

AUDIOVISUAL SPEECH PERCEPTION IN
COCHLEAR IMPLANT PATIENTS

Inaugural Dissertation

zur

Erlangung des Doktorgrades
philosophiae doctor (PhD) in Health Sciences
der Medizinischen Fakultät
der Universität zu Köln

vorgelegt von

Natalie Layer

aus Nischnij Tagil (Russland)

Comouth - Druckerei & Buchbinderei, Aachen

2023

Betreuer*in:

PD Dr. Pascale Sandmann

Gutachter*in:

PD Dr. Jürgen-Christoph von Kleist-Retzow

Prof. Dr. Silvia Daun

Datum der Mündlichen Prüfung: 09.03.2023

ACKNOWLEDGEMENTS

This thesis and the included publications would not have been possible without the contribution of numerous people who supported me throughout my PhD-time. There were both joyful and difficult moments during the last three and a half years, that were heavily influenced by the Corona pandemic, and I am very grateful that so many people were there for me, both for professional and mental support. There are many people I would like to thank, but there are a few in particular for whom I would like to express my deepest gratitude.

First and foremost, I would like to thank my supervisor, **Pascale Sandmann**, who guided me through all stages of my PhD. Despite her many other obligations (especially in the clinic), she always did her best to provide the best possible supervision, and she took me and my projects very seriously. I believe that a supervisor like her is extremely rare, which is why I am grateful to have been part of her research group. The helpful discussions, the personal conversations, her patience, knowledge, empathy, and encouraging words contributed significantly to my personal and professional development. I wish her the best of luck, success, and happiness in her next career step in Oldenburg.

I would also like to thank my colleagues and friends from my research group for always being there for me, for interesting discussions (both personal and professional), for their endless time and motivation to participate as pilot participants and listen to my problems, their advice, proofreading, and all the other things that young researchers do. **Anna Weglage** began her PhD at the same time as me, and we constantly supported and motivated each other, as we progressed through the stages of our PhDs, which is great. Besides, she was the best room mate at every conference we attended. **Pauline Burkhardt** was always extremely helpful, because she was the first participant to test our experiments, and she always provided the best insights, as she knows exactly what it is like to be a CI user. I really admire her for achieving such academic success despite the challenges of her hearing impairment. **Lina Wiesel** always created a funny and enjoyable atmosphere with her nice manner. It's a pity that we had to hand her over to the audio team later on. **Annabell Urban** has only recently joined our team, but we really have grown fond of her over the last few months. I think we are a very special team that supports each other in good and bad times, we always know we can count on each other, and we benefit greatly from our diverse backgrounds. I wish you the best of luck with finishing your degrees.

A special thanks goes to **Verena Müller** and the other CI-technicians at the CIK for always being available to me when I had technical CI-related questions, for assisting in the recruitment of suitable CI participants, for checking the CI-settings of the CI users and adjusting them to my needs, and for simply sharing stories about interesting patients with me.

Of course, I would also like to thank all of the other members of the CIK and the audiology team, particularly **Ruth Lang-Roth**, **Martin Walger** and **Irina Schierholz**,

who gave me helpful advice, listened to many of my talks, explained things to me, were always available when I had clinical questions, and made me feel valuable as a PhD student in this team. Furthermore, they provided a clinical perspective and allowed me to participate in the *Newborn Hearing Screening*.

Further, I would like to thank my colleagues at the Jean-Uhrmacher Institute, **Hartmut Meister**, **Moritz Wächtler** and **Khaled Abdel-Latif** with whom we met on a regular basis to have many insightful discussions. We showed much interest in each other's projects, and we always offered advice, technical assistance, or participated in experiments to help each other advance our projects.

Aside from those already mentioned, I would like to thank all of my co-authors. Thanks to **Micah Murray** and **Jan-Ole Radecke** for their help, motivation, ideas and feedback that highly improved the quality of our papers.

I would also like to thank my IPHS programme tutors, **Silvia Daun** and **Jürgen-Christoph von Kleist-Retzow** for giving advice and providing helpful comments for improving my projects. Speaking about the IPHS, I would like to thank the IPHS programme coordinators for their untiring assistance in answering my thousands of questions and providing excellent guidance throughout my PhD.

