns, 2006; Miller et al., 1999b; Rattay et al., 2001b). Similar delays (approx. 150 μs) might occur in humans, as shown by the latency of electrically evoked auditory brainstem response (eABR, Undurraga et al., 2013). It is yet unclear if those differences in latency for different polarities are significant enough to influence the perception of ITDs in CI listeners.

In experiment 1, we presented pulses at both ears with various ITDs, and investigated if changing the polarity of the pulse in the right ear changed the perceived location of the stimulus (Figure 3.1A). As the latency difference between anodic and cathodic stimulation is level dependent in both, animal recordings and human eABRs, we investigated the effect of polarity on ITD at two different levels on the dynamic range. To remove the contribution of an interaural level difference (ILD) cue to localization, we loudness-balanced the pulses with opposite polarity. In experiment 2, we investigated the stability of the effects of polarity against small ILD cues that could have been created by an imperfect loudness balancing between the pulses with opposite polarity.

3.2 Experiment 1: Methods

3.2.1 Listeners

Six participants, all bilateral recipients of Cochlear CIs, took part in the study. The Danish Science-Ethics Committee approved this experimental procedure (ref. number H-16036391), and all listeners signed a participation agreement before data collection. Table 1 shows details on the etiology of those listeners.

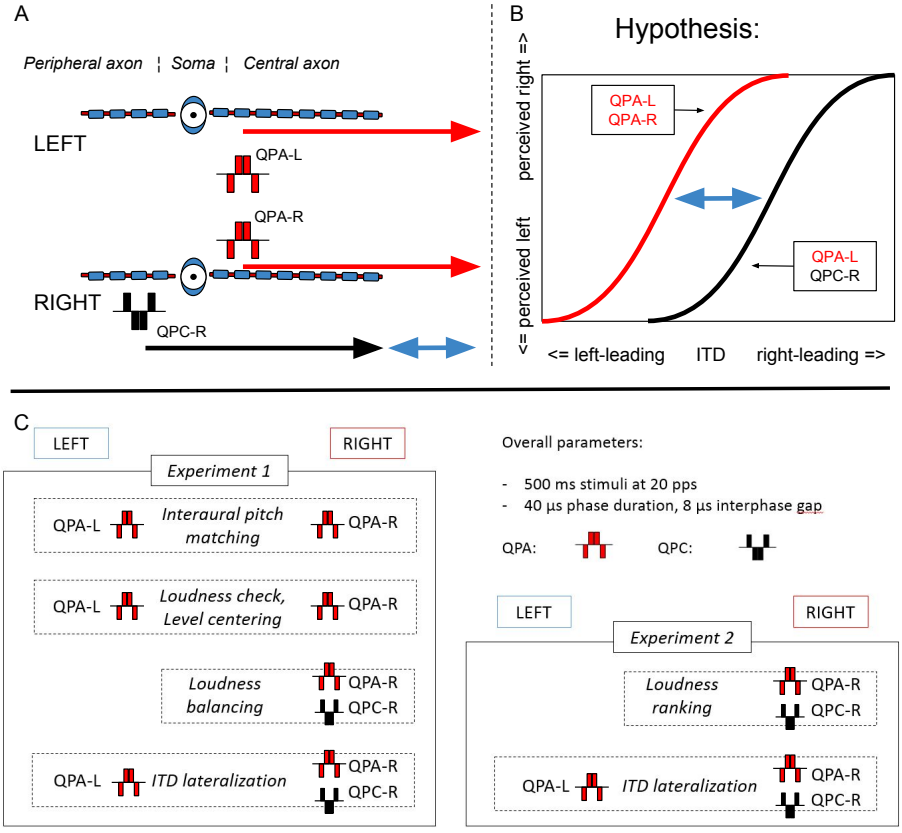

Figure 3.1: **A.** Cathodic stimulation has been suggested to elicit a response more peripherally than anodic stimulation (Miller et al., 1999b), with a latency difference of around 200 µs (blue arrow). **B.** Changing the polarity of a pulse in one ear should therefore shift the perceived location of a binaural stimulus. **C.** Schematics of the experimental paradigms. QPA-R and QPC-R likely have different loudness, hence the necessity of loudness balancing them to remove loudness cues in the lateralization task. Experiment 2 investigates the stability of the lateralization results against a small error in the loudness balancing. This was investigated by using a different method for the balancing, based on ranking.

Table 3.1: Demographics of the CI listeners. CIC3 and CIC4 implants have different chips, which differences in the spacing and values of the current steps. Oscilloscope recordings with test implants showed no difference in terms of bilateral synchronization when using different implants on each side.

Listener ID	Age (y)	Duration of implant use (y)		Electrode used		Implant type		Onset of hearing loss
		Left	Right	Left	Right	Left	Right	
S1	75	14	8	18	18	CIC3	CIC4	Post-lingual
S2	65	8	16	21	18	CIC4	CIC3	Post-lingual
S3	45	6	5	17	17	CIC4	CIC4	Pre-lingual
S4	74	2	4	18	19	CIC4	CIC4	Post-lingual
S5	59	3	2	10	10	CIC4	CIC4	Post-lingual
S6	68	1	3	17	17	CIC4	CIC4	Post-lingual

3.2.2 Setup and stimuli

We collected data using direct stimulation, with a dedicated research platform (RFGenXS box, courtesy of Cochlear Ltd) and software (NIC3, Cochlear Ltd; PPS toolbox, Guérit, 2018) to bypass the clinical speech processor of the CI users. This equipment allowed for perfectly synchronized bilateral stimulation. Some listeners had a different implant version (CIC3 and CIC4) in each ear, which differed in the available range of levels and phase durations. Recordings with a digital oscilloscope and corresponding test implants showed no difference regarding bilateral synchronization when using the RFGenXS box.

We ensured not to present any electrical charge over 212 nC (Shannon, 1992). We measured impedances for each electrode at the beginning and end of each testing session, in order to keep all stimulation below limits of compliance (7 V). For each listener, the selected electrodes were in a region with overall low and homogeneous impedances for both ears.

For all experiments, stimuli consisted of 20-pps trains of quadraphasic pulses with a total duration of 500 ms. These were generated by concatenating two biphasic pulses (separated by an 8-µs gap) with opposite leading polarity (Figure 3.1). Each biphasic pulse had a phase and interphase duration of 40 and 8 µs, respectively. With this configuration, we assumed that most of the neural excitation was caused by the 80-µs long central phase (Carlyon et al., 2013; Karg et al., 2013; Macherey and Cazals, 2016; Macherey et al., 2017). We therefore refer to the pulses with the central phase being anodic as QPA, and with the central phase being cathodic as QPC. When combined for the ITD task, we always presented a QPA stimulus on the left side ("QPA-L"), while presenting either a QPA or a QPC stimulus on the right side ("QPA-R" and "QPC-R," respectively).

3.2.3 Overall procedure

Figure 3.1C shows the experimental workflow. To measure the difference in lateralization between QPA-R and QPC-R stimuli in experiment 1, we first picked a binaural pair of place-matched electrodes using an interaural place-pitch matching task. This leads to lower ITD JNDs in both NH listeners (Henning, 1974; Scharf

et al., 1976) and CI listeners (Hoesel, 2004; Long et al., 2003; Poon et al., 2009). We ensured that with 0-μs ITD, the level of QPA-L was such that, when combined with QPA-R, the stimulus had a centered image and a comfortable loudness. We then balanced the levels such that QPA-R and QPC-R had the same loudness (there is typically a difference of 1 to 2 dB between QPA and QPC pulses in humans, see Carlyon et al., 2013). We created a pool of bilateral stimuli by combining either QPA-L with QPA-R, or QPA-L with QPC-R, with different ITDs. Finally, in the lateralization task, we presented successively two of those bilateral stimuli, and asked the listeners whether they could perceive a change in location.

3.2.4 Electrode Pitch Matching

We picked two and eight electrodes on the left and right side, respectively. These had the same geometrical center, e.g. electrodes 8-9 on the left and electrodes 5 to 12 on the right. The listeners first ranked all these single electrodes in terms of pitch, with QPA stimuli presented at their most comfortable level (MCL). For the ranking procedure, we used an optimally efficient mid-point comparison (MPC) procedure (Long et al., 2005) with ten trials. We then analysed the results as follows: for each electrode on the right side, we computed the percentage of trials having a higher rank than each of the two electrodes on the left side. We fitted a probit line to these curves to estimate the 50 percent point of subjective equality (PSE, Ihlefeld et al., 2015). From these two curves (one for each reference electrode on the left), we picked a pair of electrodes for which the PSE was the closest to a real electrode. We used this pair of electrodes (cf. Table 1) in all following experiments[a].

The pulse rate used (20 pps) was below the lower limit of temporal pitch (Krumbholz et al., 2000), and therefore avoided idiosyncratic temporal pitch cues to bias

[a] For listeners S1 and S2, we obtained the probability of perceiving the right electrode higher in pitch by using a slightly different method. In each presentation, the listeners compared the same electrode (on the left side) to varied electrodes on the right side (same as Ihlefeld et al., 2015). Even though we used two different ranges of electrodes on the right side in different blocks, as recommended by Ilhefeld et al., S2 consistently picked the left electrode to be the highest in pitch in most trials. For the following listeners, we therefore switched to the MPC procedure as described above to avoid such procedural effects. For listener S2, as a backup, we picked a pair of electrodes previously derived from another study (Janssen et al., 2017) that used an MPC procedure.

the results of place (i.e. electrode) pitch matching (Carlyon et al., 2010). We kept this low rate for all subsequent experiments, at the exception of listener S5, for whom we used 100 pps to stay within compliance limits of the device.

3.2.5 Loudness balancing between QPA-R and QPC-R

In each run of loudness balancing, the listener heard two stimuli and indicated on a chart whether the first or second stimulus was the loudest. The experimenter adjusted the level of the second stimulus to find the point of subjective equality (PSE), taking care to bracket several times above and below it. We computed the final value from the average of four runs, two of them with QPA-R being the reference, and two of them with QPC-R being the reference.

We repeated this procedure at two levels on the dynamic range, equivalent to levels 3 and 6 ("soft" and "most comfortable," respectively) on a loudness-scaling chart having 11 values. Before the monaural loudness balancing at MCL, we checked that the simultaneous presentation of QPA-L and QPA-R did not yield any uncomfortable loudness caused by binaural loudness summation (Blamey et al., 2000; Kordus and Żera, 2017).

3.2.6 ITD lateralization

We created a pool of fourteen bilateral stimuli, by combining either QPA-L with QPA-R, or QPA-L with QPC-R, with seven different ITDs. For listener S1, we picked a range of ITDs from -600 to +600 μs, in 200-μs steps (negative ITDs for left-leading stimuli). For all the following listeners, we used twice that range (-1200 to +1200 in 400-μs steps), as some of them showed larger ITD JNDs.

The listeners ranked all the stimuli from left-most to right-most. This was done with an MPC procedure with ten trials. For each stimulus presentation, listeners heard two bilateral stimuli separated by a 500 ms gap. The task was to indicate whether the second stimulus was perceived left or right from the first stimulus.

We then computed the percentage of trials where the QPC-R stimuli ranked to the right (i.e., higher) of the QPA-R stimuli as a function of the difference in ITD between QPC-R and QPA-R. We merged the seven curves (one for each reference

QPA-R stimulus) and fitted a binomial distribution to obtain the PSE (i.e., the relative ITD at which listeners perceived both stimuli at the same location). The fit was done with the glmfit function (Matlab, Mathworks Inc.) using a probit link, and taking into account the variable number of presentation for each relative ITD.

3.3 Experiment 1: Results

3.3.1 Loudness matching between opposite polarities

Figure 3.2 shows the results of the loudness balancing between QPA-R and QPC-R stimuli. A multilevel approach (Field et al. 2012) showed no effects of condition ("soft" vs "MCL", $\chi^2(1) = 2.35$, $p = 0.12$). When averaging both conditions together, QPA-R stimuli required significantly higher levels than the QPC-R stimuli (+2.2 dB, $t(4) = 4.72$, $p = 0.009$), in order to elicit the same loudness.

3.3.2 Lateralization results

Left panels of Figure 3.3 show the mean ranks (± one standard error) obtained at MCL for the lateralization task. Ranks go from 1 (left-most) to 14 (right-most), the number of stimuli. Note that the MPC procedure included all stimuli in each of the ten trials (i.e., both polarities), and we can therefore directly compare the ranks. We fitted a mixed-effects linear model to the mean ranks (see Kuznetsova et al. 2015), with the ITD and polarity of the stimulus as fixed effects and listener-related effects as random effects. This showed an effect of ITD ($F(1, 31.4) = 382$, $p < 0.001$), but no effect of polarity ($F(1, 7.59) = 0.587$, $p = 0.467$) nor an interaction between ITD and polarity ($F(1, 28.4) = 0.179$, $p = 0.675$). Some listeners (S1, S2, and S6) ranked the QPC-R stimuli consistently higher than (i.e., to the right of) QPA-R stimuli, while the opposite occurred for listeners S4 and S5. Listener S3 could not perform the task.

Right panels of Figure 3.3 show the results of fitting a binomial distribution to the proportion of QPC-R stimuli ranked higher (i.e., to the right) of QPA-R stimuli, as a function of the ITD between QPC-R and QPA-R. The shaded area shows the

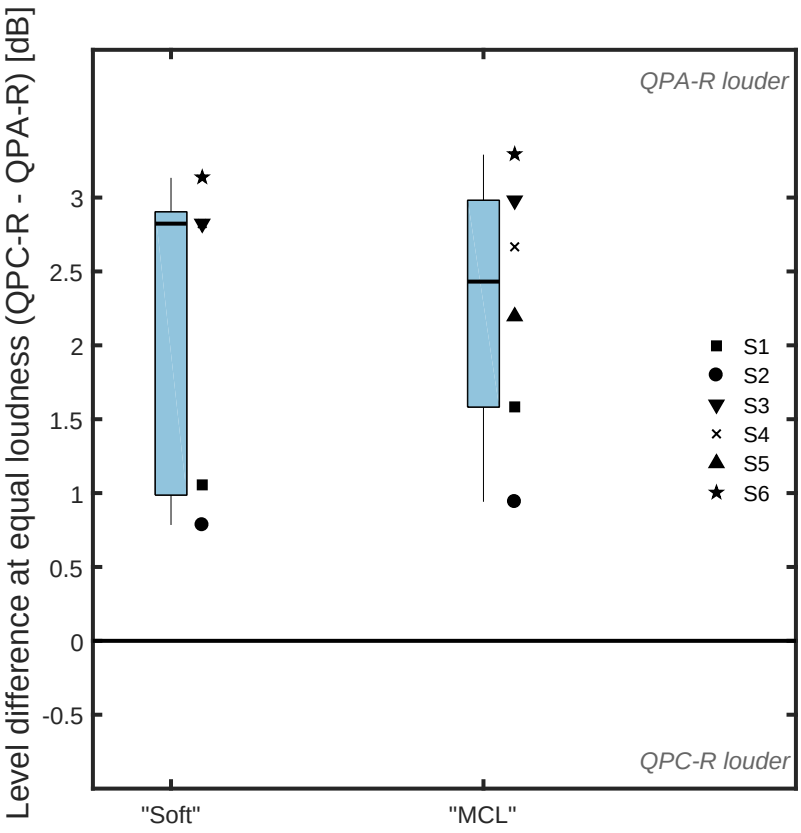

Figure 3.2: Results of the loudness balancing between QPA-R and QPC-R pulses. Positive values indicate that QPA-R pulse are louder than QPC-R at the same level.

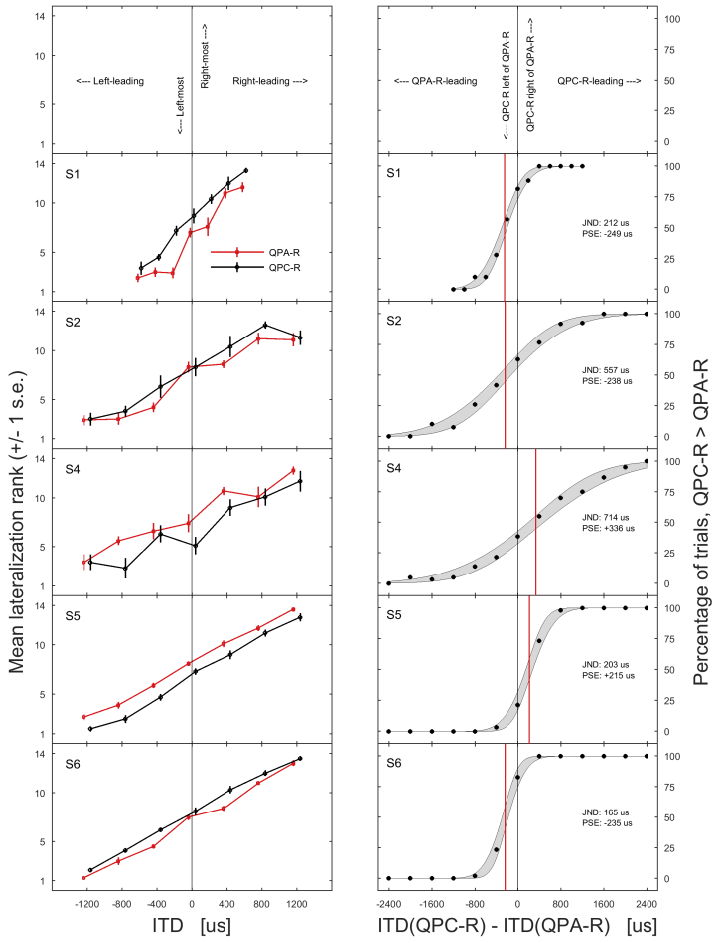

Figure 3.3: Results of the lateralization task at MCL. **Left panels:** mean ranks (+/- 1 standard error), with lowest and highest ranks corresponding to the stimuli being perceived left-most and right-most, respectively. **Right panels:** percentage of the QPC-R stimuli ranked higher than (i.e. more to the right of) QPA-R stimuli, as a function of the ITD between QPC-R and QPA-R. Grey-shaded area shows 95% confidence interval of the binomial fit, red vertical line shows the 50% of the fit (point of subjective equality between both stimuli). Just noticeable differences were computed as half the difference between the 25 and 75% point of the binomial fit.

95% confidence interval of the fit, and the vertical red line marks the PSE (i.e. the 50% point of the fit). PSEs with a negative value reflect that QPA-R had to lead QPC-R in order to elicit the same location, when combined with a QPA-L stimulus at the left ear. Positive values reflect the opposite (i.e. QPC-R had to lead QPA-R in order to elicit the same location). We quantified ITD JNDs by computing half of the difference in ITD between the 25% and 75% of the curve (similar as Francart et al., 2009). Excluding listener S3 that could not do the task, the average JND was 370 µs, with a standard deviation of 249 µs.

Figure 3.4 shows the same lateralization results as Figure 3.3, but when presenting stimuli at the "soft" level. Similar to the "MCL" condition, statistical analysis of the mean ranks showed an effect of ITD ($F(1, 24.7) = 55.3$, $p < 0.001$), no effect of polarity ($F(1, 5.54) = 3.49$ $p = 0.115$), and no interaction between ITD and polarity ($F(1, 22.6) = 0.052$, $p = 0.822$). There was a significant interaction between ITD and random effects of listeners ($\chi^2(1) = 17.3$, $p < 0.001$) and a significant interaction between polarity and random effects of listeners ($\chi^2(1) = 13.9$, $p < 0.001$). JNDs were either higher than for the "MCL" condition or even non-measurable (no 25% and/or 75 % point in the binomial fit). This reflects that the task was more challenging at these lower levels.

Figure 3.5 summarizes the PSEs and their 95% confidence interval obtained from the individual binomial fits, excluding conditions where we could not estimate JNDs. For the two listeners who could reliably do the task at both levels, the PSE was numerically higher (i.e. more towards QPC-R leading) in the "soft" condition, compared to the "MCL" condition.

3.4 Experiment 2

Effects of polarity in experiment 1 might stem from an imperfect loudness balancing between QPA-R and QPC-R stimuli. That is, in order to elicit the same location, QPA-R would require to be leading QPC-R because of being quieter than QPC-R (and vice-versa). In a second experiment, we therefore used a different balancing procedure, which involved ranking the loudness of the two QPA-R stimuli ("Soft"

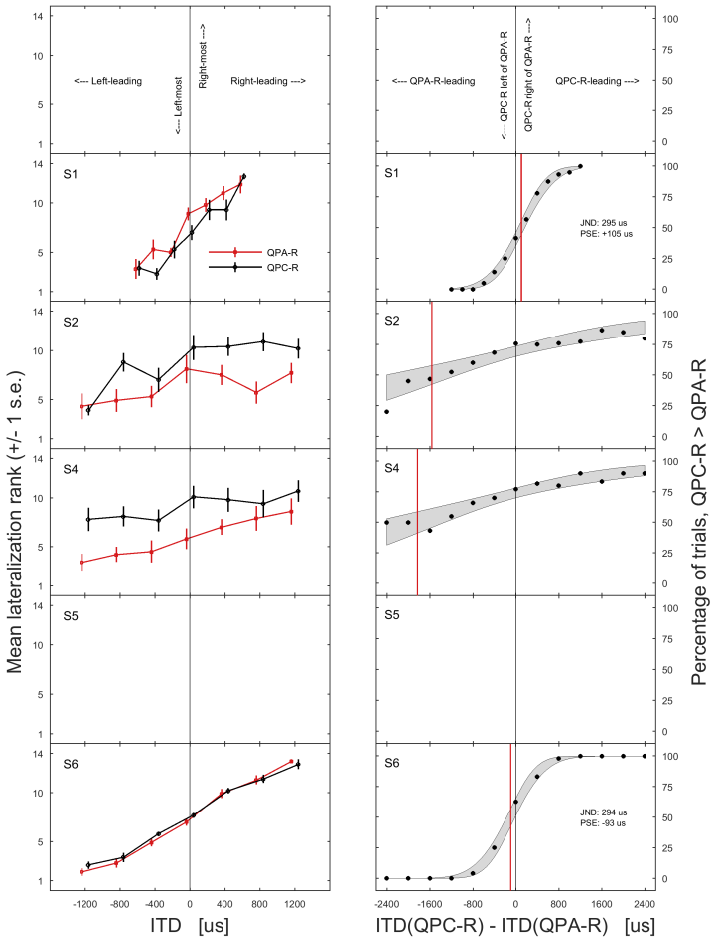

Figure 3.4: Results of the lateralization task at the "soft" level. Left panels: Mean ranks (+/- 1 standard error), with lowest and highest ranks corresponding to the stimuli being perceived left-most and right-most, respectively. Right panels: percentage of the QPC-R stimuli ranked higher than (i.e. more to the right of) QPA-R stimuli, as a function of the ITD between QPC-R and QPA-R. Grey-shaded area shows 95% confidence interval of the binomial fit, red vertical line shows the 50% of the fit (point of subjective equality between both stimuli). Just noticeable differences were calculated as half the difference between the 25 and 75% point of the binomial fit.

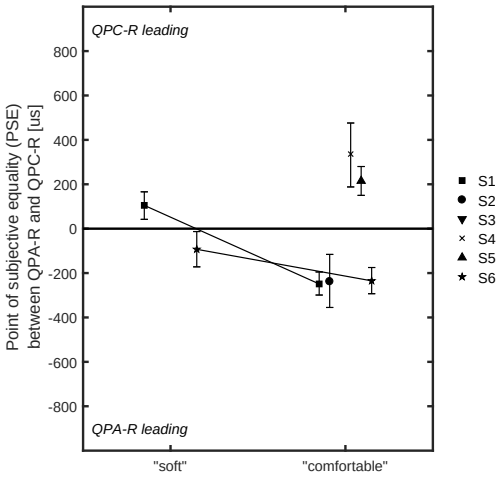

Figure 3.5: Summary of the points of subjective equality between QPA-R and QPC-R, across the two level conditions. Positive values indicate that QPC-R had to lead QPA-R in order to elicit the same perceived location (and vice-versa).

and "MCL"), and eleven QPC-R stimuli with levels distributed across the dynamic range. The objectives were three-fold: 1) to measure the effects of changing the balancing procedure on the obtained levels; 2) to evaluate the effects of such differences on the lateralization results; 3) to check for the monotonicity of the QPC loudness growth. Indeed, Macherey et al. (2017) recently reported non-monotonic QPC loudness growths in 40 % of their electrodes and subjects.

3.4.1 Methods

The same listeners participated in experiment 2 (except S3 for the lateralization task). We measured the QPC-R loudness growth at the right ear by dividing the dynamic range in eleven steps[b] and ranking the stimuli using a MPC procedure (similar as in Macherey et al., 2017). For the MPC procedure, we used ten repetitions

[b] For listener S4, a very narrow dynamic range (10 CU) only allowed for ten steps

and added the two QPA-R stimuli ("soft" and "MCL") to the list of stimuli being ranked. These had the same levels used in the previous lateralization task.

For each of the two QPA-R stimuli, we derived a new matched level for QPC-R. To do so, we fitted a binomial distribution to the percentage of MPC trials where the QPC-R stimuli were ranked higher (i.e. louder) than the QPA-R stimuli. For almost all listeners, the obtained PSEs differed slightly from the values obtained with the original loudness balancing procedure (cf. Figure 3.6). We therefore ran the lateralization procedure again with the level of QPC-R set at the PSE, only for the "MCL" condition.

3.4.2 Results

Left panels in Figure 3.6 show the individual results of the loudness ranking procedure. The extent of the dynamic range varied largely across the listeners, from 1.5 (S4) to almost 10 dB (S5-S6). We found one incidence of a significant non-monotonicity in the QPC-R loudness growth for listener S6 (Figure 3.6). This was revealed by paired-sample t-tests on the individual ranks between the levels i and i + 1 (Macherey et al., 2017). This non-monotonicity was below the levels we used in all other procedures (loudness balancing and lateralization).

The right panels in Figure 3.6 show the percentage of QPC-R stimuli ranked louder than the "soft" QPA-R stimulus (blue) and louder than the QPA-R stimulus at "MCL" (black). The thick vertical line shows the PSE (50 % of the binomial fit), while the dashed line shows the value derived using the original loudness balancing procedure (Lexp1). Fitting a mixed-effects linear model to those matched levels (in % of the dynamic range) showed only an effect of condition ("Soft" vs "MCL", $F(1, 16.0) = 162$, $p < 0.001$), but no effect of experiment ("Exp 1" vs "Exp 2", $F(1, 15.0) = 0.225$, $p = 0.642$) nor an interaction between experiment and condition ($F(1, 14.0) = 0.0017$, $p = 0.967$). For some listeners, the difference between the two experiments amounted for a fair amount of the dynamic range (e.g. 14 % for S2 at the "MCL" condition).