Of course, I would like to thank all the **CI and NH participants** who took part in my experiments. None of the studies would have been possible without their interest, patience, time, and motivation.

A great thanks also goes to my **friends and my in-laws** (Layer, Unruh, Gropp), who have always been supportive, proud of me, friendly, loving, and welcoming. I am happy to be a member of the Layer family.

Finally, I would like to thank my family (Leschinski, Wegner, Katjukov, Friedmann), especially my father **Andreas Leschinski**, my mother **Marina Leschinski**, and my sister **Nicole Leschinski**. Vielen Dank, dass Ihr mich immer unterstützt, motiviert und mir so oft gezeigt habt, dass Ihr stolz auf mich seid. Ich kann mich glücklich schätzen, Euch zu haben.

Above all, I would like to thank my husband **Moritz Layer**, who is simply the best husband on earth. Without him, I would probably starve or die of thirst when I work super concentrated and forget to look after myself. He supports me so much, believes in me, is always there for me, helps me with challenging programming issues, did a lot of proof reading, he always puts up with my temper and appreciates my quirks, and he understands me like nobody else. Thank you so much! I love you.

SUMMARY

Hearing with a cochlear implant (CI) is very different compared to a normal-hearing (NH) experience, as the CI can only provide limited auditory input. Nevertheless, the central auditory system is capable of learning how to interpret such limited auditory input such that it can extract meaningful information within a few months after implant switch-on. The capacity of the auditory cortex to adapt to new auditory stimuli is an example of *intra-modal* plasticity — changes within a sensory cortical region as a result of altered statistics of the respective sensory input. However, hearing deprivation before implantation and restoration of hearing capacities after implantation can also induce *cross-modal* plasticity — changes within a sensory cortical region as a result of altered statistics of a different sensory input. Thereby, a preserved cortical region can, for example, support a deprived cortical region, as in the case of CI users which have been shown to exhibit cross-modal visual-cortex activation for purely auditory stimuli. Before implantation, during the period of hearing deprivation, CI users typically rely on additional visual cues like lip-movements for understanding speech. Therefore, it has been suggested that CI users show a pronounced binding of the auditory and visual systems, which may allow them to integrate auditory and visual speech information more efficiently.

The projects included in this thesis investigate auditory, and particularly audiovisual speech processing in CI users. Four event-related potential (ERP) studies approach the matter from different perspectives, each with a distinct focus.

The first project investigates how audiovisually presented syllables are processed by CI users with bilateral hearing loss compared to NH controls. Previous ERP studies employing non-linguistic stimuli and studies using different neuroimaging techniques found distinct audiovisual interactions in CI users. However, the precise timecourse of cross-modal visual-cortex recruitment and enhanced audiovisual interaction for speech related stimuli is unknown. With our ERP study we fill this gap, and we present differences in the timecourse of audiovisual interactions as well as in cortical source configurations between CI users and NH controls.

The second study focuses on auditory processing in single-sided deaf (SSD) CI users. SSD CI patients experience a maximally asymmetric hearing condition, as they have a CI on one ear and a contralateral NH ear. Despite the intact ear, several behavioural studies have demonstrated a variety of beneficial effects of restoring binaural hearing, but there are only few ERP studies which investigate auditory processing in SSD CI users. Our study investigates whether the side of implantation affects auditory processing and whether auditory processing via the NH ear of SSD CI users works similarly as in NH controls.

Given the distinct hearing conditions of SSD CI users, the question arises whether there are any quantifiable differences between CI user with *unilateral* hearing loss and *bilateral* hearing loss. In general, ERP studies on SSD CI users are rather scarce, and there is no study on audiovisual processing in particular. Furthermore, there are

no reports on lip-reading abilities of SSD CI users. To this end, in the third project we extend the first study by including SSD CI users as a third experimental group. The study discusses both differences and similarities between CI users with bilateral hearing loss and CI users with unilateral hearing loss as well as NH controls and provides — for the first time — insights into audiovisual interactions in SSD CI users.

The fourth project investigates the influence of background noise on audiovisual interactions in CI users and whether a noise-reduction algorithm can modulate these interactions. It is known that in environments with competing background noise listeners generally rely more strongly on visual cues for understanding speech and that such situations are particularly difficult for CI users. As shown in previous auditory behavioural studies, the recently introduced noise-reduction algorithm "ForwardFocus" can be a useful aid in such cases. However, the questions whether employing the algorithm is beneficial in audiovisual conditions as well and whether using the algorithm has a measurable effect on cortical processing have not been investigated yet. In this ERP study, we address these questions with an auditory and audiovisual syllable discrimination task.

Taken together, the projects included in this thesis contribute to a better understanding of auditory and especially audiovisual speech processing in CI users, revealing distinct processing strategies employed to overcome the limited input provided by a CI. The results have clinical implications, as they suggest that clinical hearing assessments, which are currently purely auditory, should be extended to audiovisual assessments. Furthermore, they imply that rehabilitation including audiovisual training methods may be beneficial for all CI user groups for quickly achieving the most effective CI implantation outcome.

ZUSAMMENFASSUNG

Das Hören mit einem Cochlea-Implantat (CI) unterscheidet sich stark von dem Höreindruck eines Normalhörenden (NH), da das CI nur einen begrenzten auditiven Input liefern kann. Dennoch kann das zentrale auditorische System lernen, diesen begrenzten Input als sinnvolle Information zu interpretieren. Dies gelingt bereits innerhalb weniger Monate nach dem Einsetzen des Implantats. Die Fähigkeit des auditorischen Kortex, sich an neue auditorische Reize anzupassen, ist ein Beispiel für *intra-modale* Plastizität — Veränderungen innerhalb einer sensorischen kortikalen Region als Folge einer veränderten Inputstatistik des entsprechenden Sinnes. Das Fehlen von auditorischen Reizen vor der Implantation und die Wiederherstellung des Hörvermögens nach der Implantation können jedoch auch zu *cross-modaler* Plastizität führen — Veränderungen innerhalb einer sensorischen kortikalen Region als Folge veränderter Inputstatistik eines anderen Sinnes. Zum Beispiel kann eine sensorisch deprivierte kortikale Region von einer anderen Region in ihrer Funktion unterstützt werden, wie im Fall von CI-Trägern, die eine cross-modale Aktivierung des visuellen Kortex für rein auditive Reize aufweisen. Vor der Implantation, wenn das Hören stark beeinträchtigt ist, sind CI-Träger in der Regel auf zusätzliche visuelle Inputs, wie Lippenbewegungen, angewiesen, um Sprache zu verstehen. Daher wird vermutet, dass CI-Träger eine ausgeprägte Interaktion zwischen dem auditiven und dem visuellen System aufweisen, die es ihnen ermöglicht, auditive und visuelle Sprachinformationen effizienter zu integrieren.

Die in dieser Arbeit präsentierten Projekte untersuchen die auditive und insbesondere die audiovisuelle Sprachverarbeitung bei CI-Trägern. Vier Studien zu ereigniskorrelierten Potenzialen (EKPs) betrachten das Thema aus unterschiedlichen Perspektiven, jede mit einem anderen Schwerpunkt.

Im ersten Projekt wird untersucht, wie audiovisuell präsentierte Silben von CI-Trägern mit beidseitigem Hörverlust im Vergleich zur NH-Gruppe verarbeitet werden. Frühere EKP-Studien mit nicht-sprachlichen Stimuli und Studien, die andere Messverfahren angewendet haben, zeigen deutliche audiovisuelle Interaktionen bei CI-Trägern. Der genaue zeitliche Verlauf der cross-modalen Rekrutierung des visuellen Kortex und der verstärkten audiovisuellen Interaktion bei sprachlichen Reizen ist jedoch bisher nicht bekannt. Mit unserer EKP-Studie schließen wir diese Lücke und beschreiben die Unterschiede im zeitlichen Verlauf der audiovisuellen Interaktionen, sowie in den kortikalen Quellkonfigurationen, zwischen CI-Trägern und der NH-Kontrollgruppe.