Figure 3.7 shows the PSEs between QPA-R and QPC-R for the lateralization experiments 1 and 2. For listener S1 and S6, changing the level changed moderately

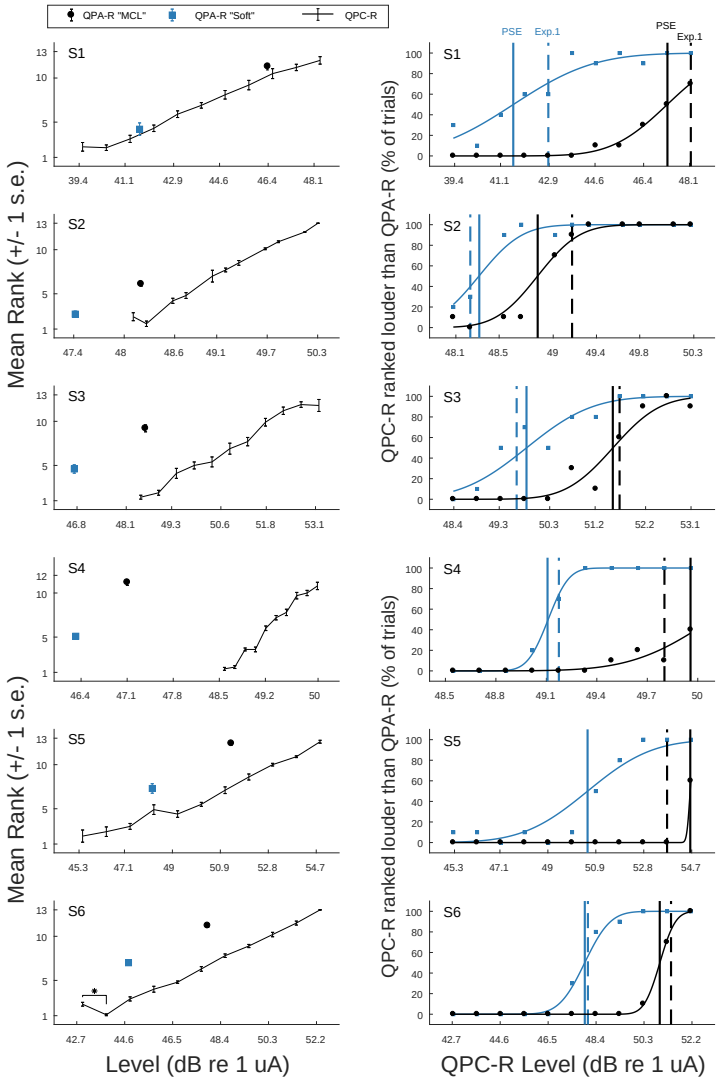

Figure 3.6: Results of the loudness ranking, experiment 2. Left panels: Mean ranks (+/- 1 standard error) of the QPC-R stimuli, with levels distributed along the dynamic range. Two QPA-R stimuli were also included, at the levels used in experiment 1 ("Soft", in blue, and "MCL", in black). Subject S6 exhibited a non-monotonicity at the lower end of the QPC-R loudness growth. Right panels: Percentage of QPC-R stimuli perceived louder the QPA-R stimuli (blue: "Soft" QPA-R as a reference; black: "MCL" QPA-R). The vertical, thick line indicates the 50% point (PSE) from the binomial fit. The dashed vertical line indicates the result of the loudness balancing in experiment 1.

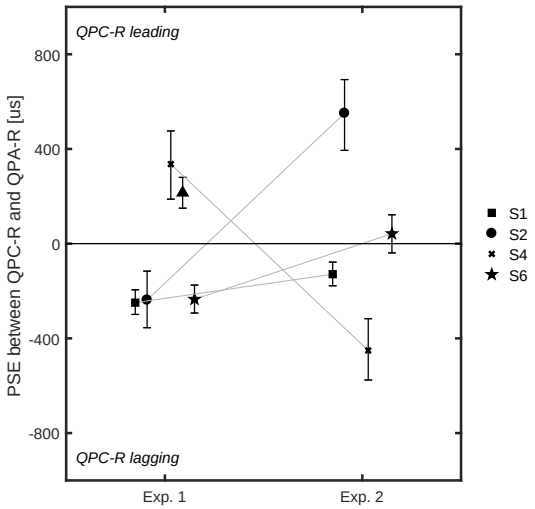

Figure 3.7: Individual lateralization results across experiments 1 and 2. Shown are the fitted point of subjective equality between QPA-R and QPC-R. A positive value indicates that QPC-R had to lead QPA-R in order to elicit the same location. In both cases, the stimuli on the right ear were combined with a QPA stimulus on the left side (QPA-L)

the PSE (+120 µs and +277 µs, respectively) across both experiments. For listeners S2 and S4 however, changing the level changed the PSE dramatically by +789 and -787 µs, respectively.

Figure 3.8 shows the difference in PSE between experiment 1 and 2 as a function of the difference in level between experiment 1 and 2 (in percent of the dynamic range). From ITD-ILD trading experiments, one would expect that for QPC-R to elicit the same location, reducing its level would need to be compensated for by having QPC-R leading. There was indeed a significant correlation between both measures (only 4 listeners, $r^2 = 0.85$, $F(1, 3) = 16.46$, $p = 0.270$). The slope of this correlation was 50.1 µs per level change of 1% of the dynamic range.

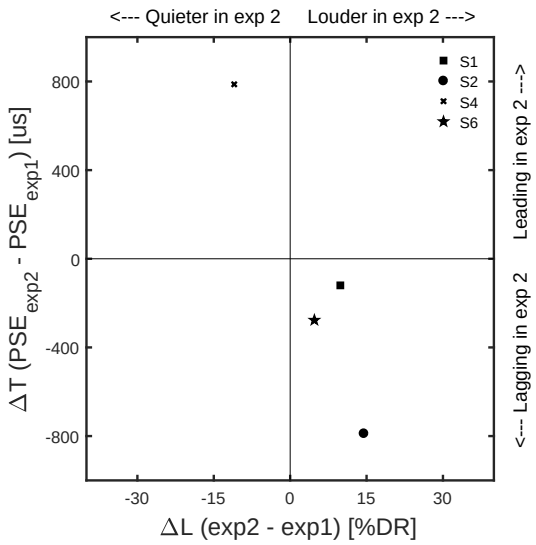

Figure 3.8: Effect of changing the level of QPC-R on the PSE between QPA-R and QPC-R. From ITD-ILD trading experiments, one would expect that for QPC-R to elicit the same location, reducing its level would need to be compensated for by having QPC-R leading. Such trend is present with the four subjects we collected data with (slope of 50.1 µs per level change of 1 % of the dynamic range).

3.5 Discussion

3.5.1 Loudness differences between QPA-R and QPC-R

In experiment 1, averaged across listeners, QPC-R stimuli required 2.2 dB more current to elicit the same loudness as QPA-R stimuli. This is consistent with previous studies that used quadraphasic pulses and a similar loudness balancing procedure (Carlyon et al., 2013; Macherey et al., 2017). Moreover, changing the balancing procedure in experiment 2 yielded similar results. Rattay (1999) suggested that absent and/or demyelinated peripheral axons could cause such decreased efficiency of cathodic currents. Their model suggests indeed that cathodic currents are more likely to depolarize nodes of Ranvier near the electrode and to target peripheral axons.

Macherey et al. (2017) reported the presence of a non-monotonic loudness

growth with level in 40 % of their listeners and electrodes when using QPC stimuli. That is, within a short region of the dynamic range, increasing the level decreased the overall loudness. This would violate the assumption of a monotonic loudness growth that we used in the loudness-balancing task in experiment 1. However, for the listeners and electrodes reported here, increasing the level of QPC-R stimuli only led to a non-monotonic loudness growth function for listener S6, at the lower end of the dynamic range (Figure 3.6). It is therefore unlikely that a non-monotonic behaviour biased the loudness balancing in experiment 1.

3.5.2 Effects of ITD and polarity on lateralization

In experiment 1, all listeners but S3 perceived a change in location with ITD. Listener S3 reported having a pre-lingual onset of hearing loss, which has been shown to correlate with poor ITD perception (Litovsky et al., 2010). Excluding that listener, mean JNDs at MCL were 370 μs, similar to values previously reported in the literature (e.g., Laback et al., 2015). The ITD task was more difficult at lower levels ("soft" condition), leading to larger or non-measurable JNDs (Figure 3.4). Such effect of level is similar as reported by Hoesel (2007) and Egger et al. (2016). This unfortunately limits the interpretation of our data in regards to the effects of polarity at different stimulus levels.

On a group level, changing the stimulus polarity did not change its perceived location (Figures 3.3, 3.4, 3.5 and 3.7). It did have an effect, however, on an individual basis. In experiment 1 and at MCL, listeners S4 and S5 needed QPC-R stimuli to lead QPA-R stimuli by respectively 336 μs (± 144 μs, 95% c.i.) and 203 μs (± 65 μs) in order to be perceived at the same location (red lines in Figure 3.3). This is consistent with the hypothesis presented in Figure 3.1, whereby the neural response to QPC stimulation is delayed compared to that of QPA. However, listeners S1, S2 and S6 showed the opposite effect, whereby QPC-R stimuli had to be delayed by respectively 249 μs (± 52 μs), 238 μs (± 120 μs) and 235 μs (± 59 μs) to elicit the same location as QPA-R stimuli. For those listeners, this suggests that the neural response to QPC-R stimuli led that of QPA-R, in opposition to previous reports in animals (Miller et al., 1999b), in humans (Undurraga et al., 2013) and modeling

(Rattay et al., 2001b). Such individual effects going in both directions also occurred in experiment 1 at the "soft" level, and in experiment 2.

The size of the individual polarity effects was large for some listeners (above ± 400 µs, Figure 3.8). This is above the reported effects of polarity on the response latency in single-neuron recordings (Miller et al., 1999b). Our results might, therefore, not correlate directly with a difference in the neural response latency. For example, a bias might come from the contribution of the flanking phases in the QP stimuli. Indeed, even a weak neural contribution of the flanking phases could shift the perceived ITD by a few tenths of microseconds. Furthermore, small loudness differences between the two polarities might bias the results, as discussed in the following section.

3.5.3 Effects of changing the loudness balancing procedure

At a group level, results of the loudness-balancing between QPC-R and QPA-R stimuli were similar in experiment 1 (manual balancing) and experiment 2 (ranking and binomial fitting). However, in some listeners, differences amounted up to 15 % of the dynamic range. This might create enough of an ILD cue to bias the results in the lateralization experiment. One indication for such bias is the correlation between the difference in level and the difference in PSE between experiments 1 and 2 (based on only four listeners). The slope of this correlation is surprisingly high (50 µs for a level change of 1 % of the DR). This suggests that an ILD cue as small as 5 % of the dynamic range corresponds to an ITD cue of 250 µs (which is the expected size of expected polarity effects). More testing is however required to confirm this poor ITD-ILD trading, when compared to NH listeners (David et al., 1959).

The absence of a "true" reference makes it difficult to judge whether one of the two loudness-balancing methods is more accurate than the other. However, a few aspects suggest that the loudness balancing procedure in experiment 1 is more prone to procedural bias. First, the experimenter decides on a final value after bracketing several times around the PSE. This gives room for interpretation, particularly for shallow loudness functions, where changing the level does not

change much the perceived loudness. The value obtained from the loudness ranking is thus more objective, as derived from a binomial fit. Furthermore, the loudness-ranking task is likely clearer in showing which percept to focus on (and to ignore pitch cues for example). This is because the two signals being compared are constantly changing, and because the differences in loudness are often very clear (comparison between extremes on the dynamic range). Finally, with the manual balancing, the smallest step size (1 CU in this study, approx. 0.13 dB) is typically the same for all listeners, and does not reflect the variability in the size of the dynamic range across listeners.

3.6 Conclusion

With quadraphasic pulses, there was no consistent effect of polarity on the perceived localization using ITD cues. The absence of a consistent effect could not be explained by nonmonotonic loudness growth functions in individual listeners. Either, the quadraphasic pulses with opposite polarity do not show the same latency differences as monophasic pulses, or small deviations in the loudness balancing provide dominating ILD cues.

Acknowledgements

The authors would like to thank Ruth Litovsky for helpful comments during the design of the experiment.

4

Effects of stimulation order of opposite-polarity pulses on loudness for cochlear implant listeners[a]

Abstract

Most cochlear implant (CI) strategies use charge-balanced rectangular biphasic pulses. However, both anodic and cathodic currents can elicit action potentials, and possibly do so at different sites on the spiral ganglions neurons. Here, we investigated the effect of the order of anodic and cathodic stimulation on loudness at short (0 to 800 μs) inter-pulse intervals (IPIs). We used pairs of pseudomonophasic (PS) pulses to mimic a biphasic pulse where we could manipulate the amplitude of each phase independently. In experiment 1 the two opposite-polarity PS pulses had the same loudness, thereby preventing either of the two PS pulses from dominating the percept. Six users of the Advanced Bionics CI loudness-ranked trains of the pulse pairs using a mid-point comparison procedure. Stimuli with anodic leading polarity were louder than with cathodic-leading polarity for IPIs shorter than 400 μs. This effect was small - about 0.3 dB - but consistent across listeners. When running the same procedure with both PS pulses having the same level, anodic-leading stimuli were still louder than cathodic-leading stimuli at very short IPIs. However, when using clinical, symmetric, biphasic pulses, the effect disappeared at short IPIs and reversed at long IPIs.

[a] This chapter is based on Guérit et al. (2018).

We discuss possible peripheral sources of such polarity interactions.

4.1 Introduction

In normal-hearing listeners, action potentials (APs) in response to sounds are generated at the very peripheral end of the spiral ganglion neurons (SGNs) that constitute the auditory nerve (Kim and Rutherford, 2016). For cochlear implant (CI) users however, electrical current can theoretically elicit APs at both the peripheral and central axons of the SGNs (Honert and Stypulkowski, 1984; Javel and Shepherd, 2000).

4.1.1 Latency distribution in animal recordings

One method to determine whether action potentials have been generated at the peripheral or central process of the SGNs is to compare the latency of APs elicited by electric pulses of different intensity and polarity (Honert and Stypulkowski, 1984; Javel and Shepherd, 2000; Miller et al., 1999b; Undurraga et al., 2013). Javel and Shepherd (2000) measured single-neuron spike latencies at the level of the Inferior Colliculus (IC) in cats, and observed a multimodal distribution of latencies. They attributed these different latencies to different generation sites, including the hair cells and the peripheral and central processes of the auditory nerve. They estimated the latency difference between spikes elicited at peripheral and central processes to be in the range of 100 to 200 μs.

Changing the polarity of the electrical stimulus can also alter spike latencies: Miller et al. (1999b) measured cat single-neuron responses to cathodic and anodic monophasic pulses, presented in monopolar mode (with the ground outside the cochlea). Responses to cathodic currents exhibited longer latencies and lower thresholds than for anodic currents, suggesting that cathodic currents evoke APs more peripherally than anodic currents. This is consistent with modelling work of Rattay et al. (2001a), based on observations from Ranck (1975), which suggests that a locally positive second derivative of the voltage along the axons of the SGNs can trigger APs. The location of those areas of positive second derivative changes

with polarity, being near the electrode with cathodic currents and further away for anodic currents (Ranck, 1975).

The afore-mentioned studies suggest that anodic currents activate the neurons more centrally than cathodic currents. Thus, by using anodic versus cathodic currents, one could target central and peripheral processes, respectively. It could also be that both polarities excite nodes of Ranvier on the same side (either peripheral or central) of the soma. Miller et al. (1999b) hypothesized that most of the neurons they studied had been excited for both polarities along the central axons. Cartee et al. (2006) suggested a greater peripheral activation, at least with cathodic currents.

4.1.2 Polarity studies in human CI listeners

Monophasic pulses cannot be used in humans, because the charge imbalance would cause electro-chemical damage to the tissues in the cochlea (Brummer and Turner, 1977). However, the effect of stimulus polarity has been studied using triphasic or asymmetric biphasic pulses (Carlyon et al., 2013; Macherey et al., 2006, 2008, 2010; Undurraga et al., 2013). Using those pulses, it has been shown that anodic currents are more efficient (i.e. require less current) than cathodic currents in eliciting a response at comfortable levels (Macherey et al., 2008). The difference between the two polarities is greatest at supra-threshold levels (Undurraga et al., 2013), and is consistent across devices and listeners (Carlyon et al., 2013). At threshold, although there is no overall difference between the two polarities, significant and consistent differences can be seen across listeners and electrodes (Macherey et al., 2017). Only one study reported an effect of polarity on latency in human CI listeners (Undurraga et al., 2013). In that study, the latency of the wave V of the electrically evoked auditory brainstem response (eABR) was significantly longer for cathodic than for anodic currents (153 μs in average, at equal level between anodic and cathodic stimulation). The difference was largest at lower levels.

4.1.3 Perceptual effects of stimulation at different sites

Because APs elicited by peripheral and central sites are likely to interact and to arrive at the brain with different latencies, they potentially disrupt the information coded in the timing of the neural response. In addition, the hyperpolarisation of a central site on a neuron may affect the propagation of spikes elicited at a peripheral site, and this could increase the current needed for the stimulus to be heard and/or to reach a comfortable loudness (Macherey et al., 2017). Overall, knowledge on the site of activation may reveal the pattern of neural survival close to each electrode, allowing the audiologist to select subsets of electrodes for stimulation (for a review of the importance of cochlear health across electrodes, see Pfingst et al., 2015).

In the present study, we examined the interactions between the effects of anodic and cathodic stimulation at short (below 800 μs) inter-pulse intervals (IPIs) on loudness. One obstacle to doing so is that, for many stimuli such as the symmetric biphasic pulses used clinically, the anodic phase is likely to dominate the loudness. Therefore, experiment 1 used a paradigm with pairs of equally loud opposite-polarity pseudomonophasic pulses (Figure 4.1). We measured the perceived loudness as a function of the order of those pulses and of the duration of the IPI. Experiments 2 and 3 studied the same interactions with stimuli where the levels, rather than the loudness, of cathodic and anodic stimulation were equated (see Figure 4.1). We hypothesized that a difference in site of AP generation with polarity would create order effects for the perceived loudness of anodic and cathodic currents presented sequentially.

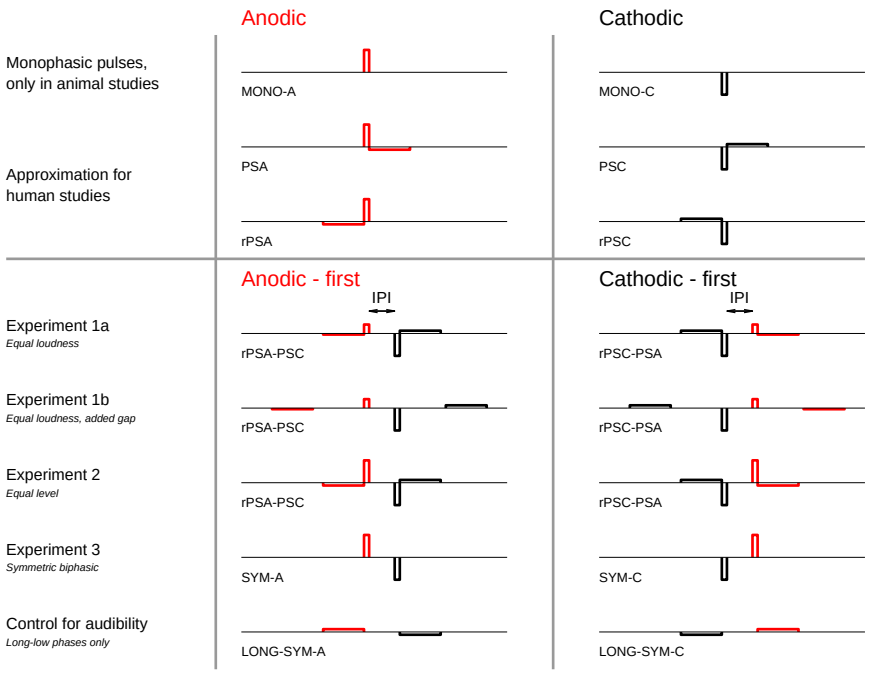

Figure 4.1: **Top panel**: Pulse shapes commonly used for polarity studies (e.g. for humans, pseudomonophasic anodic and cathodic, respectively PSA and PSC). Reverted version of PSA and PSC are labelled with a "r" (rPSA and rPSC). **Bottom panel**: Two-pulse stimuli used for the different experiments of this study. By using pairs of pseudomonophasic pulses, we could mimic biphasic pulses having different levels for each phase, while staying charge-balanced.

4.2 Experiments 1a and 1b: Equally loud asymmetric pulses

4.2.1 Methods

Listeners

The listeners were five post-lingually deaf recipients of an Advanced Bionics CI (including one bilateral CI user), amounting to six ears being tested. Their details

are shown in Table 4.1. Listeners were recruited both in Cambridge (UK) and Copenhagen (DK) and the experimental procedure was approved respectively by the National Research Ethics Committee for the East of England (ref. number 00/327) and the Danish Science-Ethics Committee (ref. number H-16036391). All listeners signed a participation agreement before data collection.

Setup and safety

All data collection was achieved by means of direct stimulation, using research hardware and software that bypassed the clinical speech processor of the CI user. Current levels were limited by ensuring that the voltage at the electrode stayed below limits of compliance (7 V in the HiRes90k Advanced Bionics implant) and that charge density stayed below 100 μC/cm2 (Litovsky et al., 2017). Stimuli were checked using a test implant and digital storage oscilloscope. Impedance checks were performed at the beginning and end of each testing session.

Stimulus

The stimuli consisted of pseudo-monophasic pulses, with a 43-μs short-high phase preceded (rPSA, rPSC) or followed (PSA, PSC) by a 344-μs 1/8th amplitude phase of opposite polarity (Figure 4.1). With such asymmetric pulses, most neural excitation comes from the short-high phase (Miller et al., 2001a; Undurraga et al., 2013). We therefore refer to the asymmetric pulses with the short-high phase being anodic or cathodic as the "anodic" and "cathodic" pulse, respectively.

A two-pulse paradigm (rPSA-PSC and rPSC-PSA, Figure 4.1) allowed us to adjust the relative level of each pulse so that both polarities elicited an equal loudness when presented separately (experiment 1). Those anodic-first and cathodic-first two-pulse stimuli were created with eight different IPIs of 0, 50, 100, 200, 400 and 800 μs. One subject, AB1, was additionally tested at an IPI of 1600 μs. For all experiments, a single electrode (number 9 or 10) in the middle of the array was used, and each two-pulse group was presented at a 100-Hz repetition rate for a duration of 400 ms.

Table 4.1: Demographics of the CI listeners. All listeners were post-lingually deaf recipients of an Advanced Bionics HiRes90k device. 1j is a straight array, Helix is perimodiolar.

ID	Age [yrs]	Duration of implant use at surgery [yrs]	Electrode used [yrs]	CI Side	Electrode array	Etiology
S1-L	60	9	9	left	1j	Pendred Syndrom
S1-R		10	9	right	Helix	
AB1	72	8	9	left	1j	Unknown
AB2	57	9	9	left	1j	Ototoxicity
AB3	71	10	9	left	1j	Otosclerosis
AB5	75	7	10	left	1j	Otosclerosis

Loudness matching of the single pulses

We initially obtained the Most Comfortable Levels (MCLs) for single anodic (PSA) and cathodic (PSC) pulse trains using a categorical loudness-scaling chart. The chart had 11 points ranging from "inaudible" (1) to "too loud" (11); MCL was defined as point 7 ("most comfortable"). Combining the single pulses, each at their MCL, into a two-pulse stimulus likely yields a louder percept. We therefore measured the MCLs of the two-pulse stimuli, where each pulse had an equal level. We used the lowest of all afore-mentioned MCLs as a reference for the loudness matching procedure of the single pulses. In this procedure, the experimenter adjusted the level of a signal stimulus (either rPSC, PSA or rPSA, Figure 4.1), and the subject indicated whether the sound was quieter, louder or at the same loudness as the reference stimulus (PSC). In each run, the experimenter obtained the point of equal subjective loudness by bracketing several times around it. The final value was computed from the mean difference (in dB) of two runs, with the PSC pulse train being the reference in one run and the adjusted stimulus in the other. The resulting equally-loud pulses were then combined into the two-pulse (rPSA-PSC and rPSC-PSA) stimuli shown in Figure 4.1. Finally, we checked that none of these levels caused loudness to exceed the MCL for any of the IPIs in the two-pulse stimuli.

Loudness ranking

Anodic-first and cathodic-first two-pulse stimuli at all IPIs were loudness ranked using the optimally efficient Mid-Point Comparison algorithm (MPC, Long et al., 2005). This procedure was repeated twelve times, in two blocks of six repetitions. A single PSC pulse was included in the loudness-ranking procedure for listeners AB3, S1-L, S1-R and AB5. This PSC pulse had the same level as in the rPSA-PSC two-pulse stimulus, and the same loudness as all other component pulses of the two-pulse stimuli. This tested whether both pulses contributed to the overall loudness. If the two-pulse stimuli were louder than their component single pulses, we could conclude that both pulses contributed to loudness. If the two-pulse stimuli were not louder, the results would be inconclusive: either one pulse dominated loudness,

or both pulses contributed but partially counteracted each other, for example by charge cancellation.

Loudness matching of the paired pulses

As loudness ranking only gives a relative indication (which stimuli are louder than others), additional loudness matching was run between the two-pulse stimuli with opposite polarity at IPIs of 50 and 200 μs. The difference (in dB) needed to equate loudness was computed from the average of four runs (two runs with anodic-first as the reference, two runs with cathodic-first), where the experimenter bracketed the level around the point of subjective equality. The level difference (in dB) between anodic and cathodic pulses was kept constant throughout the procedure.

Experiment 1b: interphase gap

Even though we assume that most of the neural excitation comes from the short-high phases in our stimuli, the long-low phases could theoretically influence the results as well, for example by interacting with the short-high phases. To control for this, experiment 1b repeated the loudness-balancing procedures from experiment 1a with five of the listeners and added a 600-μs inter-phase gap between the long-low and the short-high phase of each pulse (cf. stimulus in Figure 4.1, experiment 1b).