Die zweite Studie befasst sich mit der auditorischen Verarbeitung bei einseitig ertaubten (single-sided deaf; SSD) CI-Trägern. SSD CI-Patienten erleben einen maximal asymmetrischen Höreindruck, da sie ein CI auf einem Ohr und ein kontralaterales NH-Ohr haben. Trotz des intakten Ohres haben mehrere Verhaltensstudien eine Reihe von positiven Auswirkungen der Wiederherstellung des binauralen Hörens gezeigt. Es gibt aber nur wenige EKP-Studien, die die auditorische Verarbeitung bei SSD CI-Trägern analysiert haben. Unsere Studie untersucht, ob die Implantationsseite die

Hörverarbeitung beeinflusst und ob die Hörverarbeitung über das NH-Ohr bei SSD CI-Trägern ähnlich funktioniert wie bei Normalhörenden.

Angesichts der besonderen Hörsituation von SSD CI-Trägern stellt sich die Frage, ob es quantifizierbare Unterschiede zwischen CI-Trägern mit *einseitigem* und *beidseitigem* Hörverlust gibt. Im Allgemeinen gibt es nur wenige EKP-Studien mit SSD CI-Trägern und insbesondere gibt es keine Studie zur audiovisuellen Verarbeitung. Außerdem gibt es keine uns bekannten Studien zur Lippenlesefähigkeit von SSD CI-Trägern. Daher erweitern wir im dritten Projekt unsere erste Studie um SSD CI-Träger als dritte Versuchsgruppe. Die Studie diskutiert sowohl Unterschiede als auch Gemeinsamkeiten zwischen CI-Trägern mit beidseitigem Hörverlust und CI-Trägern mit einseitigem Hörverlust sowie NH-Kontrollen und liefert erstmalig Einblicke in audiovisuelle Interaktionen bei SSD CI-Trägern.

Das vierte Projekt untersucht den Einfluss von Störgeräuschen auf audiovisuelle Interaktionen bei CI-Trägern und ob ein Algorithmus zur Störgeräuschreduktion diese Interaktionen modulieren kann. Es ist bekannt, dass sich Menschen in Umgebungen mit zusätzlichen Hintergrundgeräuschen generell stärker auf visuelle Informationen verlassen, um Sprache besser zu verstehen und dass solche Situationen für CI-Träger besonders schwierig sind. Wie in früheren auditiven Verhaltensstudien gezeigt wurde, kann der kürzlich eingeführte Störgeräuschunterdrückungsalgorithmus "ForwardFocus" in solchen Fällen das Sprachverständnis erleichtern. Die Fragen, ob die Anwendung des Algorithmus auch unter audiovisuellen Bedingungen vorteilhaft ist und ob die Anwendung des Algorithmus einen messbaren Effekt auf die kortikale Verarbeitung hat, sind jedoch bisher noch nicht untersucht worden. In der hier präsentierten EKP-Studie beantworten wir diese Fragen mit Hilfe eines auditiven und audiovisuellen Silbenunterscheidungsparadigmas.

Insgesamt tragen die Projekte dieser Arbeit zu einem besseren Verständnis der auditiven und insbesondere der audiovisuellen Sprachverarbeitung bei CI-Trägern bei und zeigen unterschiedliche kortikale Verarbeitungsstrategien auf, um den begrenzten Input eines CIs zu kompensieren. Die Ergebnisse sind relevant für den klinischen Alltag, da sie nahelegen, dass klinische Hörmessverfahren, die derzeit rein auditiv sind, durch audiovisuelle Messverfahren erweitert werden sollten. Darüber hinaus deuten sie darauf hin, dass eine Rehabilitation, die audiovisuelle Trainingsmethoden einschließt, für alle CI-Nutzergruppen von Vorteil sein kann, um schnell ein möglichst effektives Ergebnis der CI-Implantation zu erzielen.

CONTENTS

Acknowledgements	iii
Summary	iv
Zusammenfassung	vii
1 Introduction	1
1.1 Hearing disorders	3
1.2 Hearing with a cochlear implant	7
1.3 Neuroplasticity in cochlear implant users	11
1.4 Audiovisual interactions	15
1.5 Difference between cochlear implant user groups	16
1.6 The influence of background noise	17
1.7 Methods	19
1.7.1 Electroencephalography	19
1.7.2 Auditory event-related potentials	22
1.7.3 ERP data analysis	24
1.8 Aim of this thesis	27
2 "The timecourse of multisensory speech processing in unilaterally stimulated cochlear implant users revealed by ERPs"	31
Author contributions	31
Synopsis	31
Publication	33
Supplementary material	51
3 "Side-of-implantation effect on functional asymmetry in the auditory cortex of single-sided deaf cochlear-implant users"	53
Author contributions	53
Synopsis	53
Publication	55
Supplementary material	78
4 "Electrophysiological differences and similarities in audiovisual speech processing in cochlear implant users with unilateral and bilateral hearing loss"	79
Author contributions	79
Synopsis	79
Publication	81
Supplementary material	99
5 "The influence of background noise and noise reduction by 'ForwardFocus' on audiovisual speech perception in bimodal CI users: an ERP study"	101

Author contributions	101
Synopsis	101
Submitted manuscript	103
6 Discussion	143
6.1 Summary	143
6.2 Altered auditory processing in CI users with unilateral hearing loss	143
6.3 Altered audiovisual interactions in CI users with bilateral hearing loss .	146
6.4 Altered audiovisual interactions in different CI user groups	149
6.5 Background noise influences audiovisual interactions in CI users	151
6.6 Outlook	153
6.7 Conclusion	153
Bibliography	155
Colophon	xi
Eidesstattliche Erklärung	xiii

INTRODUCTION

Hearing loss is a very common issue that many people have to deal with. According to the world report on hearing (World Health Organization, 2021), there are more than 1.5 billion people worldwide who undergo some loss of their hearing capacity at some point in their lives, with at least 430 million requiring treatment. Everyone knows someone, most likely a parent, grandparent, aunt or uncle, who clearly needs a hearing aid but refuses to wear one because he or she claims to hear well enough. As a result, the relatives and friends must speak more clearly, repeat what they said, and have to suffer from a television that runs on maximum volume. Unfortunately, in a conversation between several NH persons and a hearing impaired person, it is easy to forget about the hearing impairment. This is dangerous because a hearing impaired person will quickly lose track in a group conversation. No longer being able to follow conversations can force hearing impaired people into isolation, and it has been shown to increase the risk of mental disorders, cognitive impairments, and even dementia (Fellinger et al., 2007; Lin et al., 2011; Gates et al., 2011; Lin et al., 2013; Humes et al., 2013). This has, of course, a tremendous impact on the quality of life of an individual with hearing impairment (Nordvik et al., 2018; Ciorba et al., 2012). To prevent these adverse effects, it is critical to raise awareness of the importance of seeking professional help when hearing abilities start to decrease or are already impaired.

Up to now, there exist various techniques for detecting hearing impairments, including both subjective and objective tools. Subjective tests require an overt response from the patient, such as in tone audiometry, where the patient must press a button as soon as a tone is heard, or in speech audiometry, where a word is presented at a certain intensity level and the patient is required to repeat the word. Objective measures, on the other hand, such as brainstem evoked response audiometry (BERA), do not require a behavioural response. The clinical professional records auditory brainstem responses using electrodes on specific positions on a patient's head while acoustic stimuli are played. As a result, the professional measures automatic responses, which can be evaluated. However, current clinical procedures for detecting hearing impairments are restricted to purely auditory measures.