4.2.2 Results

Loudness matching of the single pulses

Figure 4.2.A shows the results of matching the loudness of rPSA, PSA and rPSC pulse trains to a PSC pulse train in experiment 1a. A two-way (polarity vs "reversing" of pulses) repeated-measures ANOVA on the levels in dB re 1 uA showed a significant effect of polarity ($F(1, 5) = 49.9$, $p < 0.001$), reversing ($F(1, 5) = 10.81$, $p = 0.022$) but no interaction ($F(1, 5) = 1.58$, $p = 0.265$). Post-hoc t-tests showed a significant difference between anodic (rPSA, PSA) and cathodic (rPSC, PSC) stimuli amounting to 2.1 dB ($t(5) = 7.06$, $p < 0.001$), and a small but significant difference of 0.1 dB

between the reversed versions of each pulse (rPSA and rPSC quieter than PSA and PSC when pooled together, $t(5) = 3.29$, $p = 0.0218$).

When adding an extra 600-μs inter-phase gap between the long-low and short-high phases (Experiment 1b, Figure 4.2.B), the difference between anodic and cathodic pulses was similar (2.1 dB versus 1.9 dB in experiment 1a for all listeners who performed both experiments). A repeated-measures ANOVA including the single pulse levels from experiments 1a and 1b showed a significant effect of polarity ($F(1, 4) = 36.2$, $p = 0.004$), an interaction between polarity and reversing the pulses ($F(1, 4) = 20.8$, $p = 0.01$) and an interaction between experiment, polarity and reversing ($F(1, 4) = 13.5$, $p = 0.02$). These interactions reflect the fact that, to reach the same loudness, PSA needed more current than rPSA in experiment 1b but not in experiment 1a.

Loudness ranking and matching of the two-pulse stimuli

Mean ranks[a] and standard errors across trials for all listeners are shown in Figure 4.3. As some listeners did not have the same number of conditions in this experiment, ranks were scaled between 1 and 10 for comparison across listeners. Note that, although the anodic- and cathodic-first data are plotted in separate panels, all stimuli were loudness-ranked together as part of the same MPC procedure.

A repeated-measure ANOVA on the mean ranks (excluding the single PSC stimulus) showed significant effects of polarity ($F(1, 5) = 131.1$, $p < 0.001$) and of IPI ($F(5, 25) = 113.8$, $p < 0.001$), and a significant interaction between IPI and polarity ($F(5, 25) = 34.72$, $p < 0.001$). Quietest stimuli were to be found at lowest IPIs (0 and 50 μs). Interestingly, all two-pulse stimuli with an IPI over 0 μs were louder than the single PSC pulse, indicating that both pulses must contribute to loudness. At 0 μs, the two-pulse stimuli had a similar loudness to that of the single PSC pulse. Anodic-first stimuli were always ranked louder than the corresponding cathodic-first stimuli for IPIs ranging from 0 to 400 μs (Figure 4.4.A).

[a] The distribution of ranks across trials deviated in some occasions from normality, particularly for stimuli ranked at the loudest or quietest end of the range. Running the statistical analysis with the median ranks instead did not change the main conclusions.

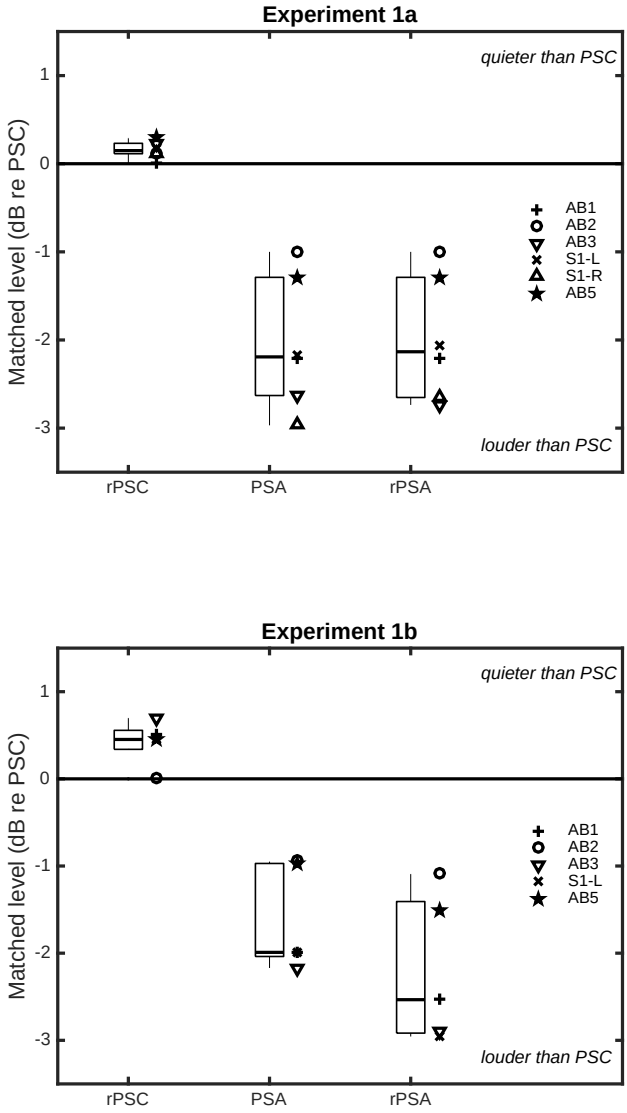

Figure 4.2: **A**. Levels of the single pulses used in experiment 1a, relative to the PSC pulse. Anodic pulses required in average 2.1 dB less current to elicit the same loudness than cathodic pulses. **B**. Levels used in experiment 1b. Subject S1-R did not participate in that experiment.

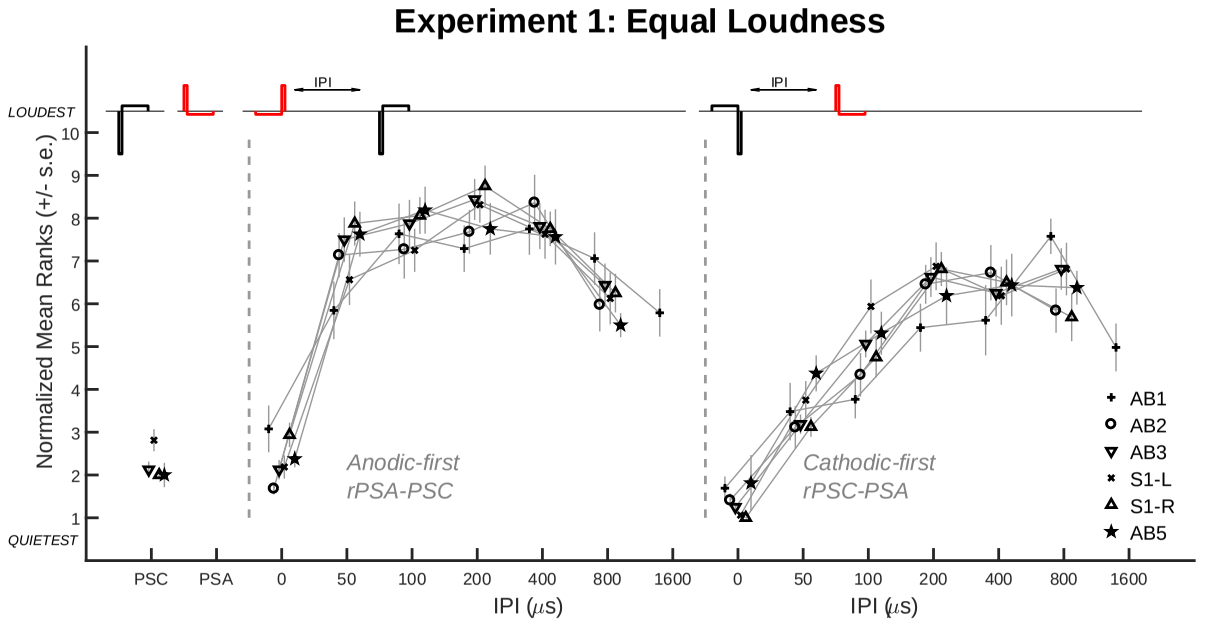

Figure 4.3: Results of the loudness ranking procedure for experiment 1 (equal loudness between anodic and cathodic stimulation). Ranks for each subject were scaled between 1 and 10. Single PSA was not included in this experiment, as it was loudness matched to PSC.

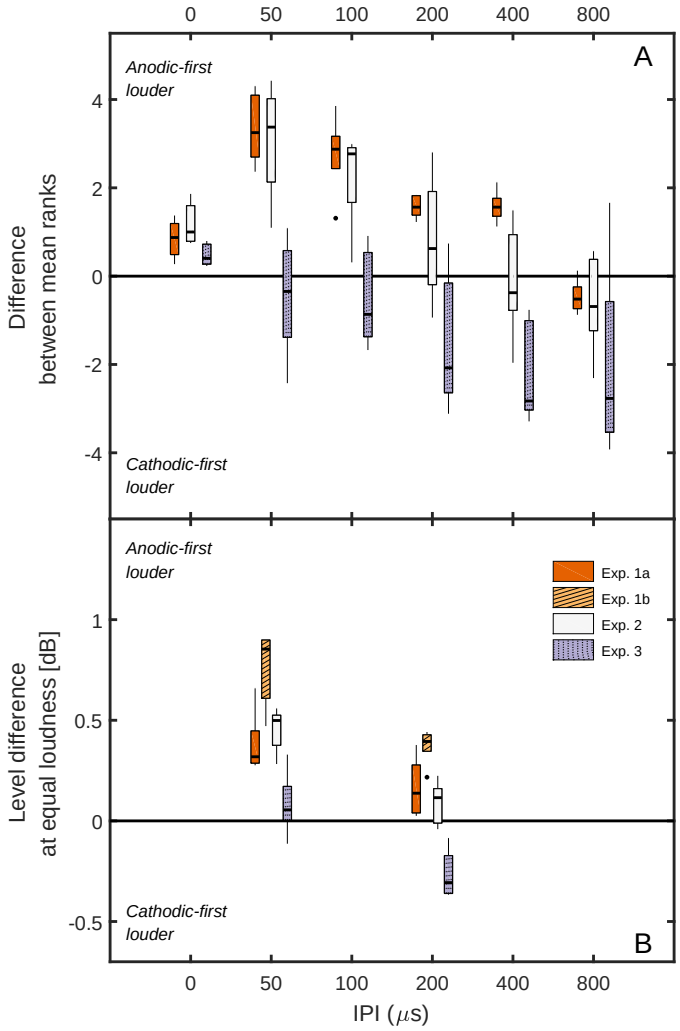

Figure 4.4: **A**. Difference between the mean ranks obtained with anodic-first versus cathodic-first stimuli. The boxes show the distribution of individual results (N = 5, subject S1-R not shown in experiment 1a), with positive values indicating a higher rank given to anodic-first stimuli. Experiment 1a: equal loudness between cathodic and anodic pulses in isolation. Experiment 2: equal level. Experiment 3: symmetric biphasic pulses. **B**. Results of the loudness matching between anodic-first and cathodic-first stimuli at 50- and 200-μs IPI. Positive values indicate that anodic-first stimulus is louder than cathodic-first. Experiment 1b: equal loudness between cathodic and anodic pulses in isolation, with a 600-μs inter-phase gap between the long-low and short-high phases of the pseudomonophasic pulses. *Lower and Upper limits of the boxes*: 25th and 75th percentiles of the ranks. *Horizontal black line*: median rank. *Whiskers*: 25th (or 75th) percentile minus (or plus) 1.5 the interquartile range. Dots correspond to data points with values outside the range delimited by the whiskers.

Subsequent loudness matching (Figure 4.4.B) between the two-pulse stimuli at 50- and 200-μs IPI confirmed that the anodic-first stimuli were louder than cathodic stimuli (average of 0.28 dB when pooling both IPIs, $t = 4.71$, $df = 9$, $p = 0.001$). The difference was numerically larger at 50- than 200-μs IPI (0.38 dB vs 0.17 dB, respectively), but did not differ significantly between the two IPIs ($F(1, 4) = 3.01$, $p = 0.16$).

Loudness matching results, experiment 1b

In experiment 1b, we added an inter-phase gap of 600 μs between the long low and short high phases (Figure 4.1, experiment 1b), and performed loudness matching at 50- and 200-μs IPIs. Results are shown in Figure 4.4B, and exhibit the same trend, whereby anodic-first stimuli were louder than cathodic-first stimuli (averaged across IPIs, $t(4) = 9.25$, $p < 0.001$). There was a significant effect of IPI ($F(1, 4) = 12.8$, $p = 0.023$) and experiment ($F(1, 4) = 16.4$, $p = 0.015$) on the level differences between anodic- and cathodic-first pulses, but no interaction between experiment and IPI ($F(1, 4) = 3.4$, $p = 0.15$). The main effect of IPI reflects that when combining results from experiments 1a and 1b, the difference between anodic-first stimuli and cathodic-first stimuli becomes larger at 50- than 200-μs IPI ($t(4) = 3.57$, $p = 0.023$). The main effect of experiment reflects that overall, the difference between opposite polarities stimuli was larger in experiment 1b than 1a ($t(4) = 4.05$, $p = 0.016$).

4.3 Experiments 2 and 3: equal level and symmetric biphasic pulses

In experiment 1, the loudness was matched between anodic and cathodic asymmetric pulses, with level differences between each polarity amounting to 2.1 dB on average (ranging from 1.1 to 2.9 dB). It is therefore possible that the effects observed in experiment 1 were driven by the relative levels of the first and second pulses, rather than by the polarity of each pulse. Therefore, experiment 2 presented both pulses at the same level. We would then expect most of the excitation to arise

from the anodic pulse. Experiment 3 used symmetric biphasic pulses (SYM-A and SYM-C, Figure 4.1), similar to those used clinically. SYM-A and SYM-C were effectively the same stimuli as in experiment 2, but without the flanking long-low phases, and again, we expected most of the excitation to come from the anodic phase. In the particular case of experiment 3, changing the IPI is equivalent to changing the inter-phase gap of a symmetric biphasic pulse.

4.3.1 Methods

Listeners

The same listeners as in experiment 1b participated in experiments 2 and 3. Subject S1-R had an implant failure that precluded her from further listening experiments.

Loudness ranking and matching

Trains of anodic-first and cathodic-first pulse pairs with IPIs ranging from 0 to 800 μs, as well as trains of single PSC pulses were ranked from softest to loudest using the same ranking procedure as described in experiment 1 (with 12 repetitions). In addition, trains of single PSA pulses were included in the loudness ranking for listeners AB1 and AB3 in experiment 2. This was done because we expected the PSA pulse to be louder than the PSC pulse, and so that we could determine whether the PSA pulse dominated the loudness of the two-pulse stimuli. The PSA pulse was also included for all listeners in experiment 3. For both experiments loudness matching was performed for IPIs of 50 and 200 μs.

Detection thresholds of the long-low phases

To assess the possibility that the long low phases contributed to loudness in experiment 2, we measured the detection thresholds of those long-low phases in isolation. The stimulus was a biphasic pulse with 344-μs phase duration and an inter-phase gap of 140 μs, corresponding to the interval between the long-low phases in the paired-pulse stimuli of the main part of experiment 2 when the IPI was 50 μs. We used a 2-alternative forced-choice procedure, with a 1-up-3-down rule. Each run

consisted of two reversals with a 1-dB step size, followed by six reversals with a 0.25- dB step size. We measured the thresholds twice for each leading polarity, averaging from the last six reversals in each run.

4.3.2 Results

Figure 4.5 shows the mean ranks (and standard errors) obtained with the stimuli of experiment 2. The anodic pulse (PSA) was ranked louder than the cathodic pulse (PSC), consistent with the results of experiment 1a (e.g. Figure 4.2.A). For the two listeners tested with the PSA pulse, the loudness was roughly equal to that of the maximum obtained with any of the two-pulse stimuli. Hence, unlike in experiment 1, we cannot conclude that the cathodic pulse increased the overall loudness of any of the two-pulse stimuli. A repeated-measure ANOVA (with only the paired pulses) showed significant main effects of polarity ($F(1, 4) = 11.56$, $p = 0.027$), IPI ($F(5, 20) = 75.45$, $p < 0.001$) and an interaction between polarity and IPI ($F(5, 20) = 9.14$, $p < 0.001$). Anodic-first stimuli were ranked louder than cathodic-first stimuli, but this was only the case for all listeners at 0, 50 and 100-µs IPI (Figure 4.4.A). Figure 4.4.B shows the results of the subsequent loudness matching at 50- and 200-µs IPI. There was a significant effect of IPI on the level difference between equally loud anodic- and cathodic-first pulses ($F(1, 4) = 35.7$, $p = 0.0039$). This reflects that anodic-first stimuli were louder than cathodic-first stimuli by 0.45 and 0.09 dB at 50- and 200-µs IPI, respectively.

Figure 4.6 shows the mean ranks for all listeners when using clinical-like, symmetric biphasic pulses and single PSA and PSC pulses in the pitch-ranking procedure. The symmetric biphasic pulses had either the anodic (SYM-A) or cathodic (SYM-C) phase leading. Repeated-measure ANOVA performed on the ranks given to the biphasic pulses (without PSA and PSC) showed a significant main effect of IPI ($F(5, 20) = 48.88$, $p < 0.001$) and an interaction between polarity and IPI ($F(5, 20) = 3.50$, $p = 0.020$), but no main effect of polarity ($F(1, 4) = 6.81$, $p = 0.059$). Unlike the results of experiments 1 and 2, SYM-A and SYM-C were ranked similarly up to 100-200-µs IPIs, while SYM-C was ranked louder than SYM-A at 400- and 800-µs IPIs (Figure 4.4A). This is also reflected in the loudness matching results (Figure

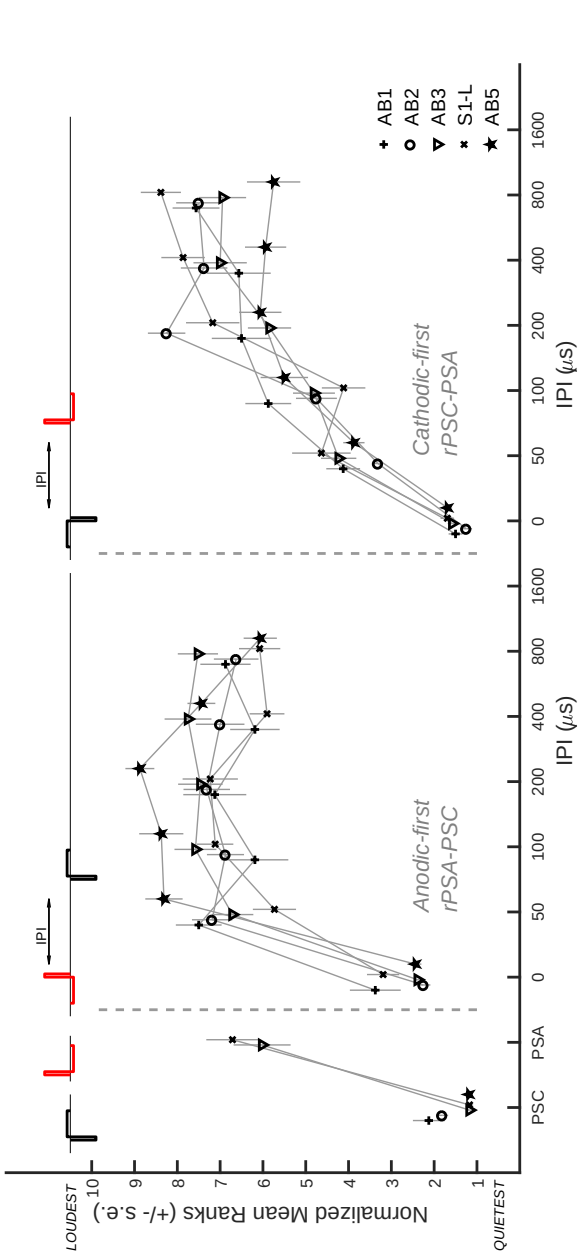

Figure 4.5: Results of the loudness ranking procedure for experiment 2 (equal level between anodic and cathodic stimulation). Mean loudness ranks for each subject are scaled between 1 and 10. As expected from previous studies with human CI listeners, PSA is louder than PSC at equal level. When combined, the two-pulse stimuli are similar in loudness to PSA in isolation.

4.4B). Similar to experiment 2, there was a significant effect of IPI ($F(1, 4) = 74.7$, $p < 0.001$), but this time the anodic-first pulses were only louder by 0.09 dB at 50-µs IPI, and cathodic-first pulses were louder by 0.26 dB at 200-µs IPI.

At 0-µs IPI, SYM-A was ranked quieter than PSA (Figure 4.6), as expected if the long-low cathodic phase of the PSA pulse does less charge cancellation of the anodic phase, compared to the charge cancellation produced by the short-high cathodic phase of SYM-A. Loudness of SYM-C increased even for IPIs above 100 µs (which was the longest interval used by McKay and Henshall, 2003).

The left pair of bars in Figure 4.7 shows the levels of the long-low phases of the stimuli used in experiment 2 and with a 50-µs IPI, relative to the detection thresholds measured here. It can be seen that, in experiment 2, the long-low phases were 5.4 dB above their detection thresholds in isolation ($t(4) = -3.68$, $p = 0.02$). This level was however equivalent to the step 2 on the loudness scaling chart (a.k.a. "just noticeable"), as shown by the right-hand pair of bars.

4.3.3 Across-experiment comparisons

Experiments 2 and 3 differed only by the presence of the long-low phases, which, based on the data shown in Fig. 7, should not contribute substantially to the overall loudness. A repeated-measures ANOVA on the ranking results across those two experiments (excluding single pulses) showed an effect of experiment ($F(1, 4) = 285$, $p < 0.001$). This reflects that the paired pulses had overall higher ranks than the single PSA in experiment 3, but not in experiment 2. There was an interaction between polarity and experiment ($F(1, 4) = 9.34$, $p = 0.0378$), consistent with anodic-first stimuli being overall louder in experiment 2, and quieter in experiment 3. Although there was a trend for anodic-first stimuli to be louder at short IPIs in experiment 2, and quieter at long IPIs in experiment 3, there was no interaction between IPI, experiment and polarity ($F(5, 20) = 2.12$, $p = 0.11$).

There was a significant effect of experiment for the loudness matching results across experiments 2 and 3 (Figure 4.4B, $F(1, 4) = 19.62$, $p = 0.01$). This indicates an effect of experiment on the difference in loudness between the two polarities. Finally, there was no interaction in the loudness matching results between IPI and

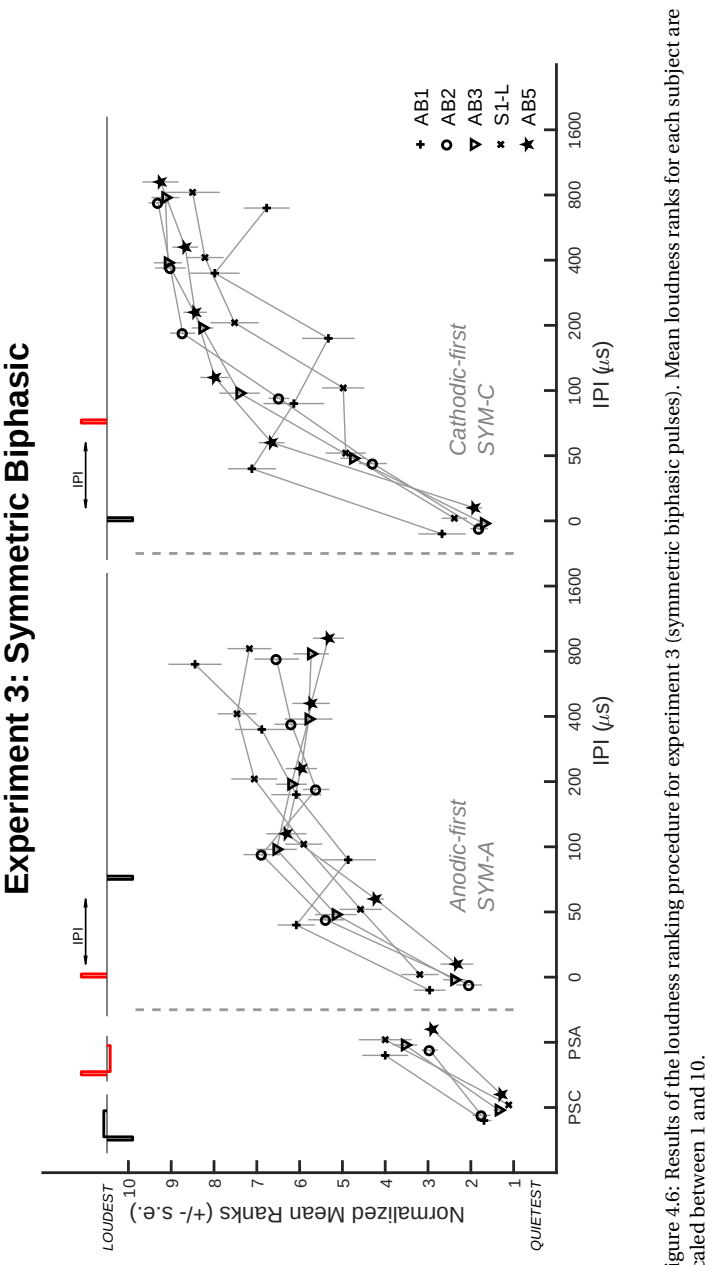

Figure 4.6: Results of the loudness ranking procedure for experiment 3 (symmetric biphasic pulses). Mean loudness ranks for each subject are scaled between 1 and 10.

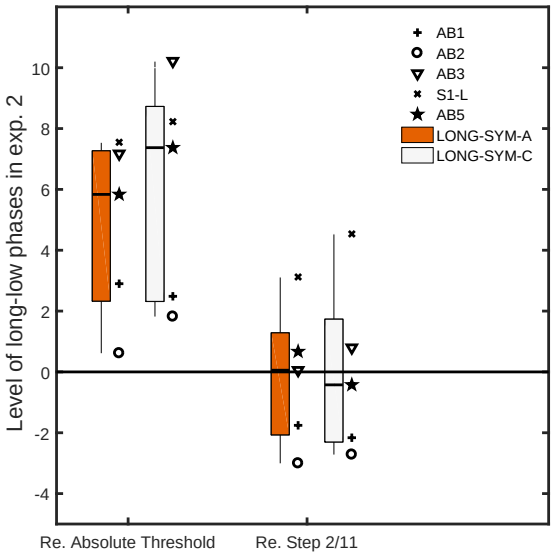

Figure 4.7: **Left**. Levels of the long-low phases used in experiment 2, relative to their detection thresholds in isolation (i.e. without the two central short-high phases). Filled boxes show the results for the first phase of the long-low phases being anodic (LONG-SYM-A), corresponding to the rPSC-PSA stimulus without the short-high phases. Empty boxes show the results for cathodic-leading long-low phases (LONG-SYM-C, or rPSA-PSC without the short high phases). **Right**. Same levels, relative to the "Just noticeable" percept of the long-low phases in isolation, obtained with a loudness-scaling chart (step 2 out of 11).

experiment ($F(1, 4) = 0.05$, $p = 0.83$).