If a hearing impairment is detected, there are multiple technical solutions that can be used to improve hearing, such as a hearing aid (for mild-to-moderate and sometimes severe hearing impairment) or, when a hearing aid is insufficient, a cochlear implant (for severe-to-profound hearing impairment). Very simply put, hearing aids amplify the sounds, whereas a CI is an auditory prosthesis that can (partially) restore the hearing of people with severe to profound sensorineural hearing loss through

direct electrical stimulation of the auditory nerve (Zeng et al., 2011; House and Urban, 1973; Simmons et al., 1965). A CI currently is the only clinically established prosthesis capable of restoring a human sense, and, if certain conditions are met, a CI even allows congenitally deaf children to hear (Macherey and Carlyon, 2014). The implantation of a CI now is a standard clinical procedure, with approximately 750,000 patients who have received this prosthesis worldwide (Cullington et al., 2022) and around 50,000 individuals in Germany (Lenarz, 2018). Because CI users are the experimental patient group this thesis focuses on, more detailed information on how a CI exactly works is provided later in this thesis.

Aside from the importance of timely supply with a hearing system, achieving the best possible outcome with the hearing system is critical. Due to the limited sound frequency resolution of CIs, speech comprehension can be hampered, although the subjective difficulty of understanding speech varies from patient to patient. The success of a CI implantation highly depends on how willing patients are to train their hearing with their new prosthesis. This requires a significant amount of effort, patience, and an extensive rehabilitation in order to gain the largest possible benefit with a CI (Diller, 2009). Despite intensive hearing and speech training during the rehabilitation phase, speech comprehension varies greatly between individual patients and is dependent on numerous factors. Among those are surgical factors (e.g. electrode placement), auditory deprivation-induced reorganisation of the brain (before implantation), and the ability to interpret the limited auditory CI-input (after implantation) (Lazard et al., 2012a,b,c). It is crucial to investigate all factors that influence the outcome of an implantation to be able to better predict the success of an implantation, and to develop better interventions by offering tailored therapeutic solutions.

Apart from researching the factors influencing rehabilitation outcomes, it is critical to understand how hearing with a CI can be optimised, for example by focusing on *audiovisual* listening situations that better reflect a CI user's everyday life situation in contrast to auditory-only situations. There is accumulating evidence that CI users are influenced more strongly by the visual aspect of speech (Rouger et al., 2007; Desai et al., 2008; Strelnikov et al., 2013), such as lip-movements and that this influence is reflected in alterations in auditory and audiovisual speech processing (Giraud et al., 2001c; Chen et al., 2016; Schierholz et al., 2015; Radecke et al., 2022). Given these remarkable CI-user specific effects, the development of audiovisual procedures may allow for better and more realistic hearing assessments in the clinic, or training set-ups during the rehabilitation phase. Particularly in difficult hearing situations (background noise, concurrent speakers), it is important to recognise the benefits of audiovisual information and to develop beneficial technical optimisations, such as improved algorithms to process incoming speech in noisy environments. Taken together, a lot has been done so far to obtain satisfactory hearing results and to understand speech processing in CI users, but there is still much room for improvements and unanswered questions to be solved.

The present chapter provides all the basic information required to understand the single projects which are part of this thesis. It covers general information about CIs, what happens in the brain of an implanted patient, what kinds of problems these individuals face in difficult hearing situations, and whether there is a difference

between distinct groups of implanted patients. Moreover, this chapter introduces the issue of audiovisual speech processing, which is the focus of this thesis, as well as the significance of audiovisual information for CI users. Finally, an overview of the analytical methods used in this thesis' projects will be provided.

1.1 HEARING DISORDERS

The current world report on hearing of the World Health Organisation (WHO; World Health Organization, 2021) announced that over 1.5 billion people currently have some degree of hearing loss (approximately 20 % of the global population), with that number expected to rise to 2.5 billion by 2050. According to current estimates, more than 42 % of people with any degree of hearing loss are over the age of 60, making age-related hearing loss (presbycusis) one of the most prevalent cause for hearing loss across lifetime, and the number is predicted to rise due to increasing life expectancy.