4.4 Discussion

All experiments reported here showed significant effects of inter-pulse interval on loudness. Furthermore, in all experiments, the order of the anodic and cathodic pulses within each pair significantly influenced the loudness. Those order effects were similar across two different tasks, loudness balancing and loudness ranking (Figure 4.4A vs 4B). They occurred at short IPIs (below 200 μs) in experiments 1a, 1b

and 2, where anodic-first stimuli were the loudest. In experiment 3 however, there was only order effects at the longest IPIs, and in the opposite direction (cathodic-first louder).

4.4.1 Order effects at short IPIs (below 200 μs)

The IPIs between 0 and 200 μs, where order effects occurred in experiments 1 and 2, fall well within the 7-ms central integration window proposed by McKay and McDermott (1998). Hence, although central mechanisms may influence the effect of IPI over longer time ranges, the greatest insight into the findings for IPIs below 200 μs can be achieved by considering the different possible types of peripheral interactions. These could be interactions between APs generated by each pulse, or interactions at the neural membrane before any generation of an AP.

In the equal-loudness experiment (1a), the two-pulse stimuli were louder than the single pulse stimuli at all non-zero IPIs, indicating that both pulses must contribute to the overall loudness. Anodic-first pulse pairs were consistently ranked louder than cathodic-first pairs for IPIs below 400 μs (Figure 4.4.A). This effect was small (0.4 dB at 50 μs and decreasing for larger IPIs, Figure 4.4.B) but significant and consistent across the listeners tested here.

When adding an extra inter-phase gap of 600 μs between the long-low and short-high phases (experiment 1b), the order effect was significantly larger (0.76 dB at 50-μs IPI). In experiment 2, there was an order effect similar in amplitude as in experiment 1a, even though the two pulses had the same level. Only when removing completely the long-low phases (experiment 3) did order effects disappear at short IPIs.

In the following, we discuss two phenomena that, in principle, could result in order effects at short IPIs: a spike collision and charge summation at the level of the neural membrane. We also discuss why the effect disappeared when using symmetric biphasic pulses.

Spike collision hypothesis

Anodic stimulation likely generates action potentials (APs) more centrally than cathodic stimulation (Macherey et al., 2017; Miller et al., 1999b; Ranck, 1975; Rattay et al., 2001a; Undurraga et al., 2013). If cathodic stimulation were to create an action potential (AP) at a peripheral node of Ranvier in the SGNs, it would likely be delayed by the presence of the soma (with a high capacitance) between the central and peripheral processes (Adamo and Daigneault, 1973; Liberman, 1984; Robertson, 1976). Assuming that loudness is connected to the number of spikes transmitted from the SGN to the brain, the lower loudness for cathodic-first stimuli in this experiment is therefore consistent with a "collision" hypothesis: APs created at the periphery by the cathodic pulse travel across the soma and get blocked (or block) the APs created more centrally by the anodic pulse. Conversely, for anodic-first stimuli, APs generated by the anodic pulse would propagate centrally, before the APs generated by the cathodic pulse (at the peripheral processes) could catch up. This would increase the chance of APs elicited by both pulses reaching the brain. Even though 50-100 μs is below the average absolute refractory period of 400 μs (Boulet et al., 2016), a small number of neurons might have the ability to fire twice with such short inter-pulse intervals (Miller et al., 2001b).

If the order effects presented here are due to a latency difference between spikes elicited by anodic and cathodic stimulation, then this difference (largest at 50-100 μs) falls within the lowest range of that observed in animal recordings, which is typically 200 μs or more but with a large variability (Figure 4.5 in Miller et al., 1999b).

One phenomenon that the spike collision hypothesis does not take into account is the propagation of spikes from central to peripheral processes, also called antidromic propagation. This would reduce the size of the effects observed here, because the anodic pulse, which excites the central axon, would block the spikes initiated at the peripheral process by the cathodic pulse. Additionally, if the effects of antidromic propagation had a different time course than the main effect, this would disrupt our estimate of the temporal dynamics. This cannot be ruled out, although it is worth noting evidence that antidromic propagation is not stable,

particularly when it comes to traveling across the soma (Brown, 1994).

Charge summation at the membrane

The neural membrane behaves approximately as a leaky integrator (Lapicque, 1907). Furthermore, for SGNs, the time constant of this integrator is estimated to be around or above 100 μs (Balthasar et al., 2003; Cosentino et al., 2015; Kwon and Honert, 2009; Macherey et al., 2007; Middlebrooks, 2004). This is longer than the duration of the short-high phases used in this study. Hence, at short IPIs, the absolute peak value of the transmembrane potential will be larger for the first pulse than the second pulse. In other words, the first pulse will mask/cancel the second pulse.

As shown in Figure 4.2, anodic pulses (PSA and rPSA) required on average 2.1 dB less current to yield the same loudness as the cathodic pulses (PSC and rPSC). This might interact with the cancellation of the second pulse by the first pulse. For anodic-first stimuli in experiment 1, the second pulse might be less cancelled than with the cathodic-first stimuli. This would explain the results in experiment 1. However, this does not explain the presence of similar order effects in experiment 2, where both pulses had the same level.

More complex charge summations might also stem from the multiplicity of nodes of Ranvier on the SGNs and their interconnection (Joucla and Yvert, 2012; Rattay et al., 2001a). For example, hyperpolarization at central nodes by cathodic currents can create a so-called cathodal block (Frijns et al., 1996; Macherey et al., 2017). The order of presentation of anodic and cathodic pulses could affect the presence of such block, and, more generally, affect the integration of charge across the various nodes of Ranvier (Rattay et al., 2001a).

Absence of order effect at short IPIs in experiment 3

In experiment 2, anodic-first stimuli were louder than cathodic-first stimuli by 0.45 dB at 50-μs IPI. In experiment 3, the difference reached only a bare 0.09 dB. Those two experiments differed by the presence or absence of the long-low phases. Long-low phases likely have the following effects: they cancel the charge injected by

the short-high phases, but might additionally elicit a neural response on their own. They might even facilitate and/or interact with the short-high phases, because of active ion channel dynamics (Boulet et al., 2016).

MCLs for the stimuli used in experiment 3 (SYM-A and SYM-C) were lower by 0.7 dB ($t(4) = 3.7$, $p = 0.0212$) compared to the MCLs experiment 2 (rPSA-PSC and rPSC-PSA). Furthermore, the two-pulse stimuli had higher ranks than the PSA pulse (at and above 50-μs IPI) in experiment 3, but not in experiment 2. Both effects suggest the long-low phases did some charge cancellation. When taking the stimulus from experiment 2 with an IPI of 50 μs and removing the short-high phases, the long-low phases were significantly above their detection threshold by an average of 5.4 dB ($t(4) = -3.68$, $p = 0.02$, Figure 4.7). This was however a weak percept, equivalent on average to the step 2 ("Just Noticeable") on our loudness scaling chart. This suggests overall that the long-low phases created more charge cancellation than they added to the overall response.

The long-low phases might, furthermore, affect the ratio of contribution between the anodic and cathodic pulses in our different experiments. For example, it is likely that in experiment 3 the contribution from the cathodic pulse was much weaker than that from the anodic pulse (MONO-C vs MONO-A), and that this difference was larger than in experiment 2 (PSC vs PSA). A very weak cathodic response could therefore explain the absence of order effects at short IPIS in experiment 3, because the anodic pulse would produce nearly all of the excitation.

4.4.2 Effects at longer IPIs (above 400 μs)

At longer IPIs (above 300-400 μs), there were no polarity order effects in experiments 1a and 2, but cathodic-first stimuli were louder than anodic-first stimuli in experiment 3 (Figure 4.4A). Furthermore, in experiment 1a there was a tendency for the overall loudness to decrease after 400 μs.

At those interval durations, the underlying mechanisms are likely to be driven by refractoriness and central integration rather than charge cancellation at the level of the neural membrane (Cosentino et al., 2015; McKay and McDermott, 1998). In other words, there is a higher chance for both pulses to elicit a neural response on

their own, rather than being integrated at the level of the neural membrane.

In experiment 3, the only experiment where a polarity order effect is present at long IPIs, the anodic phase likely elicits a much stronger neural response than its cathodic counterpart (e.g., Undurraga et al., 2013). Following anodic stimulation, a large proportion of the neurons might therefore be in their refractory period, and mask the cathodic response. This is consistent with anodic-first stimuli being ranked quieter (Figure 4.4A) than cathodic-first stimuli. These refractory effects could occur either in the auditory nerve or more centrally. In experiments 1a and 2, the ratio of contribution from each pulse might have been closer to unity, explaining why there were no order effects at the longest intervals.

In experiment 1a only, the overall loudness decreased from 400- to 800-μs IPI. The model from McKay and McDermott (1998) suggests a very shallow decrease in loudness at those IPIs, because of the central integration window. We do not see the same pattern in the other experiments, suggesting again that the anodic pulse was dominating the percept.

4.5 Conclusion

At very short IPIs (below 100 μs) and when equating loudness by means of asymmetric pulses, anodic-first stimulation is louder than cathodic-first stimulation. This effect is in agreement with (but does not prove) a hypothesis based on a difference in latency between anodic and cathodic stimulation. Alternative explanations such as charge cancellation or cathodal blocking can however not be excluded, as they would all affect the loudness judgements in the same direction. A similar result was obtained using asymmetric pulses of equal level, rather than equal loudness.

For symmetric biphasic pulses, no effect of polarity order was observed at very short IPIs. This may be due to the anodic phase dominating the neural response. At longer IPIs the anodic-first stimulus was quieter than the cathodic-first stimulus. This is consistent with the idea that, at these longer IPIs, the polarity order effects are driven by refractoriness.

Acknowledgements

FG is supported by a grant from the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS). The authors would like to thank the listeners for participating in the several sessions of this study.

5

General discussion

Abstract

The mechanisms of charge integration in the electrically activated auditory nerve are still poorly understood (for a review, see Boulet et al. 2016). This thesis focused on better characterizing the temporal charge integration in human recipients of a CI. This was achieved by measuring the loudness, detection thresholds and localization abilities with pulse pairs, while varying the inter-pulse interval (IPI) and the polarity of each pulse. Overall, results showed a variety of IPI and polarity effects, as well as interactions between both factors. Because of the short IPIs used (0 to 800 μs), those effects likely originated at a very peripheral level. This is discussed in section 5.1. Furthermore, the results suggested procedural limitations in the use of asymmetric pulses, that are necessary when studying polarity effects in humans. This is summarized in section 5.2. Finally, section 5.3 proposes further applications of the findings.

5.1 Effects of polarity on temporal charge integration

5.1.1 Single pulses

With human CI listeners, both anodic and cathodic polarities can depolarize the SGNs (Chapter 2 and Undurraga et al., 2013). However, each polarity typically differs in its efficiency, i.e. in how much current is needed to elicit a sound sensation, or the same loudness. Figure 5.1 summarizes the effects of polarity on

loudness and thresholds from Chapters 2 to 4, when using single pulses. At most comfortable levels (MCLs) and for all experiments, anodic pulses (PSA with various inter-phase gaps, QPA) required less current than cathodic pulses (PSC and QPC) for eliciting the same loudness. This effect of polarity did not differ significantly across experiments (MCLs only, $\chi^2(3) = 4.11$, $p = 0.25$), and amounted to an average value of 2.25 dB. This is consistent with previous reports of polarity effects at MCL with different pulse shapes and brands of CIs (Carlyon et al., 2013). We only tested the effects of polarity for detection thresholds in Chapter 2. There, the effects were opposite to that at MCL, with cathodic pulses eliciting lower thresholds than anodic pulses for all subjects (Figure 5.1), i.e. requiring less current to trigger detection. Although no general effects of polarity are usually reported at threshold, a recent study (Mesnildrey, 2017) showed that up to 78% of the tested electrodes and subjects had lower thresholds for anodic pulses.

Two main mechanisms are usually proposed to explain the overall lower efficiency of cathodic stimuli at MCL, and the more individual effects at threshold. First, a modelling approach (Rattay et al., 2001b) suggested that cathodic stimuli activate the neurons close to the electrode. Since peripheral axons are the closest to the electrode, this should lead to lower cathodic thresholds overall. However, peripheral axons are prone to degeneration (Leake and Hradek, 1988). This is illustrated in Figure 5.2. Such degeneration might not be detrimental for detecting cathodic currents, as such task likely requires only a few peripheral axons. At MCL however, the perception of loudness requires a higher number of neurons to spike (Florentine, 2011; Macherey et al., 2007). There, degeneration of the peripheral axons might be detrimental for cathodic stimulation, but not anodic stimulation. Another mechanism for explaining the lower efficiency of cathodic currents at MCL is the presence of a "cathodal block" (Frijns et al., 1996; Macherey et al., 2017; Ranck, 1975), whereby cathodic currents efficiently depolarize the peripheral axons, but hyperpolarize the central axons, thus blocking the propagation of action potentials. According to modelling from Frijns et al. (1996), such central hyperpolarization would only occur at the highest levels of stimulation. This is consistent with cathodic currents being consistently less efficient than anodic currents at MCLs but not at thresholds (Figure 5.1 and Macherey et al., 2006, 2017; Mesnildrey, 2017).

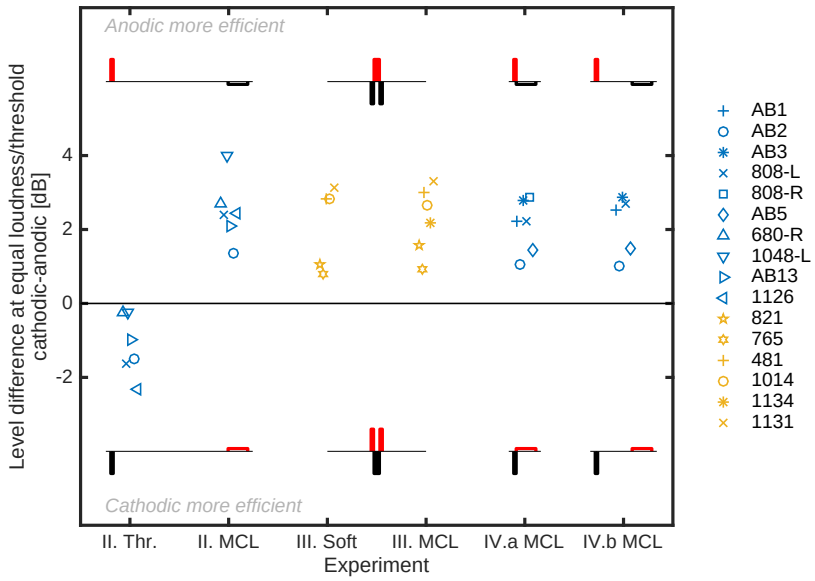

Figure 5.1: Across-chapter comparison of difference in efficiency (detection threshold and loudness matched MCL) between single anodic and cathodic pulses. Chapter IV.a: no inter-phase gap. Chapter IV.b: 600-µs inter-phase gap. Chapter II: 2-ms inter-phase gap. Chapter III: quadraphasic pulses.

5.1.2 Paired pulses

In Chapter 2, reducing the interval between two pulses that had the same polarity increased the loudness in an exponential manner. In Chapter 4, doing the same with two pulses that had an opposite polarity decreased the loudness in an exponential manner. In both studies, the increase/decrease had a similar time constant (summarized in Figure 5.3). These results are consistent with the neural membrane behaving approximately as a leaky integrator (Lapicque, 1907), as simulated and shown in Figure 5.4.

In Chapter 2 (same-polarity interactions, Figure 5.3A), there was an interaction between polarity and IPI, with the loudness of anodic pulses pairs decreasing more rapidly than that of cathodic stimulation. Both animal recordings and modelling

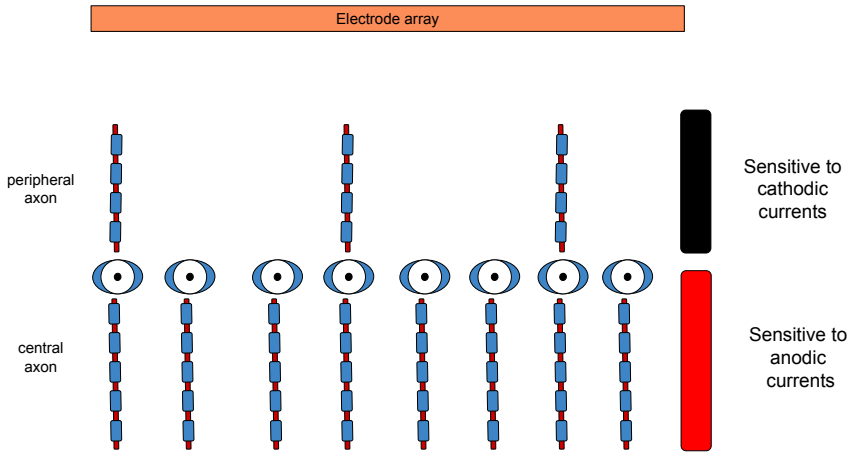

Figure 5.2: Schematics of the interaction between peripheral degeneration and polarity-specific stimulation of the neurons. At threshold, only a few neurons might be needed to perceive a sound, hence peripheral degeneration (5 peripheral axons have been removed in this drawing) might not be detrimental, even with cathodic stimulation. However, if a certain number of spikes is required (i.e. a certain loudness), peripheral degeneration might affect more cathodic stimulation than anodic stimulation.

suggest longer time constants of facilitation for nodes of Ranvier on the peripheral axons, which cathodic stimulation might preferentially activate. Hence, at the longest IPIs tested here, the second cathodic pulse might be facilitated, at least more than the second pulse in the anodic pulse pair. However, we suggested in Chapter 2 that the non-zero values at the longest IPIs could also be explained by a simple probabilistic model. Some neurons that did not spike for the first pulse might do so for the second pulse even without facilitation.

In Chapter 4 (opposite-polarity interactions), decreasing the IPI between the pulse pairs decreased the loudness. When both pulses had the same loudness in isolation, pulse pairs were louder than the single pulse in isolation (Figure 5.2B). Furthermore, the increase in loudness with IPI was slower for cathodic-first than for anodic-first pulses. The cathodic neural response is likely delayed compared to the anodic response because it stems from a more peripheral location (Miller et al., 1999b). By presenting anodic and then cathodic stimulation, the “neural

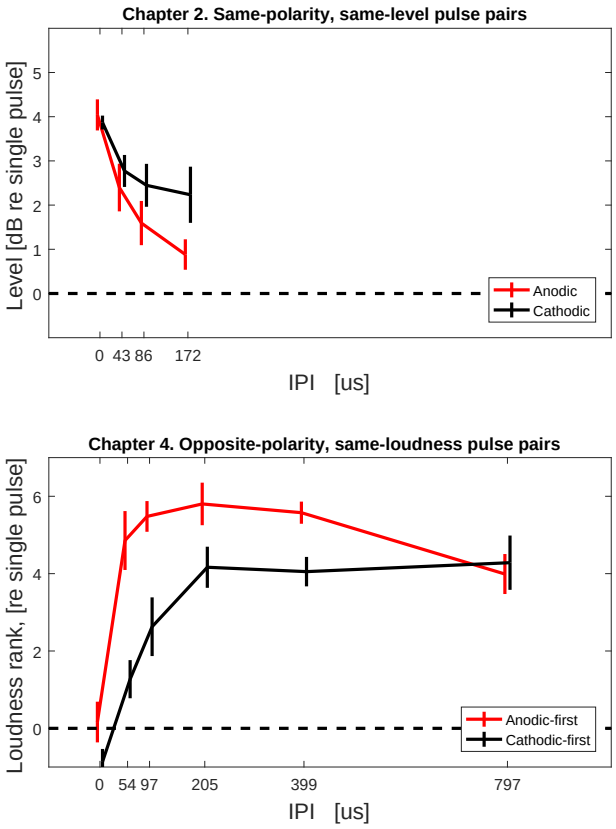

Figure 5.3: Results at MCL with the pulse pairs. **A.** Results of Chapter 2, with pulses having the same polarity. **B.** Results of Chapter 4, with pulses having an opposite polarity. Dashed line shows the loudness of the single pulse.

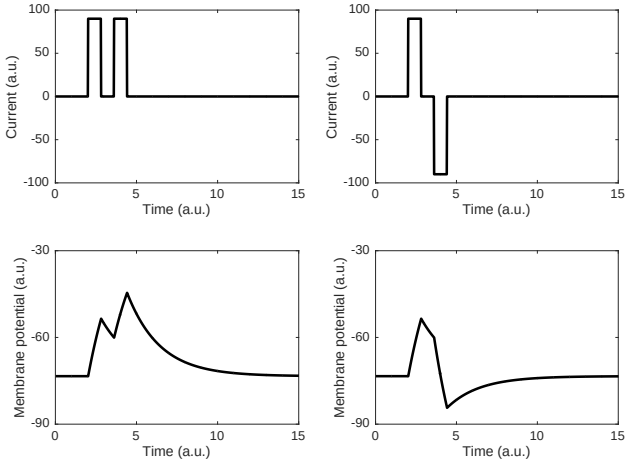

Figure 5.4: Simulated output given by a leaky integrator, when presented with pairs of pulses with the same (left) or opposite polarities (right). As the transmembrane potential goes up, the likelihood of reaching spiking threshold increases.

IPI" would therefore be longer than the actual IPI because of the delayed cathodic response (and vice-versa). This would explain the fast increase in loudness with IPI with anodic-first stimulation. However, alternative explanations such as charge cancellation or cathodal blocking at different nodes of Ranvier can not be excluded, and are discussed in Chapter 4.

Chapter 3 intended to get more insights in the overall latency difference between cathodic and anodic stimulation, by attempting its measure using ITDs. Results showed no effect of polarity on ITD-based localization at a group level (N = 5). For some subjects, the presentation of QPA in the right ear had to be delayed significantly to elicit the same perceived location as QPC. This indeed suggests a delayed neural response to QPC stimuli, compared to QPA stimuli. However, some subjects also showed the opposite effect, suggesting that the neural response to QPC leads that of QPA stimuli. No physiological study supports such behaviour. As discussed in Chapter 3, for some subjects, small differences in level (in the range of their loudness difference limens) changed the lateralization results dramatically.

The results might therefore reflect small errors in the loudness balancing between QPA and QPC, rather than a difference in the latency of the neural response.

Overall, all Chapters suggest an important role of the polarity of the stimulus on the interactions between pulses at short gaps. This is relevant for clinical contexts, as high pulse rates (approx. 6000 to 10000 pps when considering all channels) lead to similar IPIs as we tested. Thus, it is likely that changing the polarity of the pulses in a high-rate pulse train will have perceptual effects (e.g. on modulation detection thresholds, as suggested by Joshi et al., 2017).

5.2 Limitations of using asymmetric pulses

5.2.1 Main assumptions and corroborating results

Monophasic (i.e. pure cathodic or anodic) stimulation is unsafe for the auditory nerve (Brummer and Turner, 1977; Huang et al., 1999). However, different asymmetric shapes are achievable, depending on the CI brand. These asymmetric pulses are charge-balanced within a short window, but intend to give more weight to one specific polarity. In Chapters 2 and 4, with Advanced Bionics CIs, these were pseudo-monophasic (PS) pulses (pictograms, Figure 5.1). With such pulses, it is expected that the short-high phase dominates the neural response. This assumption relies on the neural membrane behaving as a leaky integrator. That is, a substantially larger amount of the charge will be leaked out for the long-low phase than for the short-high. This implies a careful choice of phase duration and asymmetry ratio in order to fit the membrane capacitive-resistive properties. Modelling and single-neuron recordings suggest that a ratio of 8 is enough to give more weight to the short-high phase (Frijns et al., 1996; Miller et al., 2001a). Furthermore, with Advanced Bionics devices, larger ratios would limit the minimum current step size achievable (cf. Appendix C). In Chapters 2 and 4, the ratio was 8, combined with a short-high phase duration of 43 μs.

Several studies used PS pulses in human CI listeners (Macherey et al., 2006, 2008, 2010, 2011; Undurraga et al., 2013; Wieringen et al., 2005). Undurraga et al. (2013) showed clear evidence that the short-high phase dominated the neural response

for both PSA and PSC pulses, as their eABR recordings stayed synchronized with the short-high phase when reverting their PS pulses. Our results in Chapter 2 support this finding at supra-threshold levels: changing the gap between the two short-high phases had a strong effect on the loudness for both anodic and cathodic short-high phases.

In Chapter 3, with Cochlear devices, we used quadraphasic (QP) pulses, assuming that the central phase dominated the neural response. There is indeed twice more charge in the central phase, compared to the flanking phases. Such QP pulses create similar polarity effects as PS pulses (Figure 5.1 and Carlyon et al., 2013).

Two main limitations arise from the use of asymmetric pulses. On one hand, the phases that are assumed not to contribute (long-low and flanking phases for PS and QP pulses, respectively) might still elicit some neural response. On the other hand, because of the neural membrane integrating the charge over a few hundred microseconds, those same phases will partially cancel the central or short-high phases. This will complicate the interpretation of the latter as equivalent to monophasic pulses. The next sections discuss those limitations, with or without the presence of an inter-phase gap.

5.2.2 Limitations when having no or short inter-phase gaps

In Chapters 3 (ITD) and 4 (opposite-polarities), there was no or only a short inter-phase gap between the hypothesized dominating phase (PS: short-high; QP: central phase) and the other ones (PS: long-low; QP: flanking phases).

In Chapter 4, the polarity order effects differed between experiments 2 and 3, when removing/adding the long-low phases of the PS pulses. Combined with the fact that the long-low pulses could be heard in isolation by the same CI listeners, this suggests a complex interplay between the short-high and long-low phases. Long-low phases might indeed mostly cancel the short-high phase, but also contribute to the overall response. Although we did not test for this, all the aforementioned interactions might very well depend on level, further complicating the interpretation.

Same interactions might also occur for QP pulses. Indeed, it is possible that

the flanking phases do contribute to the neural response at MCL. Since anodic currents are more efficient than cathodic currents at MCL, the neural response could be synchronized to the flanking phase of the QPC pulse, and to the central phase of the QPA pulse. The perception of the QPC pulse would then lead that of the QPA pulse, hence biasing the results of the ITD task.

It would also be interesting to study whether a contribution from the flanking phases explains the non-monotonic loudness growth reported by (Macherey et al., 2017). With QPC stimuli, the neural response could indeed exhibit a discrete change from the central to the flanking phases when increasing level.

5.2.3 Limitations when having a large inter-phase gap

In Chapter 2, there was a 2-ms inter-phase gap between the short-high and long-low phases of PS pulses.

Results at threshold showed no effect of IPI, in opposition to the results at MCL. On one hand, short-high phases might may have dominated the neural response but not interacted between each other. This could happen if each phase recruited different neurons for example. It is consistent with a study from Carlyon et al. (2005) that showed no effect of increasing the IPI between two phases of same polarity on detection thresholds. On the other hand, it could also be that the long-low phases did contribute significantly at thresholds. This is consistent with results from Macherey et al. (2006), that showed that both the long-low and short-high phases contributed to detection thresholds (with a gap of approx. 1 ms between both phases), but that MCLs were mostly driven by the short-high phases. This suggests overall that when having a long inter-phase gap between the short-high and long-low phases, there is a higher chance for each phase to contribute. This seems to be more critical at threshold, likely because of each polarity contributing more equally at the lowest end of the dynamic range (Undurraga et al., 2013).

5.2.4 Summary on the use of asymmetric pulses

Overall, our results are consistent with the assumption that the short-high phase (and central phase for the QP pulses) dominates the neural response, both for

anodic and cathodic versions of the pulses. However, even if the short-high phase (or central phase) dominates the response, the long-low phase (or flanking phases) might still play a role.

Having the long-low phase close to the short-high phase will lead to charge cancellation with pseudo-monophasic pulses. This might change the temporal interactions between pairs of short-high phases (cf. Chapter 4, experiments 2 and 3). Similar charge cancellation will happen between the central and flanking phases of the QP pulses.

Having a larger inter-phase gap reduces the respective cancellation between the short-high and the long-low phases. This, however, gives a higher chance for each phase to contribute to the neural response indpedently. Such contribution seems to be more critical at threshold than at MCL (Chapter 2 and Macherey et al., 2006).

Depending on the task (threshold vs MCL), one might therefore decide on using a different asymmetric pulses.

5.3 Applications

5.3.1 Characterization of the neural interface

The results of this thesis are applicable to the study of neural survival in CI listeners. Such neural survival might vary significantly across listeners and along the cochlea within each listener. This is supported by post-mortem SGN cell count (Hinojosa and Marion, 1983; Khan et al., 2005), and by the large variability of outcomes across electrodes for several psychophysical measurements (e.g., Bierer and Faulkner, 2010; DeVries et al., 2016; Long et al., 2014; Mesnildrey, 2017; Pfingst et al., 2015). Accordingly, a promising approach for improving speech intelligibility has been to deactivate or reprogram channels that are thought to be poor in delivering speech information (Zhou, 2017).

Various measures have been or are being proposed to detect such “bad” channels, including psychophysical measures (thresholds at low rates; modulation detection thresholds; difference between monopolar and tripolar thresholds; dif-

ference between polarity effects at thresholds) and physiological measures (effect of phase duration and inter-phase gap on the eCAP).

The paradigm of Chapter 2 (asymmetric pulses with short-high phases having the same polarity) seems well suited to study within-listener variations across electrodes. This chapter showed that at MCL, the results were consistent with the capacitive-resistive behaviour of the SGNs. A change in diameter or myelination across sites in the cochlea would impact such behaviour, and might therefore be captured by such paradigm.

Because the results at MCL were monotonic, one could reduce the number of IPIs to shorten the procedure (e.g. keep only 0 and 172-μs IPI). It might also be of interest to measure eCAPs with the paradigm of Chapter 2. If it correlates strongly with the psychophysical results, this would allow for comparison with animal models, where different types of ætiologies can be artificially induced. Similar attempts with eCAP measures have been done in animal studies by varying the inter-phase gap and phase duration of symmetric biphasic pulses (Prado-Guitierrez et al., 2006). The use of biphasic pulses however complicates the interpretation of their results (Ramekers et al., 2014). Using the paradigm of Chapter 2 would therefore largely ease the interpretation.

It is yet unclear which underlying mechanisms are responsible for the effects shown in Chapters 3 and 4. Therefore, more investigation would be needed before using such paradigms with the purpose of understanding across-electrode variability.

5.3.2 Modelling

The results presented in this thesis are appealing to model for several reasons. First, polarity effects were rather consistent across the subjects and electrodes tested here. Therefore, they likely reflect fundamental mechanisms of temporal integration. Furthermore, most of the paradigms were inspired by animal physiological recordings (Cartee et al., 2000; Cartee et al., 2006; Miller et al., 1999b). Since several models are based on those animal recordings, this offers a chance to bridge the gap between human psychophysics and animal physiology.

Recent phenomenological models may be able to account for most of the results presented in the thesis. For example, as shown in Chapter 2, both models from Joshi et al. (2017) and Macherey et al. (2007) predicted the discrepancy between the results at MCL and threshold. This is encouraging, as those models have a low computational complexity and rely on a small number of parameters. However, the cat model from Joshi et al. (2017) would need to be calibrated for human listeners, and the model from Macherey et al. (2007) would need to account for polarity effects.

Several factors related to the pattern of voltage along the SGNs have been hypothesized to play a role in the results of Chapters 2 to 4. Geometrical models calculate such voltage distribution (e.g., Frijns et al., 1996; Rattay et al., 2001b; Smit et al., 2008) and give relevant insights on the mechanisms of depolarization along each SGN. Furthermore, these models make it possible to predict the influence of the angle and distance between the electrode and neurons, peripheral axon survival, and effects of current spread along the cochlea. Finally, geometrical models allow the extraction of population characteristics of the neurons being stimulated electrically. As shown in Chapter 2, characteristics of such a population response might play an important role in the loudness summation of pulse pairs having the same polarity, even without facilitation. Geometrical models have the downside of relying on many more parameters.

Some of the differences seen between currents of opposite polarity could be explained by ion channel dynamics. For example, even a hyperpolarizing pulse can elicit a spike. This usually happens at the end of the hyperpolarizing pulse ("rebound" spike Kim and Holt, 2013, , likely because of HCN ion channels). Thus, cathodic and anodic currents might not only trigger spikes at locations on the SGNs where they depolarize the neuron. Only a physiological model coupled with a geometrical one would be able to give insights in such contributions.

Finally, a binaural model might give insights on the results of Chapter 3. Such a model would help disentangling the effects of ITD, ILD and their relation to loudness difference limens and dynamic range in each ear (Joshi, 2017). The last two factors might indeed play a larger role in lateralization tasks than the effects of polarity on the latency of the neural response.

5.4 General conclusions

This thesis investigated the temporal integration of electrical currents in CI listeners. This was done by presenting pulses pairs with short inter-pulse intervals (IPIs), while controlling for the polarity of each pulse. The main conclusions are as follows:

- Patterns at MCL and short IPIs were consistent with an underlying leaky integration of currents by the neurons. This is not the case at threshold, where integration of pulse pairs might be based on recruiting different neurons, rather than changing the operating point on the same neurons.
- When each pulse in the pulse pairs had the same loudness in isolation, the effect of IPI was polarity-dependent (both in Chapter 2 and 4). This is consistent with the hypothesis that each polarity activates different locations on the SGNs, with these locations having different capacitive-resistive properties and different latencies.
- The necessary use of charge-balanced asymmetric pulses can complicate the interpretation of the effects of IPI. For example, in Chapter 4, the order effects differed with the absence/presence of long-low phases. It is therefore recommended to push the long-low phases away when studying the interactions between short-high phases.
- There was overall no effect of polarity on ITD-based lateralization with the five subjects tested here. This however does not prove that the neural response has the same latency for currents with opposite polarities. Indeed, Chapter 3 suggests that such latency differences might have been shadowed by small ILD cues.
- The presented results are relevant for modelling, as they are consistent across subjects (particularly at MCL) and mimic paradigms from single-neuron studies. Furthermore, results from Chapter 2 at MCL are free from the influence of charge cancellations, largely simplifying the interpretation.

Bibliography

Adamo, N. J. and E. A. Daigneault (1973). “Ultrastructural features of neurons and nerve fibres in the spiral ganglia of cats”. In: *Journal of neurocytology* 2.1, pp. 91–103.

Balthasar, C. de, C. Boëx, G. Cosendai, G. Valentini, A. Sigrist, and M. Pelizzone (2003). “Channel interactions with high-rate biphasic electrical stimulation in cochlear implant subjects”. In: *Hearing Research* 182.1-2, pp. 77–87.

Bierer, J. A. (2010). “Probing the electrode-neuron interface with focused cochlear implant stimulation.” In: *Trends in amplification* 14.2, pp. 84–95.

Bierer, J. A. and K. F. Faulkner (2010). “Identifying cochlear implant channels with poor electrode-neuron interface: partial tripolar, single-channel thresholds and psychophysical tuning curves.” In: *Ear and hearing* 31.2, pp. 247–258.

Blamey, P. J., G. J. Dooley, C. J. James, and E. S. Parisi (2000). “Monaural and binaural loudness measures in cochlear implant users with contralateral residual hearing”. In: *Ear and Hearing* 21.1, pp. 6–17.

Blamey, P. et al. (2013). “Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 Patients”. In: *Audiology and Neurotology* 18.1, pp. 36–47.

Bostock, H., T. A. Sears, and R. M. Sherratt (1983). “Spatial distribution of excitability”. In: *The Journal of Physiology* 341, pp. 41–58.

Boulet, J., M. White, and I. C. Bruce (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.

Brännström, K. J. and J. Lantz (2010). “Interaural attenuation for Sennheiser HDA 200 circumaural earphones.” In: *International Journal of Audiology* 49.6, pp. 467–471.

Briaire, J. J. and J. H. M. Frijns (2006). “The consequences of neural degeneration regarding optimal cochlear implant position in scala tympani: A model approach”. In: *Hearing Research* 214.1-2, pp. 17–27.

Broadbent, D. E. and P. Ladefoged (1957). “On the fusion of sounds reaching different sense organs”. In: *The Journal of the Acoustical Society of America* 29.6, pp. 708–710.

Brown, M. C. (1994). “Antidromic responses of single units from the spiral ganglion.” In: *Journal of Neurophysiology* 71.5, pp. 1835–1847.

Brummer, S. B. and M. J. Turner (1977). “Electrochemical considerations for safe electrical stimulation of the nervous system with platinum electrodes”. In: *IEEE Transactions on Biomedical Engineering* BME-24.1, pp. 59–63.

Carlyon, R. P., A. Van Wieringen, J. M. Deeks, C. J. Long, J. Lyzenga, and J. Wouters (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.

Carlyon, R. P. et al. (2010). “Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear.” In: *Journal of the Association for Research in Otolaryngology* 11.4, pp. 625–40.

Carlyon, R. P., J. M. Deeks, and O. Macherey (2013). “Polarity effects on place pitch and loudness for three cochlear-implant designs and at different cochlear sites.” In: *The Journal of the Acoustical Society of America* 134.1, pp. 503–509.

Cartee, L. A. (2000). “Evaluation of a model of the cochlear neural membrane. II: Comparison of model and physiological measures of membrane properties measured in response to intrameatal electrical stimulation”. In: *Hearing research* 146.1, pp. 153–166.

Cartee, L. A. (2006). “Spiral ganglion cell site of excitation II: Numerical model analysis”. In: *Hearing Research* 215.1-2, pp. 22–30.

Cartee, L. A., C. van den Honert, C. C. Finley, and R. L. Miller (2000). “Evaluation of a model of the cochlear neural membrane. I. Physiological measurement of

membrane characteristics in response to intrameatal electrical stimulation". In: *Hearing research* 146.1, pp. 143–152.

Cartee, L. A., C. A. Miller, and C. van den Honert (2006). "Spiral ganglion cell site of excitation I: comparison of scala tympani and intrameatal electrode responses." In: *Hearing research* 215.1-2, pp. 10–21.

Chang, S.-A. et al. (2010). "Performance over time on adults with simultaneous bilateral cochlear implants". In: *Journal of the American Academy of Audiology* 21.1, pp. 35–43.

Ching, T. Y. C., E. van Wanrooy, and H. Dillon (2007). "Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: a review." In: *Trends in amplification* 11.3, pp. 161–92.

Colburn, H. S., Y. Chung, Y. Zhou, and A. Brughera (2008). "Models of brainstem responses to bilateral electrical stimulation". In: *Journal of the Association for Research in Otolaryngology* 10.1, pp. 91–110.

Colombo, J. and C. W. Parkins (1987). "A model of electrical excitation of the mammalian auditory-nerve neuron". In: *Hearing Research* 31.3, pp. 287–311.

Cosentino, S., J. M. Deeks, and R. P. Carlyon (2015). "Procedural Factors That Affect Psychophysical Measures of Spatial Selectivity in Cochlear Implant Users". In: *Trends in Hearing* 19.0, pp. 1–16.

Cosentino, S., O. Stakhovskaya, G. Bernstein, and M. Goupell (2017). "Optimizing pitch matching for bilateral cochlear implant users". In: *Conference on Implantable Auditory Prostheses*. Lake Tahoe, CA, USA.

Cutting, J. E. (1976). "Auditory and linguistic processes in speech perception: inferences from six fusions in dichotic listening." In: *Psychological Review* 83.2, pp. 114–140.

David, E. E., N. Guttman, and W. A. van Bergeijk (1959). "Binaural Interaction of High-Frequency Complex Stimuli". In: *The Journal of the Acoustical Society of America* 31.6, pp. 774–782.

DeVries, L., R. Scheperle, and J. A. Bierer (2016). "Assessing the Electrode-Neuron Interface with the Electrically Evoked Compound Action Potential, Electrode Position, and Behavioral Thresholds". In: *Journal of the Association for Research in Otolaryngology* 17.3, pp. 237–252.

Dobie, R. A. and M. J. Wilson (1996). “A comparison of t test, F test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids.” In: *The Journal of the Acoustical Society of America* 100.4 Pt 1, pp. 2236–2246.

Dorman, M. F. and R. H. Gifford (2010). “Combining acoustic and electric stimulation in the service of speech recognition.” In: *International Journal of Audiology* 49.12, pp. 912–919.

Dynes, S. B. and B. Delgutte (1992). “Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea”. In: *Hearing Research* 58.1, pp. 79–90.

Egger, K., P. Majdak, and B. Laback (2016). “Channel Interaction and Current Level Affect Across-Electrode Integration of Interaural Time Differences in Bilateral Cochlear-Implant Listeners”. In: *Journal of the Association for Research in Otolaryngology* 17.1, pp. 55–67.

Elberling, C., M. Don, M. Cebulla, and E. Stürzebecher (2007). “Auditory steady-state responses to chirp stimuli based on cochlear traveling wave delay.” In: *The Journal of the Acoustical Society of America* 122.5, pp. 2772–2785.

Field, A., J. Miles, and Z. Field (2012). *Discovering statistics using R*. Sage, Thousand Oaks.

Florentine, M. (2011). *Loudness*. Ed. by R. R. Fay and A. N. Popper. Springer.

Francart, T. (2011). “Sensitivity of Bimodal Listeners to Interaural Time Differences with Modulated Single- and Multiple-Channel Stimuli”. In: *Audiology & neuro-otology* 16.2, pp. 82 –92.

Francart, T. and H. J. McDermott (2013). “Psychophysics, fitting, and signal processing for combined hearing aid and cochlear implant stimulation.” In: *Ear and hearing* 34.6, pp. 685–700.

Francart, T., J. Brokx, and J. Wouters (2009). “Sensitivity to interaural time differences with combined cochlear implant and acoustic stimulation.” In: *Journal of the Association for Research in Otolaryngology* 10.1, pp. 131–141.

Friesen, L. M., R. V. Shannon, D. Baskent, and X. Wang (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.

Frijns, J. H. M., S. L. De Snoo, and J. H. Ten Kate (1996). "Spatial selectivity in a rotationally symmetric model of the electrically stimulated cochlea". In: *Hearing Research* 95.1-2, pp. 33–48.

Galvin, J. J., S. I. Oba, D. Başkent, M. Chatterjee, and Q.-J. Fu (2015). "Envelope Interactions in Multi-Channel Amplitude Modulation Frequency Discrimination by Cochlear Implant Users". In: *Plos One* 10.10, e0139546.

Guérit, F. (2018). "PPS Toolbox". In:

Guérit, F., S. Santurette, J. Chalupper, and T. Dau (2014). "Investigating Interaural Frequency-Place Mismatches via Bimodal Vowel Integration." In: *Trends in hearing* 18.

Guérit, F., J. Marozeau, and B. Epp (2017). "Linear combination of auditory steady-state responses evoked by co-modulated tones". In: *The Journal of the Acoustical Society of America* 142.4, EL395–400.

Guérit, F., J. Marozeau, J. M. Deeks, B. Epp, and R. P. Carlyon (2018). "Effects of the relative timing of opposite-polarity pulses on loudness for cochlear implant listeners". In: *The Journal of the Acoustical Society of America* 144.5, pp. 2751–2763.

Harnsberger, J. D., M. A. Svirsky, A. R. Kaiser, D. B. Pisoni, R. Wright, and T. A. Meyer (2001). "Perceptual "vowel spaces" of cochlear implant users: Implications for the study of auditory adaptation to spectral shift". In: *The Journal of the Acoustical Society of America* 109.5, p. 2135.

Hartmann, R., G. Topp, and R. Klinke (1984). "Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea". In: *Hearing Research* 13.1, pp. 47–62.

He, S., C. J. Brown, and P. J. Abbas (2010). "Effects of stimulation level and electrode pairing on the binaural interaction component of the electrically evoked auditory brain stem response." In: *Ear and hearing* 31.4, pp. 457–470.

He, S., C. J. Brown, and P. J. Abbas (2012). "Binaural interaction component of the electrically evoked auditory brainstem response and interaural pitch com-

parisons in bilateral cochlear implant recipients". In: *Ear and hearing* 33.1, pp. 57–68.

Henning, G. B. (1974). "Detectability of interaural delay in high-frequency complex waveforms." In: *The Journal of the Acoustical Society of America* 55.1, pp. 84–90.

Herdman, A. T., O. Lins, P. Van Roon, D. R. Stapells, M. Scherg, and T. W. Picton (2002). "Intracerebral sources of human auditory steady-state responses". In: *Brain Topography* 15.2, pp. 69–86.

Hinojosa, R. and M. Marion (1983). "Histopathology of profound sensorineural deafness". In: *Annals of the New York Academy of Sciences* 405.1, pp. 459–484.

Hoesel, R. J. M. van (2004). "Exploring the benefits of bilateral cochlear implants". In: *Audiology and Neuro-Otology* 9.4, pp. 234–246.

Hoesel, R. J. M. van (2007). "Sensitivity to binaural timing in bilateral cochlear implant users". In: *The Journal of the Acoustical Society of America* 121.4, pp. 2192–2206.

Holube, I. (2012). "Speech intelligibility in fluctuating maskers". In: *Proceedings of ISAAR 2011: Speech Perception and Auditory Disorders. 3rd International Symposium on Auditory and Audiological Research.* Ed. by T. Dau, M. L. Jepsen, T. Poulsen, and J. C. Dalsgaard. Nyborg, Denmark: The Danavox Jubilee Foundation, pp. 57–64.

Honert, C. van den and P. H. Stypulkowski (1984). "Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings". In: *Hearing research* 14.3, pp. 225–243.

Honert, C. van den and J. T. Mortimer (1979). "The Response of the Myelinated Nerve Fiber To Short Duration Biphasic Stimulating Currents". In: *Annals of biomedical engineering* 7, pp. 117–125.

Huang, C. Q., R. K. Shepherd, P. M. Carter, P. M. Seligman, and B. Tabor (1999). "Electrical stimulation of the auditory nerve: direct current measurement in vivo". In: *IEEE Transactions on Biomedical Engineering* 46.4, pp. 461–470.

Ihlefeld, A., B. G. Shinn-Cunningham, and R. P. Carlyon (2012). "Comodulation masking release in speech identification with real and simulated cochlear-implant hearing." In: *The Journal of the Acoustical Society of America* 131.2, pp. 1315–1324.

Ihlefeld, A., R. P. Carlyon, A. Kan, T. H. Churchill, and R. Y. Litovsky (2015). "Limitations on Monaural and Binaural Temporal Processing in Bilateral Cochlear Implant Listeners". In: *Journal of the Association for Research in Otolaryngology* 16, pp. 641–652.

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

James, C., P. Blamey, J. K. Shallop, P. V. Incerti, and A. M. Nicholas (2001). "Contralateral masking in cochlear implant users with residual hearing in the non-implanted ear." In: *Audiology & neuro-otology* 6.2, pp. 87–97.

Janssen, N. A., L. Bramsløw, S. Riis, and J. Marozeau (2017). "The scale illusion detection task: Objective assessment of binaural fusion in normal-hearing listeners". In: *Proceedings of the International Symposium on Auditory and Audiological Research*. Ed. by S. Santurette, T. Dau, J. C. Dalsgaard, L. Tranebjærg, T. Andersen, and T. Poulsen. Nyborg, Denmark: The Danavox Jubilee Foundation, pp. 207–213.

Javel, E and R. K. Shepherd (2000). "Electrical stimulation of the auditory nerve. III. Response initiation sites and temporal fine structure". In: *Hearing research* 140.1-2, pp. 45–76.

John, M. S., O. G. Lins, B. L. Boucher, and T. W. Picton (1998). "Multiple auditory steady-state responses (MASTER): stimulus and recording parameters". In: *Audiology* 37.2, pp. 59–82.

John, M. S., A. Dimitrijevic, and T. W. Picton (2001). "Weighted averaging of steady-state responses". In: *Clinical neurophysiology* 112.3, pp. 555–562.

John, M. S., A Dimitrijevic, and T. W. Picton (2003). "Efficient stimuli for evoking auditory steady-state responses." In: *Ear and hearing* 24.5, pp. 406–423.

Joshi, S. N. (2017). "Modeling auditory nerve responses to electrical stimulation". PhD thesis. Technical University of Denmark.

Joshi, S. N., T. Dau, and B. Epp (2017). "A Model of Electrically Stimulated Auditory Nerve Fiber Responses with Peripheral and Central Sites of Spike Generation". In: *Journal of the Association for Research in Otolaryngology* 18.2, pp. 323–342.

Joucla, S. and B. Yvert (2012). “Modeling extracellular electrical neural stimulation: From basic understanding to MEA-based applications”. In: *Journal of Physiology-Paris* 106.3-4, pp. 146–158.

Kan, A., C. Stoelb, R. Y. Litovsky, and M. J. Goupell (2013). “Effect of mismatched place-of-stimulation on binaural fusion and lateralization in bilateral cochlear-implant users.” In: *The Journal of the Acoustical Society of America* 134.4, pp. 2923–36.

Kan, A., R. Y. Litovsky, and M. J. Goupell (2015). “Effects of interaural pitch-matching and auditory image centering on binaural sensitivity in cochlear-implant users”. In: *Ear and Hearing* 36.3, e62–e68.

Karg, S. A., C. Lackner, and W. Hemmert (2013). “Temporal interaction in electrical hearing elucidates auditory nerve dynamics in humans.” In: *Hearing research* 299, pp. 10–18.

Kerber, S. and B. U. Seeber (2012). “Sound localization in noise by normal-hearing listeners and cochlear implant users”. In: *Ear and Hearing* 33.4, pp. 445–457.

Khan, A. M., O. Handzel, D. K. Eddington, D. Damian, and J. B. Nadol (2005). “Effect of cochlear implantation on residual spiral ganglion cell count as determined by comparison with the contralateral nonimplanted inner ear in humans”. In: *Annals of Otology, Rhinology and Laryngology* 114.5, pp. 381–385.

Kim, K. X. and M. A. Rutherford (2016). “Maturation of NaV and KV Channel Topographies in the Auditory Nerve Spike Initiator before and after Developmental Onset of Hearing Function”. In: *Journal of Neuroscience* 36.7, pp. 2111–2118.

Kim, Y.-H. and J. R. Holt (2013). “Functional contributions of HCN channels in the primary auditory neurons of the mouse inner ear”. In: *The Journal of General Physiology* 142.3, pp. 207–223.

Klatt, D. H. (1980). “Software for a cascade/parallel formant synthesizer”. In: *Journal of the Acoustical Society of America* 67.3, pp. 971–995.

Koles, Z. J. and M. Rasminsky (1972). “A computer simulation of conduction in demyelinated nerve fibres”. In: *The Journal of Physiology* 227.2, pp. 351–364.

Kordus, M. and J. Żera (2017). “Loudness Functions and Binaural Loudness Summation in Bilateral Cochlear Implant Users”. In: *Archives of Acoustics* 42.3, pp. 351–364.

Krumbholz, K, R. D. Patterson, and D Pressnitzer (2000). “The lower limit of pitch as determined by rate discrimination.” In: *The Journal of the Acoustical Society of America* 108.3 Pt 1, pp. 1170–1180.

Kwon, B. J. and C. van den Honert (2009). “Spatial and temporal effects of interleaved masking in cochlear implants.” In: *Journal of the Association for Research in Otolaryngology* 10.3, pp. 447–457.

Laback, B., K. Egger, and P. Majdak (2015). “Perception and coding of interaural time differences with bilateral cochlear implants”. In: *Hearing Research* 322, pp. 138–150.

Lapicque, L. (1907). “Recherches quantitatives sur l’excitation électrique des nerfs traitée comme une polarisation”. In: *Journal de Physiologie et de Pathologie Générale* 9.1, pp. 620–635.

Leake, P. A. and G. T. Hradek (1988). “Cochlear pathology of long term neomycin induced deafness in cats”. In: *Hearing Research* 33.1, pp. 11–33.

Levitt, H. (1971). “Transformed Up-Down Methods in Psychoacoustics”. en. In: *The Journal of the Acoustical Society of America* 49.2 Pt 2, pp. 467–477.

Liberman, M. C. (1984). “Single-neuron labeling and chronic cochlear pathology. I. Threshold shift and characteristic-frequency shift”. In: *Hearing Research* 16.1, pp. 33–41.