A person is considered as being affected by hearing loss when the hearing ability is reduced, making it impossible to hear as well as someone with normal hearing abilities (World Health Organization, 2021). The WHO has introduced a grading system based on tone audiometry to standardise the way hearing loss severity is categorised (Humes, 2019), for the first time in 1986 (Olusanya et al., 2019). For this purpose, the values of the tone audiometry at 500, 1000, 2000, and 4000 Hz are averaged in the better ear to obtain a hearing threshold. The grading system has been updated several times, and the current grading is shown in Table 1.1. Hearing thresholds of 20 dB or better in both ears are considered "normal" (World Health Organization, 2021). Depending on the severity of hearing loss, those with a hearing threshold above 20 dB may be classified as "hard of hearing" or "deaf." The term "hard of hearing" refers to people who have mild to severe hearing loss, whereas the term "deaf" refers to people who experience severe or profound hearing loss in both ears and can only hear at very loud intensities or absolutely nothing (World Health Organization, 2021). To get an idea of the intensity levels in daily life, Figure 1.1 illustrates some typical sounds and their corresponding intensity level. The curved area is often referred to as the "speech banana", because of its shape, and it depicts the approximate frequency and intensity range in which conversations in quiet environments take place (Ross, 2004). The most important speech sounds are produced within this range. Therefore, it is reasonable that if a person already has some moderate hearing loss, with a hearing threshold between 35 and 50 dB, understanding some speech sounds at a normal conversational level can be challenging (compare Figure 1.1). As a result, the primary goal is to choose hearing supplies that are as well-fitted as possible to fall within the range of the speech banana. Before discussing the different types and the origins of hearing loss, the anatomy of the ear and the process of normal hearing are described.

HEARING IN HEALTHY INDIVIDUALS The following description of the anatomy and physiology of the ear is based on chapter 7 of the book *Basic Otorhinolaryngology* by Probst et al. (2018). For the purpose of this thesis, only the fundamentals necessary to comprehend how a CI works (which will be explained later in the present thesis) are presented.

Grade	Hearing threshold in better hearing ear in decibels (dB)	Hearing experience in a quiet environment for most adults	Hearing experience in a noisy environment for most adults
Normal hearing	Less than 20 dB	No problem hearing sounds	No or minimal problem hearing sounds
Mild hearing loss	20 to < 35 dB	Does not have problems hearing conversational speech	May have difficulty hearing conversational speech
Moderate hearing loss	35 to < 50 dB	May have difficulty hearing conversational speech	Difficulty hearing and taking part in conversation
Moderately severe hearing loss	50 to < 65 dB	Difficulty hearing conversational speech; can hear raised voices without difficulty	Difficulty hearing most speech and taking part in conversation
Severe hearing loss	65 to < 80 dB	Does not hear most conversational speech; may have difficulty hearing and understanding raised voices	Extreme difficulty hearing speech and taking part in conversation
Profound hearing loss	80 to < 95 dB	Extreme difficulty hearing raised voices	Conversational speech cannot be heard
Complete or total hearing loss/ deafness	95 dB or greater	Cannot hear speech and most environmental sounds	Cannot hear speech and most environmental sounds
Unilateral	< 20 dB in the better ear, 35 dB or greater in the worse ear	May not have problem unless sound is near the poorer hearing ear. May have difficulty in locating sounds	May have difficulty hearing speech and taking part in conversation, and in locating sounds

Table 1.1: The current WHO grading system of hearing loss severity (adapted from the world report on hearing of the WHO (World Health Organization, 2021)). This table provides the categorisation (grade) according to the range of the hearing threshold in the better ear and the hearing experience for situations in quiet and noisy environments based the underlying threshold.

The auditory system consists of three parts: 1) the external ear, 2) the middle ear, and 3) the inner ear (see Figure 1.2). The external ear is composed of the ear conch (pinna) and a narrow passage called the ear canal. The external ear serves as an acoustic antenna, guiding sound waves to the sensitive middle ear structures. Additionally, the pinna and ear canal work like an acoustic funnel that amplifies specific frequency bands. The transmitted sound waves cause the eardrum (the part of the ear that connects the outer and middle ear) to vibrate. Other important parts of the middle ear beside the eardrum are the tympanic cavity which is mechanically connected to the inner ear by a group of auditory ossicles (Malleus, Incus, Stapes) through the round and oval windows. The eardrum causes the ossicles to vibrate, allowing sounds to be amplified and transmitted to the inner ear. The most important part of the inner ear is the cochlea, which is the actual sensory organ for hearing and which includes the basilar membrane and the hair cells. As soon as vibrations from the