Liberman, M. C. and M. E. Oliver (1984). “Morphometry of intracellularly labeled neurons of the auditory nerve: Correlations with functional properties”. In: *Journal of Comparative Neurology* 223.2, pp. 163–176.

Lilly, J. C., J. R. Hughes, E. C. Alvord Jr, and T. W. Galkin (1955). “Brief, noninjurious electric waveform for stimulation of the brain.” In: *Science* 121.3144, pp. 468–469.

Lin, P., T. Lu, and F.-G. Zeng (2013). “Central masking with bilateral cochlear implants.” In: *The Journal of the Acoustical Society of America* 133.2, pp. 962–969.

Litovsky, R. Y., G. L. Jones, S. Agrawal, and R. van Hoesel (2010). “Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans”. In: *The Journal of the Acoustical Society of America* 127.1, pp. 400–414.

Litovsky, R. Y., M. J. Goupell, A. Kan, and D. M. Landsberger (2017). "Use of Research Interfaces for Psychophysical Studies With Cochlear-Implant Users". In: *Trends in Hearing* 21, p. 233121651773646.

Litvak, L. M., A. J. Spahr, A. A. Saoji, and G. Y. Fridman (2007). "Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners." In: *The Journal of the Acoustical Society of America* 122.2, pp. 982–91.

Loizou, P. C. et al. (2009). "Speech recognition by bilateral cochlear implant users in a cocktail-party setting". In: *The Journal of the Acoustical Society of America* 125.1, pp. 372–383.

Long, C. J., D. K. Eddington, H. S. Colburn, and W. M. Rabinowitz (2003). "Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user". In: *The Journal of the Acoustical Society of America* 114.3, pp. 1565–1574.

Long, C. J. et al. (2005). "Optimizing the clinical fit of auditory brain stem implants". In: *Ear and hearing* 26.3, pp. 251–262.

Long, C. J. et al. (2014). "Examining the electro-neural interface of cochlear implant users using psychophysics, CT scans, and speech understanding". In: *JARO - Journal of the Association for Research in Otolaryngology* 15.2, pp. 293–304.

Macherey, O. and Y. Cazals (2016). "Effects of Pulse Shape and Polarity on Sensitivity to Cochlear Implant Stimulation: A Chronic Study in Guinea Pigs". In: *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*. Ed. by P. van Dijk, D. Başkent, E. Gaudrain, E. de Kleine, A. Wagner, and C. Lanting. Springer, pp. 133–142.

Macherey, O., A. van Wieringen, R. P. Carlyon, J. M. Deeks, and J. Wouters (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.

Macherey, O., R. P. Carlyon, A. van Wieringen, and J. Wouters (2007). "A dual-process integrator-resonator model of the electrically stimulated human auditory nerve." In: *Journal of the Association for Research in Otolaryngology* 8.1, pp. 84–104.

Macherey, O., R. P. Carlyon, A. van Wieringen, J. M. Deeks, and J. Wouters (2008). "Higher sensitivity of human auditory nerve fibers to positive electrical currents". In: *Journal of the Association for Research in Otolaryngology* 9, pp. 241–251.

Macherey, O., A. van Wieringen, R. P. Carlyon, I. Dhooge, and J. Wouters (2010). "Forward-masking patterns produced by symmetric and asymmetric pulse shapes in electric hearing." In: *The Journal of the Acoustical Society of America* 127.1, pp. 326–338.

Macherey, O., J. M. Deeks, and R. P. Carlyon (2011). "Extending the limits of place and temporal pitch perception in cochlear implant users." In: *Journal of the Association for Research in Otolaryngology* 12.2, pp. 233–251.

Macherey, O., R. P. Carlyon, J. Chatron, and S. Roman (2017). "Effect of pulse polarity on thresholds and on non-monotonic loudness growth in cochlear implant users". In: *Journal of the Association for Research in Otolaryngology* 18, pp. 513–527.

McDermott, H., C. Sucher, and A. Simpson (2009). "Electro-acoustic stimulation. Acoustic and electric pitch comparisons." In: *Audiology & neuro-otology* 14.suppl 1, pp. 2–7.

McKay, C. M. and H. J. McDermott (1998). "Loudness perception with pulsatile electrical stimulation: the effect of interpulse intervals." In: *The Journal of the Acoustical Society of America* 104.2 Pt 1, pp. 1061–1074.

McKay, C. M. and K. R. Henshall (2003). "The perceptual effects of interphase gap duration in cochlear implant stimulation". In: *Hearing Research* 181.1-2, pp. 94–99.

Mesnildrey, Q. (2017). "Towards a Better Understanding of the Cochlear Implant - Auditory Nerve Interface: from intracochlear electrical recordings to psychophysics." PhD thesis. Aix-Marseille Université.

Middlebrooks, J. 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.

Miller, C. A., P. J. Abbas, and J. T. Rubinstein (1999a). “An empirically based model of the electrically evoked compound action potential”. In: *Hearing Research* 135.1-2, pp. 1–18.

Miller, C. A., B. K. Robinson, J. T. Rubinstein, P. J. Abbas, and C. L. Runge-Samuelson (2001a). “Auditory nerve responses to monophasic and biphasic electric stimuli.” In: *Hearing research* 151.1-2, pp. 79–94.

Miller, C. A., M. J. Faulkner, and B. E. Pfingst (1995). “Functional responses from guinea pigs with cochlear implants II. Changes in electrophysiological and psychophysical measures over time”. In: *Hearing Research* 92.1-2, pp. 100–111.

Miller, C. A., P. J. Abbas, B. K. Robinson, J. T. Rubinstein, and A. J. Matsuoka (1999b). “Electrically evoked single-fiber action potentials from cat: responses to monopolar, monophasic stimulation.” In: *Hearing research* 130.1-2, pp. 197–218.

Miller, C. A., P. J. Abbas, and B. K. Robinson (2001b). “Response properties of the refractory auditory nerve fiber”. In: *Journal of the Association for Research in Otolaryngology* 2.3, pp. 216–232.

Nadol, J. B. (1997). “Patterns of neural degeneration in the human cochlea and auditory nerve: Implications for cochlear implantation”. In: *Otolaryngology - Head and Neck Surgery* 117.3, pp. 220–228.

Nelson, P. B., S.-H. Jin, A. E. Carney, and D. A. Nelson (2003). “Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners”. In: *The Journal of the Acoustical Society of America* 113.2, pp. 961–968.

Oxenham, A. J. and H. A. Kreft (2014). “Speech Perception in Tones and Noise via Cochlear Implants Reveals Influence of Spectral Resolution on Temporal Processing.” In: *Trends in hearing* 18.

Pfingst, B. E. et al. (2015). “Importance of cochlear health for implant function”. In: *Hearing Research* 322, pp. 77–88.

Picton, T. W., M. S. John, A. Dimitrijevic, and D. Purcell (2003). “Human auditory steady-state responses”. In: *International Journal of Audiology* 42.4, pp. 177–219.

Picton, T. W., P van Roon, and M. S. John (2007). “Human auditory steady-state responses during sweeps of intensity.” In: *Ear and Hearing* 28, pp. 542–557.

Plack, C. J. (2013). *The sense of hearing*. Psychology Press.

Poelmans, H., H. Luts, M. Vandermosten, P. Ghesquière, and J. Wouters (2012). "Hemispheric asymmetry of auditory steady-state responses to monaural and diotic stimulation". In: *Journal of the Association for Research in Otolaryngology* 13.6, pp. 867–876.

Poon, B. B., D. K. Eddington, V. Noel, and H. S. Colburn (2009). "Sensitivity to interaural time difference with bilateral cochlear implants: Development over time and effect of interaural electrode spacing". In: *The Journal of the Acoustical Society of America* 126.2, pp. 806–815.

Prado-Guitierrez, P., L. M. Fewster, J. M. Heasman, C. M. McKay, and R. K. Shepherd (2006). "Effect of interphase gap and pulse duration on electrically evoked potentials is correlated with auditory nerve survival." In: *Hearing research* 215.1-2, pp. 47–55.

Ramekers, D., H. Versnel, S. B. Strahl, E. M. Smeets, S. F. L. Klis, and W. Grolman (2014). "Auditory-nerve responses to varied inter-phase gap and phase duration of the electric pulse stimulus as predictors for neuronal degeneration". In: *Journal of the Association for Research in Otolaryngology* 15.2, pp. 187–202.

Ranck, J. B. J. (1975). "Which elements are excited in electrical stimulation of mammalian central nervous system: A review". In: *Brain Research* 98.3, pp. 417–440.

Rattay, F. (1999). "The basic mechanism for the electrical stimulation of the nervous system". In: *Neuroscience* 89.2, pp. 335–346.

Rattay, F., P. Lutter, and H. Felix (2001a). "A model of the electrically excited human cochlear neuron. I. Contribution of neural substructures to the generation and propagation of spikes". In: *Hearing Research* 153.1-2, pp. 43–63.

Rattay, F., R. N. Leao, and H. Felix (2001b). "A model of the electrically excited human cochlear neuron. II. Influence of the three-dimensional cochlear structure on neural excitability". In: *Hearing Research* 153.1-2, pp. 64–79.

Reiss, L., C. W. Turner, S. R. Erenberg, and B. J. Gantz (2007). "Changes in pitch with a cochlear implant over time." In: *Journal of the Association for Research in Otolaryngology* 8.2, pp. 241–57.

Reiss, L., R. Ito, J. L. Eggleston, and D. R. Wozny (2014). "Abnormal binaural spectral integration in cochlear implant users." In: *Journal of the Association for Research in Otolaryngology* 15.2, pp. 235–48.

Resnick, J. M., G. O'Brien, and J. T. Rubinstein (2018). "Simulated auditory nerve axon demyelination alters sensitivity and response timing to extracellular stimulation". In: *Hearing Research* 361, pp. 121–137.

Riecke, L., E. Formisano, C. S. Herrmann, and A. T. Sack (2015). "4-Hz transcranial alternating current stimulation phase modulates hearing". In: *Brain Stimulation* 8.4, pp. 777–783.

Robertson, D. (1976). "Possible relation between structure and spike shapes of neurones in guinea pig cochlear ganglion". In: *Brain research* 109.3, pp. 487–496.

Rosen, S., A. Faulkner, and L. Wilkinson (1999). "Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants." In: *The Journal of the Acoustical Society of America* 106.6, pp. 3629–36.

Scharf, B., M. Florentine, and C. H. Meiselman (1976). "Critical band in auditory lateralization." In: *Sensory processes* 1.2, pp. 109–126.

Shannon, R. V. (1992). "A model of safe levels for electrical stimulation". In: *IEEE Transactions on biomedical engineering* 39.4, pp. 424–426.

Siciliano, C. M., A. Faulkner, S. Rosen, and K. Mair (2010). "Resistance to learning binaurally mismatched frequency-to-place maps: implications for bilateral stimulation with cochlear implants." In: *The Journal of the Acoustical Society of America* 127.3, pp. 1645–60.

Skinner, M. W. et al. (2002). "CT-derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus-22 recipients." In: *Journal of the Association for Research in Otolaryngology* 3, pp. 332–350.

Smit, J. E., T. Hanekom, and J. J. Hanekom (2008). "Predicting action potential characteristics of human auditory nerve fibres through modification of the Hodgkin-Huxley equations". In: *South African Journal of Science* 104.7-8, pp. 284–292.

Smit, J. E., T. Hanekom, and J. J. Hanekom (2009). "Estimation of stimulus attenuation in cochlear implants". In: *Journal of Neuroscience Methods* 180.2, pp. 363–373.

Stickney, G. S., F.-G. Zeng, R. Litovsky, and P. Assmann (2004). "Cochlear implant speech recognition with speech maskers". In: *The Journal of the Acoustical Society of America* 116.2, pp. 1081–1091.

Strange, W., O.-S. Bohn, S. a. Trent, and K. Nishi (2004). "Acoustic and perceptual similarity of North German and American English vowels". In: *The Journal of the Acoustical Society of America* 115.4, pp. 1791–1807.

Svirsky, M. A., A. Silveira, H. Neuburger, S. W. Teoh, and H. Suarez (2004). "Long-term auditory adaptation to a modified peripheral frequency map". In: *Acta Oto-laryngologica* 124.4, pp. 381–386.

Takanen, M., T. Raitio, O. Santala, P. Alku, and V. Pulkki (2013). "Fusion of spatially separated vowel formant cues". In: *The Journal of the Acoustical Society of America* 134.6, p. 4508.

Undurraga, J. A., R. P. Carlyon, J. Wouters, and A. van Wieringen (2013). "The polarity sensitivity of the electrically stimulated human auditory nerve measured at the level of the brainstem." In: *Journal of the Association for Research in Otolaryngology* 14.3, pp. 359–377.

Vermeire, K. and P. Van de Heyning (2009). "Binaural hearing after cochlear implantation in subjects with unilateral sensorineural deafness and tinnitus." In: *Audiology & neuro-otology* 14.3, pp. 163–71.

Viemeister, N. F. and G. H. Wakefield (1991). "Temporal integration and multiple looks." In: *The Journal of the Acoustical Society of America* 90.2 Pt 1, pp. 858–865.

Wieringen, A. van, R. P. Carlyon, J. Laneau, and J. Wouters (2005). "Effects of waveform shape on human sensitivity to electrical stimulation of the inner ear". In: *Hearing Research* 200.1-2, pp. 73–86.

Wieringen, A. van, O. Macherey, R. P. Carlyon, J. M. Deeks, and J. Wouters (2008). "Alternative pulse shapes in electrical hearing." In: *Hearing research* 242.1-2, pp. 154–163.

Wilson, B. S., C. C. Finley, D. T. Lawson, R. D. Wolford, D. K. Eddington, and W. M. Rabinowitz (1991). “Better speech recognition with cochlear implants”. In: *Nature* 352.6332, pp. 236–238.

Won, J. H., W. R. Drennan, and J. T. Rubinstein (2007). “Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users”. In: *Journal of the Association for Research in Otolaryngology* 8.3, pp. 384–392.

Zhou, N. (2017). “Deactivating stimulation sites based on low-rate thresholds improves spectral ripple and speech reception thresholds in cochlear implant users”. In: *The Journal of the Acoustical Society of America* 141.3, EL243–EL248.

A

Investigating Interaural Frequency-Place Mismatches via Bimodal Vowel Integration[a]

Abstract

For patients having residual hearing in one ear and a cochlear implant (CI) in the opposite ear, interaural place-pitch mismatches might be partly responsible for the large variability in individual benefit. Behavioral pitch matching between the two ears has been suggested as a way to individualize the fitting of the frequency-to-electrode map, but is rather tedious and unreliable. Here, an alternative method using two-formant vowels was developed and tested. The interaural spectral shift was inferred by comparing vowel spaces, measured by presenting the first formant (F1) to the non-implanted ear and the second (F2) on either side. The method was first evaluated with 8 normal-hearing (NH) listeners and vocoder simulations, before being tested with 11 CI users. Average vowel distributions across subjects showed a similar pattern when presenting F2 on either side, suggesting acclimatization to the frequency map. However, individual vowel spaces with F2 presented to the implant did not allow a reliable estimation of the interaural mismatch. These results suggest that interaural frequency-place mismatches can be derived from such vowel spaces. However, the method remains limited by difficulties in bimodal fusion of the two

[a] This chapter is based on Guérit et al. (2014).

formants.

A.1 Introduction

In recent years, an increasing number of patients with residual contralateral hearing have received a cochlear implant (CI). This population is therefore combining the neural excitation coming from the CI with that from the ear stimulated acoustically. This has been shown to improve speech perception in noise, an effect likely to come from better access to the low-frequency content of the speech in the ear with preserved acoustic hearing (Dorman and Gifford, 2010). However, the extent to which patients benefit from the combination of electric and acoustic stimulation is highly variable, with some cases of interference between the modes of stimulation (for a review, see Ching et al., 2007). Several factors have been suggested to explain this variability, such as differences in the amount of residual hearing, the devices used, and their fitting. In particular, due to the variability in electrode placement in the cochlea and in cochlear duct length among patients, it is difficult to activate nerve fibers with the same frequency-to-place map as in the contralateral ear. Typically, a standard frequency-to-electrode allocation is used across subjects for the clinical fitting, assuming that the brain can adapt to a mismatch. The evolution of speech perception over time after implantation supports the theory of accommodation to a frequency shift (e.g. Skinner et al., 2002). However, a complete adaptation might not be possible in the case of large mismatches. Rosen et al. (1999) showed that, even after a long-term training period with a vocoder system simulating a 6.5 mm basalwards shift, speech recognition was worse than for the unshifted condition. Also in NH listeners, Siciliano et al. (2010) used a 6-channel vocoder and presented odd channels in the right ear, shifted 6 mm basally, while keeping the even channels unshifted in the left ear. After 10 hours of training, subjects showed poorer speech perception in this condition than when presented with the three unshifted channels only, suggesting that they did not benefit from combining the mismatched maps. More recently, and based on bilateral CI users' data, Kan et al. (2013) suggested that the salience of interaural time and level differences

was hampered for shifts greater than 3 mm. These binaural cues are essential for auditory scene analysis, and a decreased salience would imply difficulties, e.g. in the segregation of speech from several noise sources.

The above findings suggest that the electrode-array location is important for adequate fitting of, and optimal benefit from, the CI. Although electrode location can theoretically be determined from computer tomography (CT) scans, these are often unavailable in audiological practice and require an additional dose of radiation. For patients having residual hearing in the opposite ear, behavioral pitch-matching has been suggested but is rather difficult because of the different percepts elicited by the implant and the acoustic stimulation. Carlyon et al. (2010) also showed that results for behavioral pitch-matching experiments are strongly influenced by nonsensory biases and that the method is tedious and time-consuming. Other behavioral methods have been suggested, such as a contralateral masking paradigm (James et al., 2001; Lin et al., 2013) or an interaural time difference detection task (Francart, 2011; Francart et al., 2009), but the results from these two methods are not very precise and are also very time consuming to obtain. More recently, the use of the binaural interaction component of the auditory brainstem response has been proposed, based on data from cats (He et al., 2010). However, the preliminary results showed no significant correlation between the amplitude of the component and interaural pitch comparisons in humans (He et al., 2012).

In the present study, based on the ability to fuse vowel formants across ears (Broadbent and Ladefoged, 1957; Cutting, 1976), an alternative method using synthesized two-formant vowels was developed and tested. This method is potentially clinic-friendly, using stimuli that are similar to those CI users deal with in their everyday lives. The question addressed was the following: Can the second formant (F2) of a two-formant vowel be used as an indicator of interaural frequency-to-place mismatch by presenting it either to the aided/normal-hearing side or to the implanted side? If the implant is perfectly fitted, the perceived vowel distributions should not depend on the ear to which F2 is presented, when fixing the first formant (F1) on the acoustic side. In the presence of an interaural mismatch, vowel distributions should show differences when presenting F2 to the acoustic vs the electric side. To test this hypothesis, an experiment with normal-hearing (NH)

listeners using a vocoder system and simulated interaural mismatches was implemented. Then, the procedure was tested with bimodal (BM) and single-sided-deaf (SSD) CI users. Along with this procedure, speech-in-noise reception thresholds of the CI listeners were collected for each ear and both ears combined.

A.2 Methods

A.2.1 Subjects

Eight NH subjects were tested in Denmark, all of them native German speakers. Their hearing thresholds were below 20 dB HL at all audiometric frequencies, and the mean age was 25.4 years, ranging from 22 to 30 years. The experimental procedure was approved by the Danish Science-Ethics Committee (ref. number H-3-2013-004), and written informed consent was obtained from all subjects before data collection.

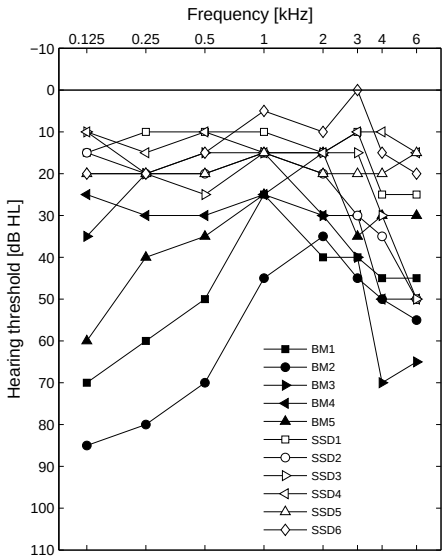

Figure A.1: Individual pure-tone thresholds of the non-implanted side for the CI listeners. Thresholds for the bimodal listeners (filled symbols) were obtained without help of the hearing aid.

Table A.1: Demographics of the CI listeners. The distinction between SSD and BM was based on whether they wore a hearing aid (BM) or not (SSD). Pure-tone threshold averages (PTA) were calculated from the thresholds at 0.5, 1 and 2 kHz, and were all below 20 dB HL for the SSD patients. Insertion depth angles were measured from post-operative CT-scans.

Subject ID	Age at surgery [yrs]	Duration of deafness at at surgery [yrs]	Duration implant use [mths]	Electrode; strategy	CI Side	Insertion depth angle [deg]	PTA of the non-implanted ear [dB HL]
BM1	74	3	15	Helix; HiRes Fidelity 120	right	350	38
BM2	73	20	21	1j; HiRes S	right	452	50
BM3	65	11	15	1j; HiRes Fidelity 120	left	355	20
BM4	60	37	19	1j; HiRes S	left	355	28
BM5	37	5	20	Helix; CIS	right	332	25
SSD1	59	5	16	Helix; HiRes Fidelity 120	right	375	12
SSD2	62	2	19	Helix; HiRes Fidelity 120	right	353	18
SSD3	53	6	20	Helix; HiRes Fidelity 120	right	355	18
SSD4	31	2	19	1j; HiRes Fidelity 120	left	375	13
SSD5	36	2	19	Helix; HiRes Fidelity 120	right	360	18
SSD6	35	20	15	1j; HiRes S	right	337	10

Eleven implant users were tested in the ENT department of the Unfallkrankenhaus in Berlin (UKB), and were all native German speakers. Five BM and six SSD implant users took part in the experiment. Detailed information can be found in Table A.1. All BM and SSD subjects were post-lingually, unilaterally deafened and had similar duration of experience with their implant (mean=18 months, std=2.2 months). The SSD group was, on average, younger (46 yrs) than the BM group (62 yrs), and had a shorter duration of deafness (6 yrs) when they received the CI, compared to the BM group (15 yrs). Individual pure-tone audiometry thresholds are shown in Figure A.1. Aided thresholds were also measured to ensure that all stimuli were audible to the patients. The BM subjects were wearing their hearing aids during all tests described here. The experimental procedure was approved by the ethics committee of Charité Berlin (ref. number EA 4 / 069 / 12), and written informed consent was obtained from all subjects before data collection.

A.2.2 Two-formant vowels test: stimuli and setup

Two-formant vowels were generated using a Matlab-based Klatt synthesizer (Klatt, 1980), and embedded between the consonants /t/ and /k/. The fundamental frequency (F0) was fixed at 110 Hz (male speaker) and the bandwidth of the first and second formant was 90 and 110 Hz, respectively. The duration of the vowels was slightly longer than normal (≈350 ms) for ease of recognition in CI users, and the stimuli were presented at 60 dB SPL. F1 was set to 250 Hz and 400 Hz, and F2 between 600 Hz and 2200 Hz in 200 Hz steps. With these settings, six different German vowels could be elicited when progressively increasing F2 with fixed F1: [u:]/[y:]/[i:] with F1 at 250 Hz and [o:]/[ø:]/[e:] with F1 at 400 Hz (cf. Table A.2).

A monaural (F1 and F2 in the left channel) and a dichotic (F1 in the left and F2 in the right channel) version were created for each stimulus. For the study with NH listeners, the right channel was processed using a vocoder mimicking Advanced Bionics CI processing (Litvak et al., 2007). 15-channel noise excitation was used for this vocoder, with noise bands having 25 dB/octave of attenuation. Three different settings were used: “Voc1”, “Voc2” and “Voc3”. For the condition “Voc1”, the synthesis filters were identical to the analysis filters in order to simulate ideal

Table A.2: Possible vowel choices for the NH subjects during the categorization task. Phonetic equivalent as well as typical F1 and F2 values (Strange et al., 2004) are indicated. 250 Hz was chosen rather than 300 Hz for F1 when synthesizing the vowels to make sure that subjects would differentiate stimuli having two different F1.

Possible choice	TUK	TÜK	TIK	TOK	TÖK	TEK
Phonetic equivalent	[u:]	[y:]	[i:]	[o:]	[ø:]	[e:]
Typical F1 [Hz]	320	301	309	415	393	393
Typical F2 [Hz]	689	1569	1986	683	1388	2010

place pitch. For the "Voc2" and "Voc3" conditions, the idea was to simulate a typical shift (about a fifth, or 7 semitones) and a worst-case shift (more than an octave). Therefore, the synthesis filters were shifted, simulating either a slight mismatch in terms of electrode placement ("Voc2", 2.5 mm shift), or a larger mismatch ("Voc3", 5 mm shift at the apex). For the "Voc3" condition, the mismatch was smaller towards the base for not losing the high frequency content, as it would be in common CI processor settings.

For the NH listeners, Sennheiser HDA 200 headphones were used, ensuring good interaural attenuation (Brännström and Lantz, 2010). Test procedures were implemented in Matlab and all tests were conducted in a double-walled sound attenuating listening booth. For the implant users, the right channel was connected to the implant using the Advanced Bionics Direct Connect® system, i.e. bypassing the microphone, but using the clinical speech processor. The left channel was connected to a loudspeaker, placed 1 meter to the left or right side of the subjects, to stimulate their non-implanted ear. Subjects indicated their responses orally to the experimenter, who was operating the customized Matlab-based interface from outside the booth.

A.2.3 Two-formant vowels test: procedure for NH listeners

NH subjects were asked to categorize each stimulus using one of six possibilities, chosen to match with the frequency range of the stimuli (Table A.2). They could listen to each stimulus up to three times if needed. No feedback was provided after choosing one of the possible choices. The different combinations of F1 and F2

resulted in two blocks of 18 stimuli each: a monaural and a dichotic block.

The first part of the test was performed using the monaural stimuli, and was organized as follows: (1) two repetitions of the stimulus block were presented for training only, (2) five repetitions were recorded (5·18=90 presentations). All stimuli were presented in a random order, and subjects were aware of the number of remaining presentations.

After this first test, the subjects were trained to fuse stimuli that were non-vocoded on one side and vocoded on the other. This was done by listening to 8 minutes of an audio-book, from which the right channel had been vocoded (with the "Voc1" settings) and the left channel lowpass filtered at 500 Hz to mimic a typical audiogram of bimodal listeners. Subjects were asked to listen carefully to both sides, with the aim to train them to combine the non-vocoded and vocoded percepts. This training was successful, as changing the frequencies of F1 and F2 elicited different vowels for all subjects. In a pilot test without listening to the audiobook, 3 out of 4 subjects based their response on F1 only (non vocoded), and therefore changing the frequency of F2 had no effect on their vowel perception.

After this training, nine dichotic sub-tests (three for each vocoder setting, presented in a random order) were administered, following the same protocol as for the monaural test: (1) two repetitions of the dichotic stimulus block were presented for training only; (2) five repetitions of the block were recorded.

A.2.4 Two-formant vowels test: procedure for CI users

The same categorization task was used, but to reduce the duration of the experiment, only stimuli with F1 at 250 Hz were presented. Accordingly, only "TUK", "TÜK", and "TIK" were possible responses during the task. The experiment was divided into two sub-tests, the first one with the monaural stimulus set, and the second with the dichotic set. For each sub-test, the stimulus set was repeated twice for training only, and then 10 repetitions were recorded, all stimuli being randomly presented.

A.2.5 Speech perception of the implant users

Two weeks prior to the vowel test, speech reception thresholds (SRT) of the CI users were measured with the Oldenburg sentence test (OLSA) and the International Female Fluctuating Masker (IFFM) as interferer, that was always fixed at 65 dB SPL (Holube, 2012). Subjects were seated in the booth, with the loudspeaker 1 meter away from the non-implanted ear. First, the SRT was measured for the non-implanted ear only (CI removed, with the HA on for the BM users). Then, the SRT was obtained for the implanted ear through the Direct Connect® system. Every measurement was carried out twice using a different sentence list from the OLSA corpus.

From these two conditions, the difference ΔEl-Ac between the electric and acoustic (non-implanted ear) SRT was calculated. Then, the SRT was measured with the stimuli presented to both sides. ΔEl-Ac was added to the speech level on the electric side in order to provide cues from both ears around the SRT. Otherwise, the combined SRT would rely mainly on the better ear. A 500 ms delay was also added to the IFFM on the electric side, reducing the interaural correlation of the interferer. Therefore, cues obtained by listening in the dips would not be accessible at the same time, further limiting the effect of having a better ear.

A.3 Results

A.3.1 NH listeners

Figure 2 shows the vowel categorization results for the 8 NH listeners. In the top panel (Figure A.2, 1st row), the results of the monaural test are plotted (F1 and F2 in the left channel). When F1 is fixed at 250 Hz (Figure A.2(a)), changing F2 from 600 Hz to 2200 Hz evokes clearly different vowels: [u:] for F2≈800 Hz; [y:] for F2≈1500 Hz; [i:] for F2≈2000 Hz. Individual distributions of the vowel [y:], obtained with only five repetitions, are shown in Figure A.2(b). These patterns are consistent with previously reported North-German vowel maps, e.g. in (Strange et al., 2004). For F1=400 Hz (Figure A.2(c)), similar distributions are observed, but with the three

vowels [o:], [ø:] and [e:].

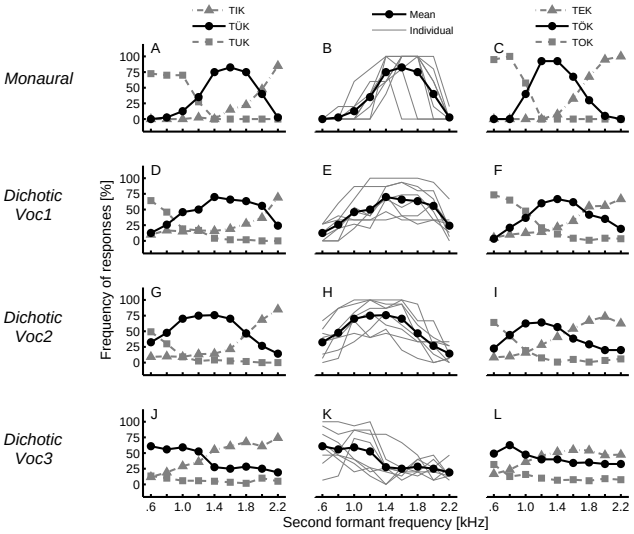

Figure A.2: Mean (N=8) and individual results of the categorization test for the NH listeners. The number of occurrences (in %) is indicated for each vowel as a function of the frequency of F2. For the monaural condition, only 5 repetitions of the stimuli were presented, whereas 15 repetitions were used for the dichotic condition. Left column (A/D/G/J): Mean results with F1 fixed at 250 Hz, therefore only the occurrence of the choices TUK, TÜK and TIK is shown. Middle column (B/E/H/K): Individual results (gray lines) when F1 is fixed at 250 Hz, for the mid-F2 vowel TÜK. The mean is also shown in black. Right column (C/F/I/L): Mean results when F1 is fixed at 400 Hz, only the occurrence of the choices TOK, TÖK and TEK is shown. First line (A-C): monaural condition, F1 and F2 are presented in the left channel. Second line (D-F): dichotic condition, F1 is presented in the left channel while F2 is in the right channel, being processed with an unshifted vocoder ("Voc1"). Third line (G-I): dichotic condition, but with a vocoder slightly shifted ("Voc2"). Last line (J-L): dichotic condition with a vocoder more pronouncedly shifted ("Voc3"). It can be seen with the mid-F2 vowel distribution (black circles) that the distribution is shifting towards the left, due to the simulated shift of the vocoder.

For the dichotic condition, when presenting F2 to the right ear, vocoded without any mismatch ("Voc1"), the three vowel distributions are broader (Figure A.2, 2nd row). This broadening is a direct consequence of the broadening of the individual distributions (Figure A.2(e)), rather than an increased variability of the peak location. This was expected, as the noise-vocoder creates a spread of excitation. However, the distributions still reflect the three different vowels centered at similar values of F2 to without the vocoder, both for F1 at 250 and 400 Hz. For example,

the mid-F2 vowel (black curve) has its distribution centered at 1400 Hz (TÜK) for both conditions. It should be noticed that, in panel E, two subjects exhibit a rather flat distribution of the mid-F2 vowel. This indicates that changing F2 does not have an effect on their vowel perception, highlighting the difficulty to fuse F1 (non-vocoded) and F2 (vocoded). These subjects may have based their choice mostly on F1 perception, rather than achieving spectral fusion. When simulating a shift with the vocoder ("Voc2" and "Voc3"), vowel distributions were affected, as seen in the 3rd and 4th rows of Figure A.2. The low-F2 vowels (TUK and TOK) progressively disappeared. Shifting the vocoder basally assigns channels to higher frequencies. Therefore, F2 frequencies at 600 Hz in the original signal are shifted, evoking vowels having a higher F2 frequency. The high-F2 vowels (TIK and TEK) are represented at more frequencies, and the mid-F2 vowels (TÜK and TÖK) have their distribution shifted towards the left using this representation. Looking at panels E, H and K, showing individual distributions of the mid-F2 vowel when F1 is fixed at 250 Hz (TÜK), it can be observed that changing F2 has an effect for most of the subjects. However, for the larger shift ("Voc3"), a higher number of subjects show a flat distribution. This can be explained both by the difficulty to achieve spectral fusion and by the fact that subjects could have been confused by perceiving only the low and mid-F2 vowels during this test condition. These individual flat distributions broaden the mean distribution (Figure A.2(j) & (l)).

A.3.2 CI listeners

Vowel distributions measured for the implant users are shown in Figure A.3. The top panels show the results of the monaural condition, in which both formants were presented acoustically, and the bottom panels present the dichotic results, for F2 presented to the implant. For the monaural condition, these are very similar to the NH listeners' distributions: the three categories (TUK, TÜK, and TIK) are similarly distributed over the F2 frequency range (Figure A.3(a)). Moreover, individual distributions of the mid-F2 vowel for the same condition (Figure A.3(b)) show a very good agreement across subjects, even though five of them wore a HA, and some of the SSD subjects had mild hearing losses in the non-implanted ear.

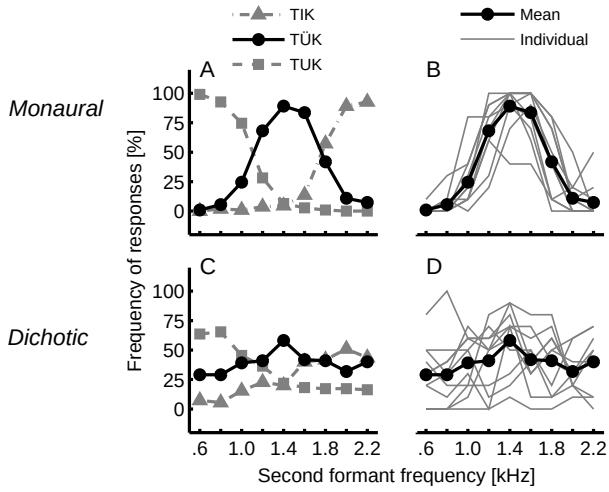

Figure A.3: Mean (N=11) and individual results of the of the categorization test for the CI listeners. Left panels (A/C): Mean results of the occurrence of the three possible choices: TUK, TÜK and TIK. Right panels (B/D): Individual (gray lines) and mean (dark circles) results for the mid-F2 vowel TÜK. Top panels (A-B): monaural condition, F1 and F2 are presented acoustically. Bottom panels (C-D): dichotic condition, F1 is presented acoustically and F2 electrically. In the bottom panels, the large variability when presenting F2 to the CI can be seen.

For the dichotic condition, when F2 is presented to the implant while F1 is kept on the acoustic side, the variability across subjects increases dramatically (Figure A.3(d)). To highlight this variability, a subset of five subjects' dichotic responses is shown in Figure A.4. Subject SSD1 was the only one with a clear pattern for the three vowels, centered at values similar to the monaural condition. Other subjects never perceived either the low-F2 (Figure A.4(d)), mid-F2 (Figure A.4(b)), or high-F2 vowel (Figure A.4(e)), and some subjects confused vowels, for example the low and mid-F2 vowel for subject BM2 (Figure A.4(c)).

The mean distributions for the eleven CI users (Figure A.3(c)) are broader than in the monaural condition. This broadening results from the individual variability, rather than from broad individual distributions, as seen in the results obtained in the NH subjects. It is also interesting to notice that the mean distributions of the dichotic condition, despite being shallower, are centered at F2 values similar to

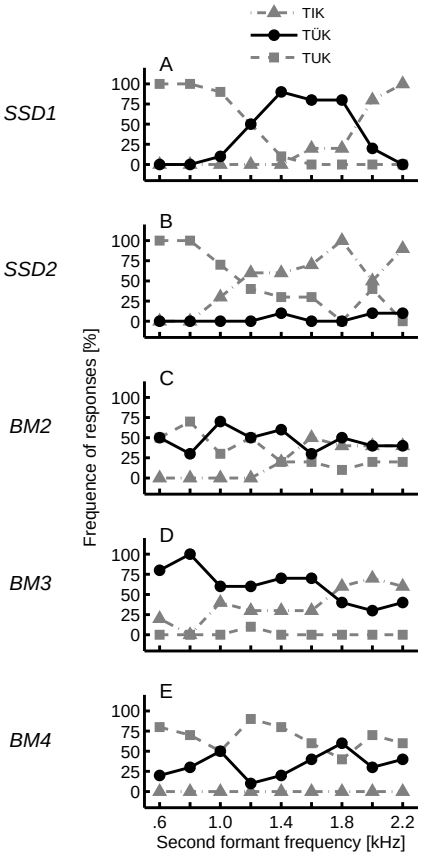

Figure A.4: Individual results of the categorization test for a subset of CI listeners for the dichotic condition (F1 presented acoustically, F2 electrically). Very different patterns can be observed, with only one (panel a) resembling the results of the monaural condition.

the monaural condition, especially for the low and high-F2 vowels.

A.3.3 Speech perception results of the CI listeners

Speech reception thresholds measured with the OLSA test in an IFFM background are shown in Figure A.5. For each subject, the gray squares indicate the SRT mea-

sured in the non-implanted ear ("Acoustic"), the gray circles the SRT measured in the CI ear, and the dark triangles show the combined acoustic-electric thresholds. The acoustic ear was significantly better than the electric ear for the whole group ($p<0.001$, Wilcoxon rank sum test), but not for the BM group ($p=0.095$, Wilcoxon). For the condition where acoustic and CI stimulation were combined (Ac.+CI), the speech level was adjusted and the IFFM was uncorrelated between the ears (cf. Methods) to encourage subjects to use cues from both sides. Using this particular setup, no significant change was observed between the Acoustic and Ac.+CI conditions for the whole group ($p=0.84$, Wilcoxon), or for each group (BM, $p=0.42$; SSD, $p=0.94$). Despite this, the difference was always towards an improvement when adding the CI for the BM group (up to 5 dB, compare triangles to squares in Figure A.5). For the SSD group, no such trend could be seen, with the effect of adding the CI being either positive or negative across subjects. This indicates that BM listeners benefited more than the SSD listeners from having the CI, with this particular setup. These results suggest that the population tested here relies mainly on the information from the acoustically stimulated ear, as previously reported with a similar cohort in Vermeire and Van de Heyning (2009).

A.4 Discussion

The results of the NH listeners showed that fusing dichotic vowels vocoded in one ear is possible, and that the two-formant-vowel method can in principle be used to derive an interaural mismatch in place of stimulation. CI listeners could perform the monaural task reliably but their dichotic results showed a large individual variability. These aspects are discussed in more details below.

A.4.1 Design of the experiment using NH listeners

The binaural fusion of different frequency bands to form an object, referred to as spectral fusion in Cutting (1976), has been shown using vowels synthesized with the same system for each ear (Broadbent and Ladefoged, 1957; Takanen et al., 2013), and therefore with a similar percept on each side. Here, NH listeners were

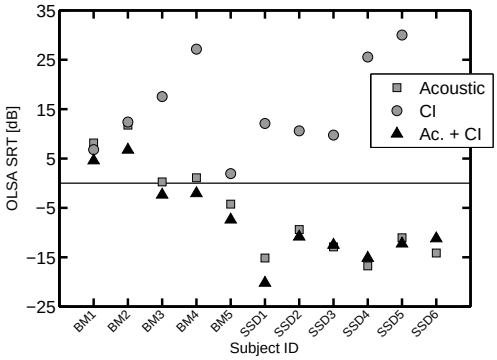

Figure A.5: Speech reception thresholds of the CI listeners, using the Oldenburg sentence test (OLSA) in an International Female Fluctuating Masker (IFFM) background. Better speech perception is indicated by lower SRTs, and is mainly driven by audibility for the acoustic presentation of the stimuli (gray squares), as the SSD subjects have lower thresholds than the BM subjects. When tested with the CI only (gray circles), the thresholds are often higher than for the acoustic condition, especially for the SSD subjects. For the combined condition (dark triangles), thresholds are similar to the ones obtained with the acoustic presentation. Subject SSD6 reached the upper limit of the dynamic range (clipping at the CI) for the CI condition.

able to achieve spectral fusion when F1 was created with a pulse-excited system (Klatt synthesizer) and F2 was noise-vocoded when presented to the opposite ear. During the first pilot tests, subjects often perceived two different auditory events, typically perceiving a vowel on the non-vocoded side, and noise on the opposite, vocoded, side. Prior training with an audiobook having the left channel low-pass filtered and the right channel vocoded was sufficient to overcome this issue (at least for the “Voc1” and “Voc2” conditions). A noise-vocoder rather than a sine-vocoder was chosen in order to simulate the difference in perceptual quality between electric and acoustic stimulation. The efficiency of this short training (8 min) with an audiobook might indicate that it is easier to fuse the two percepts in the NH procedure than fusing the electric and acoustic percepts for CI listeners.

Setting F1 at 250 Hz, three different vowels could be perceived by changing F2 from 600 to 2200 Hz ([u:], [y:] and [i:]). A similar pattern was observed having F1 set at 400 Hz, with the vowels [o:], [ø:] and [e:]. Due to this three-vowel distribution obtained by varying F2 only, an indirect measure of frequency perception can be

derived when comparing the monaural and dichotic conditions, with the monaural condition acting as a reference. When simulating a frequency mismatch by shifting the synthesis filters of the vocoder, an effect could be seen in the vowel distributions, both individually and in the group average (Figure A.2): the F2 center frequencies were shifted for each vowel, reflecting the interaural mismatch. However, this effect may be expected to be more salient in such NH listeners, who did not have time to adapt to the mismatch, than in CI listeners, who might have acclimatized to a potential interaural mismatch.

A.4.2 Monaural results in CI listeners

Implant users were able to perform the task reliably in the monaural condition, with low inter- and intra-subject variability (Figure A.3(a-b)). This part of the test was achieved in 15 minutes (30 minutes with the dichotic condition), and was easy to explain to the subjects. In comparison to classical pitch-matching experiments, where training is necessary for the subject to perform the task, this more ecological approach thus seems promising, as the population wearing implants is far from the cohort of young students usually tested in such psychoacoustic experiments.

The bimodal population tested here was atypical, as most BM subjects had a severe low-frequency hearing loss, which can lead to a distorted perception of pitch. Interestingly, the bimodal subjects (wearing their HA during the experiment) performed well in this monaural vowel discrimination task (see the low across-subjects variability in Figure A.3(b)), even though their OLSA SRT was significantly higher that of the SSD group ($p<0.01$, Wilcoxon rank sum test). This does allow conclusions about their pitch perception ability, but it indicates that formant discrimination is preserved when assessed with synthesized two-formant vowels.

A.4.3 Variability in the dichotic results of CI listeners

Individual results of the implant users for the dichotic condition (Figure A.3(d)) showed large differences across subjects. Only one subject had a similar distribution for both conditions (SSD1, Figure A.4(a)), whereas the others showed very different patterns (a few examples can be seen in Figure A.4). All subjects had a

similar experience with their implant (18 months) and a similar insertion depth (mean=363 deg, std=31 deg). In the study from Harnsberger et al. (2001), no systematic shift was reported in the monaural, individual, vowel maps they recorded, and the variability was attributed to individual differences in formant frequency discrimination. Here, only one formant was presented to the CI, therefore reducing the variability stemming from individual differences in formant frequency discrimination. The large difference in the individual results could also be caused by difficulties to fuse F1 and F2 when they have different perceptual qualities, as reported in classical pitch-matching experiments (Carlyon et al., 2010). Abnormal binaural spectral integration has also been shown in bimodal subjects, which could account for the difficulties in integrating information from both ears (Reiss et al., 2014). Here, some subjects showed a flat distribution (Figure A.4(c)), suggesting that they based their response on F1 only, presented acoustically. Moreover, subjects having speech thresholds similar to subject SSD1 on the implanted side (suggesting an equivalent vowel perception) showed very different results in the dichotic test condition. Taken together, these findings suggest an insufficient fusion between the electric and acoustic percepts as the main reason for the variability seen in the results. Whether matching the place of excitation in the presence of these very different perceptual qualities would help in terms of speech perception remains unknown, as other factors (e.g. binaural spectral integration) might also be contributing (Francart and McDermott, 2013; Reiss et al., 2014).

Shallow distributions can be seen in the mean results of the CI listeners (Figure A.3(c)), given the large individual variability. The distribution of the mid-F2 vowel (TÜK) is flat, but the low-F2 and high-F2 vowels have mean distributions with patterns centered at values similar to the monaural condition. This suggests that, after 18 months of implantation, implanted listeners may be acclimatized to the new tonotopic organization given by the implant, consistent with previous studies (McDermott et al., 2009; Reiss et al., 2007).

A.4.4 Further investigations

The two-formant-vowel method described in this study was intended to estimate interaural frequency place mismatches. However, the results suggest that abnormal bimodal vowel integration is prominent in the BM and SSD subjects tested here, which limits the possibility of estimating this interaural mismatch. Further investigations on such vowel integration would be relevant, especially regarding the growing implanted population with residual hearing in the contralateral ear. This could include a study of binaural spectral integration considering dichotic pitch fusion (Reiss et al., 2014) and formant fusion.

Furthermore, regarding the use of this method for mismatch evaluation, the effects of training, adaptive procedures, a comparison with a classical pitch-matching experiment, and testing with both formants presented to the CI should be considered to better understand the sources of the individual variability. Training with an audiobook appeared to be very efficient with the NH subjects, but might be underestimating the difficulties of bimodal spectral integration. Testing with both formants stimulated electrically would provide information on individual formant frequency discrimination, which could be a potential cause of individual variability.

The CI population tested here had the acoustic ear as the better ear, which is not typical of a BM population. This raises some challenges, e.g. in terms of assessing the speech in noise benefit of the combined electric-acoustic stimulation, where the SRT but also the percepts differ between the two ears. Methods to test both ears simultaneously at their own SRT on these listeners have not yet been described.

If the two-formant-vowel procedure was tested in a more typical BM population, with only very low frequencies preserved, F2 could be too high in frequency to be perceivable. Instead of measuring the monaural condition, a reference acoustic vowel map could be used for the estimation of the mismatch. It is also known that pitch perception changes over time after implantation (Reiss et al., 2007), or after a change in the frequency-to-electrode map (Svirsky et al., 2004). These are points that were not tested here, but should be replicated with the two-formant

vowel method for its validation. Finally, only bimodal German-speaking listeners participated in this experiment, but this protocol could also be applied to other languages with a few modifications in the stimuli, as well as listeners with residual hearing in the implanted ear or bilateral implant users.

A.5 Conclusions

NH listeners' results and mean results of the CI listeners suggest that place mismatches can be derived from vowel spaces obtained when presenting two-formant vowels monaurally and dichotically in SSD and BM listeners. This test is also easier and less time consuming to perform for subjects than a classical pitch-matching paradigm. However, the method's reliability remains very limited by the individual variability, and results mostly indicate an abnormal bimodal vowel integration within the CI population tested here.

Declaration of interest:

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper. Part of this work was presented at the 4th International Symposium on Auditory and Audiological Research held in Nyborg, Denmark, August 2013, and at the 37th Annual Midwinter Meeting of the Association for Research in Otolaryngology held in San Diego, California, February 2014.

Acknowledgements

The authors would like to thank Franziska Barthold for conducting the tests with the CI patients in Berlin, as well as Ingo Todt for analyzing CT scans to estimate insertion depth.

B

Linear combination of auditory steady-state responses evoked by co-modulated tones[a]

Abstract

Up to medium intensities and in the 80-100-Hz region, the auditory steady-state response (ASSR) to a multi-tone carrier is commonly considered to be a linear sum of the dipoles from each tone specific ASSR generator. Here, this hypothesis was investigated when a *unique* modulation frequency is used for all carrier components. Listeners were presented with a co-modulated dual-frequency carrier (1 and 4 kHz), from which the modulator starting phase Φ_i of the 1-kHz component was systematically varied. The results are supporting the hypothesis of a linear superposition of the dipoles originating from different frequency specific ASSR generators.

B.1 Introduction

The auditory steady-state response (ASSR) is an auditory evoked potential which follows the repetition rate, defined by the modulation frequency f_m, of an ongoing sound signal (Picton et al., 2003). For repetition rates between 80 and 100 Hz, the ASSR has been shown to arise from brainstem sources, while at lower rates (below 40 Hz), mostly sub-cortical and cortical sources are involved (Herdman et al., 2002).

[a] This chapter is based on Guérit et al. (2017).

When evoking the ASSR with sinusoidally amplitude-modulated (SAM) tones, an activation of auditory nerve fibers within a narrow region of the basilar membrane (Picton et al., 2003) is assumed. The response to multiple SAM tone carriers with differing modulation frequencies has been shown to be a linear combination of the responses to each SAM component in the 80-100 Hz range of repetition rates (e.g., Herdman et al., 2002). However, for modulation frequencies around and below 40 Hz, multiple ASSR components do not combine linearly (John et al., 1998), presumably because of interactions within the sub-cortical and cortical sources of the ASSR.

At higher stimulation levels, this linear combination of the ASSR components does not hold (Picton et al., 2007). This can be explained by the nonlinear mechanics of the auditory periphery: a travelling wave excited by a pure tone carrier does not only result in an isolated vibration around the peak region of the carrier, but also evokes vibrations basal to that region. Stimuli presented at higher levels and composed of multiple frequency components are thus likely interacting across different regions along the basilar membrane. The contribution of different tonotopic regions to the ASSR has also been addressed in the context of chirp-evoked ASSRs (Elberling et al., 2007), where it was found that the amplitude of the ASSR can be increased by stimulation with chirps accounting for the dispersion properties of the basilar membrane. For these stimuli, it is however not clear how each tonotopic region contributes to the measured ASSR other than that the overall amplitude increases.

For binaural stimulation with modulation frequencies around 80-Hz, a linear combination of ASSRs has been shown, also for components having the same modulation frequency, suggesting either the independence of two separate sources, or the linearity of a unique source of ASSR (e.g., Poelmans et al., 2012).

Here, the assumption of a linear, monaural superposition of multiple co-modulated sources of ASSR in the 80 Hz region was investigated. The ASSR was recorded with electroencephalography (EEG), and was evoked by two SAM tones centred respectively at 1 and 4 kHz. Both carriers were modulated with the same modulation frequency but with a relative phase that was varied across conditions. It is hypothesized that the overall response measured using EEG is the vector sum of the ASSR

evoked by each SAM tone separately, and will be sensitive to the relative modulator phase between the SAM tones.

The results will contribute to the understanding of how multiple sources of ASSR combine into the electrical signal measured at the scalp.

B.2 Methods

B.2.1 Subjects

Nine subjects participated in the experiment. Their hearing thresholds were below 20 dB HL at all audiometric frequencies (125 Hz to 8 kHz), and the mean age was 29.8 years, ranging from 25 to 40 years. The experimental procedure was approved by the Danish Science-Ethics Committee (ref. number H-3-2013-004), and written informed consent was obtained from all subjects before data collection.

B.2.2 Stimulus and apparatus

Seven different stimuli were used to elicit ASSRs, consisting of two SAM tones, s_{1k} and s_{4k}, and of five combined versions of those same tones with varied modulator starting phases of $s_1 k$. The carrier frequencies of the two tones, f_{1k} and f_{4k}, were set respectively at 1 and 4 kHz. The carriers were 100% modulated at a frequency f_m of 88 Hz, as shown in equations (1) and (2).

$$s_{1k}(t) = \quad a_{1k} \cdot \sin(2\pi f_{1k} t) \cdot \left(\frac{1+\sin(2\pi f_m t + \Phi_i)}{2} \right) \tag{B.1}$$

$$s_{4k}(t) = \quad a_{4k} \cdot \sin(2\pi f_{4k} t) \cdot \left(\frac{1+\sin(2\pi f_m t)}{2} \right) \tag{B.2}$$

When s_{1k} was presented in isolation, its modulator starting phase Φ_i was set to 0. For the five co-modulated conditions, stimuli were created by setting Φ_i to values distributed around the unit circle ($c^{\Phi_i} = s_{1k}^{\Phi_i} + s_{4k}; \quad \Phi_i = \frac{2i\pi}{5}$, $i = 0, 1, .., 4$), while s_{4k} was kept the same.

To avoid distortions in the co-modulated conditions, the two carriers were played separately through two ER-2 earphones mounted on an ER-10B+ probe (Etymotic Research, Inc.), and connected to the computer through a Phonitor mini amplifier (SPL electronics GmbH) and a Fireface UCX sound card (Audio AG). Both a_{1k} and a_{4k} were adjusted to deliver s_{1k} and s_{4k} at 65 dB SPL in isolation, using a B&K 4137 ear coupler and a B&K 2636 sound level meter (Brüel & Kjær A/S).

B.2.3 ASSR recording and analysis

Subjects were seated in a double-walled, electrically shielded, sound-attenuating booth. They were instructed to relax and stay calm. They watched a silent film with subtitles throughout the whole recording session, and were awake at all time. The stimulated ear was randomized across subjects, and the opposite ear was occluded with an ear plug, to avoid acoustical cross-talk.

EEG signals were recorded using a BioSemi ActiveTwo system (Biosemi B.V.), sampled at 8192 Hz, and analyzed offline with Matlab (The MathWorks, Inc.).

A vertical electrode montage was used, using the 10/20 system, with two electrodes: P9 or P10 at the left or right mastoid, respectively, and Cz at the vertex. If the right ear was stimulated, the difference between Cz and P10 was computed, while Cz and P9 were used for the left ear stimulation. Each stimulus condition was recorded for approximately 10 minutes (608 seconds). The signal was cut into epochs of 16 seconds, and any epoch exceeding 80 μV was discarded from the processing. A weighted averaging method based on the standard deviation in each epoch (John et al., 2001) was then applied to obtain a single 16-seconds epoch, from which the Fast Fourier Transform (FFT) was computed with a bin width of 0.0625 Hz. A F-ratio was computed between the power of the FFT bin at 88 Hz (chi-squared variable with 2 degrees of freedom) and the power of the EEG background noise (96 neighbouring bins, ±3 Hz, $96 \cdot 2$ degrees of freedom). The ASSR was deemed above the noise floor when the null hypothesis that both noise and ASSR component came from the same F distribution was rejected ($p \leqslant 0.01$, Dobie and Wilson, 1996). This corresponds to a signal-to-noise ratio above or equal to 6.73 dB ($= 10 \cdot \log_{10} \frac{P_{signal+noise}}{P_{noise}}$).

Due to anatomical differences (head size, neural sources), inter-subject variability in the group delay (and therefore the phase) is expected. Because of this, co-modulated responses are likely to be in/out of phase for different values of Φ_i across listeners. Measured amplitude responses to the c^{Φ_i} stimuli were therefore shifted to have their maximum value at $\Phi_i = 0$ rad (Riecke et al., 2015).

Before computing the phase of the co-modulated ASSRs, the response to s_{4k} in isolation was subtracted (**ASSR**$\left(c^{\Phi_i}\right)$−**ASSR**(s_{4k})). In case of linearity, the phase of this vector subtraction should therefore equal Φ_i, the phase of s_{1k}. Again, to account for inter-subject variability in group delay, the phase of the afore-mentioned subtraction was shifted to be 0 rad for Φ_0. Unless specified, the Φ_0 condition was removed from all statistical analysis, as data for this point do not satisfy independence requirements.

B.3 Results

B.3.1 Responses to single carriers

The amplitudes of the ASSRs were above the noise floor for 8 out of 9 subjects in response to s_{1k}, and 7 out of 9 in response to s_{4k}. They were similar in amplitude (figure B.1B) and comparable in value to previously reported amplitudes at those stimulation levels (Picton et al., 2007). Overall (including the responses to co-modulated carriers), mean amplitude and standard deviation of the significant ASSRs were respectively 52.5 and 26.5 nV. One subject had higher noise levels (mean/s.d. of 48.9/4.3 nV versus a mean and s.d. of respectively 10.7 and 2.5 nV for the other subjects). Since this subject did show significant ASSRs in some conditions, it was not excluded. We however controlled for every statistical analysis that removing this subject did not change the main conclusions.

The difference in phase between the responses to s_{1k} and s_{4k} was 103°, as shown in figure B.1A. In order to link the ASSR phase to an estimate of cochlear travel time, and assuming a linear phase of the frequency components along the cochlea, this phase corresponds to a latency difference of 3.3 ms for a modulation frequency of 88 Hz ($\frac{103}{360 \cdot 88}$). This difference was statistically significant (Paired-Sample t-Test

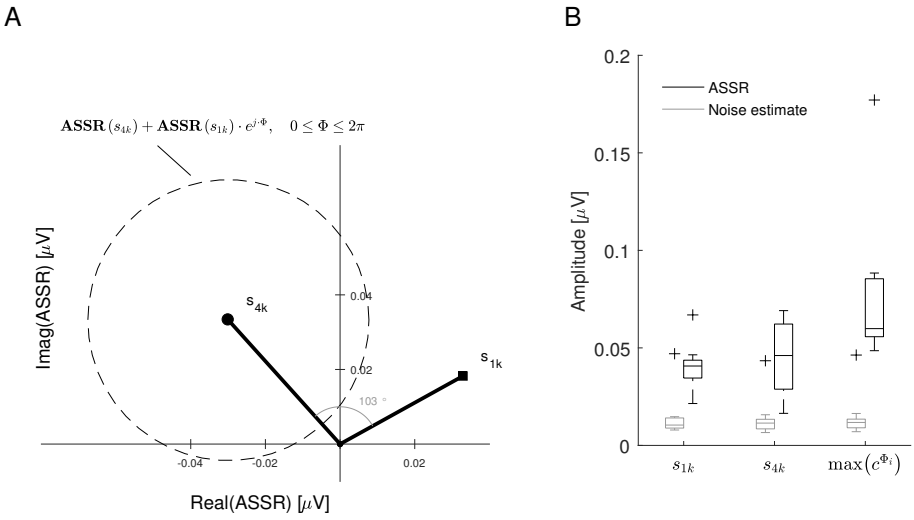

Figure B.1: **A.** Averaged complex-valued ASSRs across subjects in response to the s_{1k} (square) and s_{4k} (circle) stimuli. Shown with a dashed line is the expected co-modulated response (linear vector summation) when varying Φ_i, the modulator starting phase of s_{1k}. **B.** Across-subject distribution of ASSR amplitudes (black) and of noise estimates (grey) in response to the s_{1k}, s_{4k} and c^{Φ_i} stimuli. For c^{Φ_i}, only the maximum amplitude across all angles is shown. *Lower and Upper limits of the boxes:* 25th and 75th percentiles. *Horizontal line:* median. *Whiskers:* 25th (or 75th) percentile minus (or plus) 1.5 the interquartile range.

with the 7 subjects having both responses above significance; df = 6; p = 0.0019, t = 5.2691, 95% confidence interval = 1.7 - 4.8 ms).

B.3.2 Co-modulated responses

By combining equations (1) and (2), one can hypothesize that the mean vector sum of all co-modulated responses should equal the response to s_{4k} in isolation,

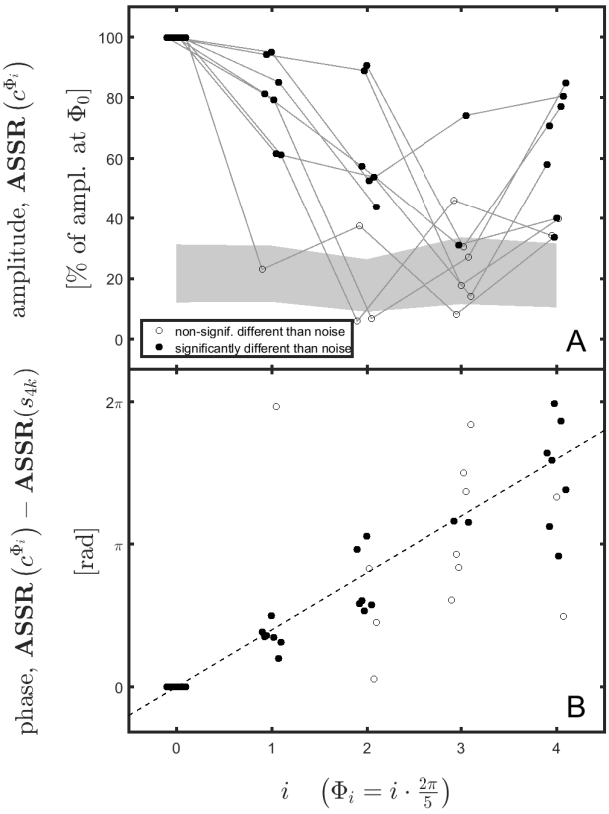

Figure B.2: **A.** Amplitudes of individual co-modulated responses, shifted to have their maximum at Φ_0, and normalized by their value at Φ_0. A small jitter has been added to the x-axis to improve readability, and the noise floor excursion (min to max) is shown with the grey shaded area. As noise levels differed across subjects, it can be seen that the amplitude (on a relative scale) required to have a significant response varies across subjects. **B.** Phase of the vector subtraction $\mathbf{ASSR}\left(c^{\Phi_i}\right) - \mathbf{ASSR}(s_{4k})$ (which should be Φ_i in case of linearity), normalized to be 0 at Φ_0 and wrapped between 0 and 2π.

as shown in equations (3) to (5) and figure B.1.

$$\frac{1}{5}\cdot\sum_{i=0}^{4}\mathbf{ASSR}\left(c^{\Phi_i}\right)=\frac{1}{5}\cdot\sum_{i=0}^{4}\left(\mathbf{ASSR}(s_{4k})+\mathbf{ASSR}(s_{1k})\cdot e^{j\cdot\frac{2i\pi}{5}}\right) \tag{B.3}$$

$$=\mathbf{ASSR}(s_{4k})+\mathbf{ASSR}(s_{1k})\cdot\sum_{i=0}^{4}e^{j\cdot\frac{2i\pi}{5}} \tag{B.4}$$

$$=\mathbf{ASSR}(s_{4k}) \tag{B.5}$$

A t-test comparing the mean vector sum of all co-modulated responses to the response to s_{4k} in isolation showed no significant difference (df = 6, $p_{real} = 0.7617$, $t_{real} = -0.3173$, $\mathrm{p}_{imaginary} = 0.5832$, $\mathrm{t}_{imaginary} = -0.5798$, Pearson's $r = 0.8322$ when pooling real and imaginary values, $r = 0.8310$ when removing the subject with high noise level). This t-test excluded two subjects who had missing data in one condition (hence df = 6), and was run on both real and imaginary parts of the ASSR, as they can be considered to be independent variables (Dobie and Wilson, 1996).

Amplitude

As shown in figure B.1B, the individual maximum for each subject across all co-modulated conditions was significantly larger than the response to the single carriers in isolation (pairwise t-Tests, paired data within subjects, Bonferroni corrections, p = 0.00023 and p = 0.000232 for s_{1k} and s_{4k}, respectively). The data was log-transformed for this test to account for the presence of a subject with higher overall amplitudes. Figure B.2A shows the amplitudes obtained for all co-modulated stimuli, with the individual responses aligned to be largest at $\Phi_0 = 0$ rad, and normalized by their value at Φ_0. A multilevel approach for repeated measures was employed, with the subjects as a random factor (Field et al., 2012), and failed to show a significant effect of Φ_i on the relative amplitude as plotted in figure B.2A ($\chi^2(3) = 6.98$, p = 0.0726, data points below the noise floor excluded). By adding the points below the noise floor in this statistical analysis, the effect of Φ_i becomes significant ($\chi^2(3) = 13.67$, p = 0.0034). It is worth noticing that conditions Φ_2 and

Φ_3 had the highest number of recordings below the noise floor (figure B.2B, where s_{1k} and s_{4k} were expected to be out of phase. To assess whether the proportion of significant points was the same between different Φ_i conditions, a Cochran Q test was used, and showed a significant effect of Φ_i (Q = 8.4000, df = 3, p = 0.0384).

Phase

Figure B.2B shows the phase values corresponding to figure B.2A, re-referenced to be 0 for Φ_0 and wrapped between 0 and 2π. Φ_i had a significant effect on the ASSR phase ($\chi^2(3) = 37.8$, $p < 0.0001$). This effect was well fitted by a linear regression (intercept = -0.1255, slope = 0.9967, 95% confidence interval = 0.75 - 1.19, r^2 = 0.7956).

B.4 Discussion

In the 80-100-Hz range, it has been hypothesized that the ASSR evoked by SAM tones with different modulation frequencies is the linear superposition of the response to the SAM tones presented alone. This assumption has been shown to hold true if the carrier frequencies are separated at least by an octave, and if medium levels are used (Herdman et al., 2002). The present study supports the hypothesis of a linear superposition, and expands it to the case of carriers modulated with a unique modulation frequency presented monaurally (this has already been shown binaurally, e.g. in Poelmans et al., 2012). However, because the ASSR measured by EEG is a gross potential, it can not be distinguished whether the observed effects in the presented paradigm are due to a superposition of two independent sources contributing to the ASSR or if the effects are due to neural interactions within a single source of the ASSR.

B.4.1 Linearity of the co-modulated conditions

Under our linearity assumption, and because Φ_i was evenly distributed around the unit circle, summing all co-modulated responses should not be significantly

different than the response to s_{4k} in isolation, and this is indeed what we could see in our recordings.

Additionally, manipulating the modulator starting phase of s_{1k} in the co-modulated conditions had a significant effect on both phase and amplitude of the ASSR. This effect was consistent with a linear sum when analyzing the co-modulated phase response (figure B.2B), while the individual patterns of the amplitude were more variable (figure B.2A). These deviations seen in the amplitude of the co-modulated responses might be due to the inherent test-retest variability of the ASSR. Finally, when both single carriers were supposedly out of phase (conditions Φ_2 and Φ_3 in figure B.2), it was often impossible to record a significant response, even with 10 minutes of recording and median noise levels of 10.9 nV.

Taken together, these results support the hypothesis of a linear superposition, and that distinct neural populations are represented in the ASSR, even when using a unique modulation frequency. John et al. (2003) measured the ASSR of 4 SAM tones at 0.5, 1, 2 and 4 kHz, co-modulated and in isolation. The co-modulated response was 25% lower than expected by a linear vector superposition of the responses in isolation. This reduction, not seen in our study, might be explained by the fact that they used four carriers separated by only one octave (while we used only two carriers separated by two octaves). This might have led to interactions at the level of the basilar membrane, such as mutual suppression.

In contrast to multi-tone carrier ASSRs, single-evoked ASSR growth functions do not show a saturation for stimulus levels above 60 dB SPL (Picton et al., 2007). Based on the results of the present study, one might however speculate that responses evoked by off-frequency regions also contribute to the measured amplitude in single-evoked ASSRs, and that the measured ASSR is a linear combination of responses evoked by on- and multiple off-frequency regions with different relative phase.

B.4.2 Further use of this paradigm

As linearity seems to be respected with this paradigm, any measured non-linearity could be used as a marker for envelope interactions at the level of the cochlea.

An example is for cochlear implant users, where the spread of electrical current produces marked envelope interactions in a behavioural task (Galvin et al., 2015).

B.5 Conclusions

This study suggests that the ASSR at 88 Hz with co-modulated carriers presented monaurally is a linear sum of the response to each carrier, as supported by the phase behaviour of the co-modulated response and the vector sum of all co-modulated responses.

Such a paradigm, where the phase difference between co-modulated carriers is varied, is therefore suitable for analyzing envelope interactions with a unique modulation frequency and at peripheral levels of the auditory system.

Acknowledgements

FG is supported by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS) and the Technical University of Denmark. The authors would like to thank Gerard Encina Llamas, for the ASSR analysis code, and James Harte for fruitful discussions. The authors would also like to thank two anonymous reviewers for constructive comments on a previous version of this manuscript.

Supplementary data

Raw data files (.bdf and stimuli files) are accessible on Zenodo.

C

Minimum step size achievable with HiRes90k device

In chapters 2 and 4, the listeners were recipient of an Advanced Bionics HiRes90k device. Stimulation was controlled using the software BEDCS 1.18 and a PSP processor. With this combination of hardware and software, the minimum step size for symmetric biphasic pulses depends on the dynamic range which is used (1 μA for 0-255 μA, 2 μA for 0-510 μA, 4 μA for 1020 μA, 8 μA for 2040 μA). For asymmetric pulses, previous studies (e.g., Macherey et al., 2006) have assumed that the minimum step size was multiplied by the ratio of asymmetry (i.e 8 μA for 0-255 μA, 16 for 0-510 μA, etc..).

However, a closer look at the oscilloscope output of a HiRes90k test implant with different asymmetry ratios and ranges (Figure C.1) reveals that smaller steps are achievable with asymmetric pulses. Table C.1 summarizes the derived minimum step size achievable for the short-high and long-low phases of asymmetric pulse.

Table C.1: Minimum step size achievable with asymmetric pulses, based on results from Figure C.1

Range	Min. step size short-high phase (μA)	Min. step size long-low phase (μA)
1: 0 to 255 μA	1	0.25
1: 0 to 510 μA	2	0.5
1: 0 to 1020 μA	4	1
1: 0 to 2040 μA	8	2

Based on those values, equation C.1 gives the minimum achievable step size as a function of the range and asymmetry ratio used.

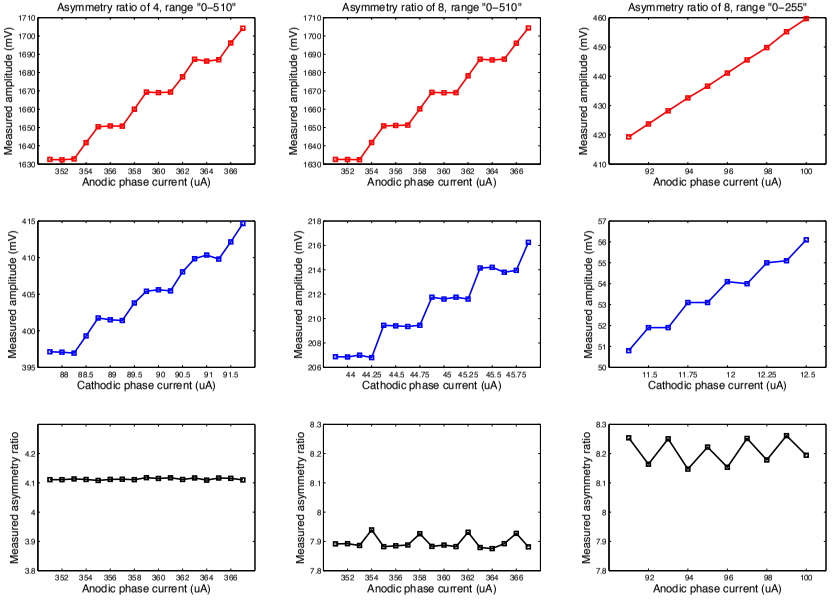

Figure C.1: Oscilloscope measurements with different asymmetry ratios and ranges. **Top line:** Amplitude of the short-high phase (anodic). **Middle line:** Amplitude of the long-low phase (cathodic). **Bottom line:** Ratio of amplitude. **Left column:** Asymmetric pulse with a ratio of 4, using range 2 (0 to 510 µA). **Middle column:** Asymmetric pulse with a ratio of 8, using range 2 (0 to 510 µA). **Right column:** Asymmetric pulse with a ratio of 8, using range 2 (0 to 255 µA).

$$step\left(range, ratio\right) = \max\left(1, \frac{ratio}{4}\right) \cdot 2^{range-1} \tag{C.1}$$

For a ratio of 8, that we used in chapters 2 and 4 this leads to a minimum step size of 2 μA, in range 1, 4 μA in range 2, etc. Figure C.2 shows the step size in dB, for a ratio of 8. This is relevant mostly for the determination of thresholds in chapter 2 where the minimum step size of the adaptive procedure was 0.25 dB. Below 36 dB re 1 μA, this value of 0.25 dB can not be achieved. As described in chapter 2, we tracked the level based on a desired value, but calculated the thresholds on the actual values, achievable with the HiRes90k device.

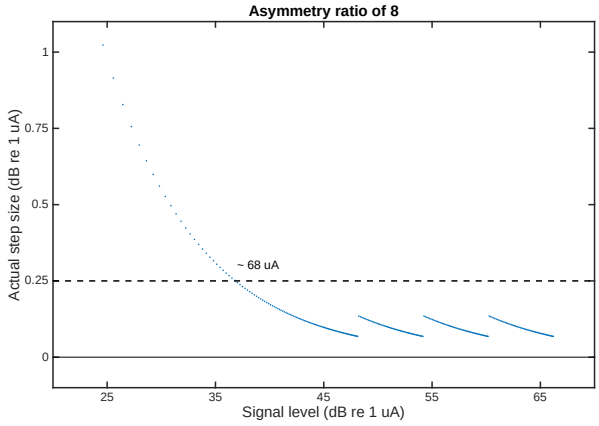

Figure C.2: Actual step size achievable with a ratio of asymmetry of 8. Below 36 dB re 1 μA (approx. 68 μA), the actual step size is larger than that of the tracking procedure used in Chapter 2 (0.25 dB, dashed line). The lowest detection threshold measured across all subjects and conditions was at 38.3 dB re 1 μA.

Contributions to Hearing Research

Vol. 1: *Gilles Pigasse*, Deriving cochlear delays in humans using otoacoustic emissions and auditory evoked potentials, 2008.

Vol. 2: *Olaf Strelcyk*, Peripheral auditory processing and speech reception in impaired hearing, 2009.

Vol. 3: *Eric R. Thompson*, Characterizing binaural processing of amplitude-modulated sounds, 2009.

Vol. 4: *Tobias Piechowiak*, Spectro-temporal analysis of complex sounds in the human auditory system, 2009.

Vol. 5: *Jens Bo Nielsen*, Assessment of speech intelligibility in background noise and reverberation, 2009.

Vol. 6: *Helen Connor*, Hearing aid amplification at soft input levels, 2010.

Vol. 7: *Morten Løve Jepsen*, Modeling auditory processing and speech perception in hearing-impaired listeners, 2010.

Vol. 8: *Sarah Verhulst*, Characterizing and modeling dynamic processes in the cochlea using otoacoustic emissions, 2010.

Vol. 9: *Sylvain Favrot*, A loudspeaker-based room auralization system for auditory research, 2010.

Vol. 10: *Sébastien Santurette*, Neural coding and perception of pitch in the normal and impaired human auditory system, 2011.

Vol. 11: *Iris Arweiler*, Processing of spatial sounds in the impaired auditory system, 2011.

Vol. 12: *Filip Munch Rønne*, Modeling auditory evoked potentials to complex stimuli, 2012.

Vol. 13: *Claus Forup Corlin Jespersgaard*, Listening in adverse conditions: Masking release and effects of hearing loss, 2012.

Vol. 14: *Rémi Decorsière*, Spectrogram inversion and potential applications for hearing research, 2013.

Vol. 15: *Søren Jørgensen*, Modeling speech intelligibility based on the signal-to-noise envelope power ration, 2014.

Vol. 16: *Kasper Eskelund*, Electrophysiological assessment of audiovisual integration in speech perception, 2014.

Vol. 17: *Simon Krogholt Christiansen*, The role of temporal coherence in auditory stream segregation, 2014.

Vol. 18: *Márton Marschall*, Capturing and reproducing realistic acoustic scenes for hearing research, 2014.

Vol. 19: *Jasmina Catic*, Human sound externalization in reverberant environments, 2014.

Vol. 20: *Michał Fereczkowski*, Design and evaluation of individualized hearing-aid signal processing and fitting, 2015.

Vol. 21: *Alexandre Chabot-Leclerc*, Computational modeling of speech intelligibility in adverse conditions, 2015.

Vol. 22: *Federica Bianchi*, Pitch representations in the impaired auditory system and implications for music perception, 2016.

Vol. 23: *Johannes Zaar*, Measures and computational models of microscopic speech perception, 2016.

Vol. 24: *Johannes Käsbach*, Characterizing apparent source width perception, 2016.

Vol. 25: *Gusztáv Löcsei*, Lateralized speech perception with normal and impaired hearing, 2016.

Vol. 26: *Suyash Narendra Joshi*, Modelling auditory nerve responses to electrical stimulation, 2017.

Vol. 27: *Henrik Gerd Hassager*, Characterizing perceptual externalization in listeners with normal, impaired and aided-impaired hearing, 2017.

Vol. 28: *Richard Ian McWalter*, Analysis of the auditory system via synthesis of natural sounds, speech and music, 2017.

Vol. 29: *Jens Cubick*, Characterizing the auditory cues for the processing and perception of spatial sounds, 2017.

Vol. 30: *Gerard Encina-Llamas*, Characterizing cochlear hearing impairment using advanced electrophysiological methods, 2017.

Vol. 31: *Christoph Scheidiger*, Assessing speech intelligibility in hearing-impaired listeners, 2018.

Vol. 32: *Alan Wiinberg*, Perceptual effects of non-linear hearing aid amplification strategies, 2018.

Vol. 33: *Thomas Bentsen*, Computational speech segregation inspired by principles of auditory processing, 2018.

Vol. 34: *François Guérit*, Temporal charge interactions in cochlear implant listeners, 2018.

Vol. 35: *Andreu Paredes Gallardo*, Behavioral and objective measures of stream segregation in cochlear implant users, 2018.

The end.

To be continued…

This thesis investigated the polarity-specific temporal integration of currents by cochlear implant listeners. This was achieved by measuring the loudness, detection thresholds and localization abilities of the listeners with pulse pairs, while varying the inter-pulse interval and the polarity of each pulse.

Overall, results showed a variety of inter-pulse and polarity effects, as well as interactions between both factors. These results are not only relevant for clinical applications, such as the estimation of neural survival in the auditory nerve, but also for the development of models of the electrically activated auditory nerve.

DTU Electrical Engineering
Department of Electrical Engineering

Ørsteds Plads
Building 348
DK-2800 Kgs. Lyngby
Denmark
Tel: (+45) 45 25 38 00
Fax: (+45) 45 93 16 34
www.elektro.dtu.dk