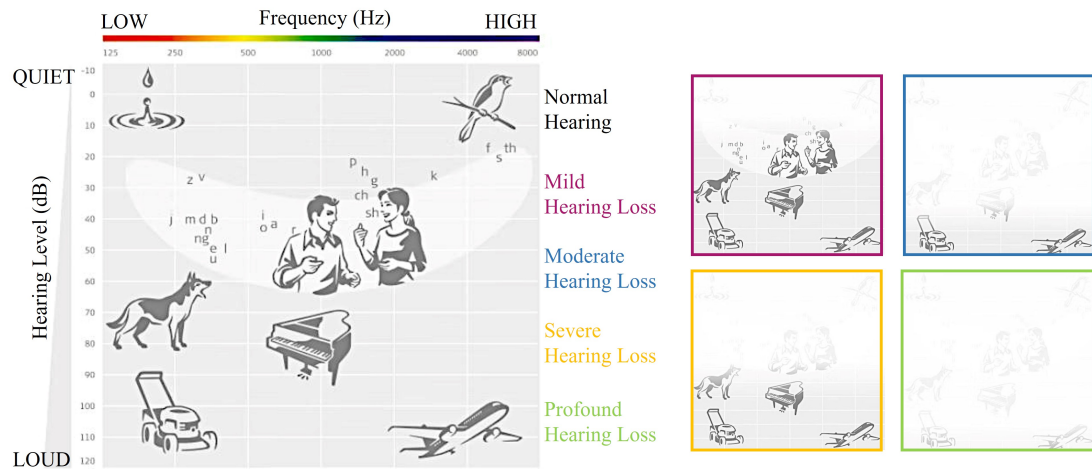


Figure 1.1: Frequency and intensity range of daily life sounds. A typical audiogram is shown on the left, with the frequency depicted on the horizontal axis, from low frequencies on the left (125 Hz) to high frequencies on the right (8,000 Hz). The hearing level is represented vertically, with low intensities at the top (-10 dB; quiet) and high intensities at the bottom (120 dB; loud). The curved area approximates the range where most conversations typically take place. The levels of hearing loss are shown on the right side of the audiogram, and the corresponding boxes with the same audiogram (in the same colour) visualise how hearing is like at different stages of hearing loss. Adapted with permission from MED-EL, Copyright © 2022 MED-EL Medical Electronics.

middle ear reach the cochlea, a travelling wave is formed along the basilar membrane. The frequency sensitive hair cells are tuned to lower frequencies in the apical region and to higher frequencies in the cochlea's basal region (tonotopic organisation, see Figure 1.3). Thus, the cochlea's primary function is to convert auditory signals into neural impulses that exactly encode the timing and the nature of the incoming sound. This is accomplished by splitting the broad frequency spectrum into narrow frequency bands. The cochlea transfers the impulses to the auditory nerve which is connected to the brain, where the neuronal signal is interpreted as a meaningful sound pattern, such as speech. However, before reaching the auditory cortex, the neuronal activity passes several stations of the auditory pathway, encompassing multiple parts of the brainstem, midbrain and the thalamus (Peterson et al., 2018).

TYPES OF HEARING LOSS Peripheral hearing loss is usually classified as either conductive, sensorineural or mixed. Mixed hearing loss occurs when a person suffers from both conductive and sensorineural hearing loss (Cunningham and Tucci, 2017). Conductive hearing loss results from damage to the outer or middle ear, whereas sensorineural hearing loss is caused by impairment of the inner ear (cochlea) or further along the auditory processing pathway. More specifically, sensorineural hearing loss can be further subdivided into 1) sensory hearing loss, which is caused within the cochlea, 2) neural hearing loss, which is caused by cochlear nerve (or auditory nerve)

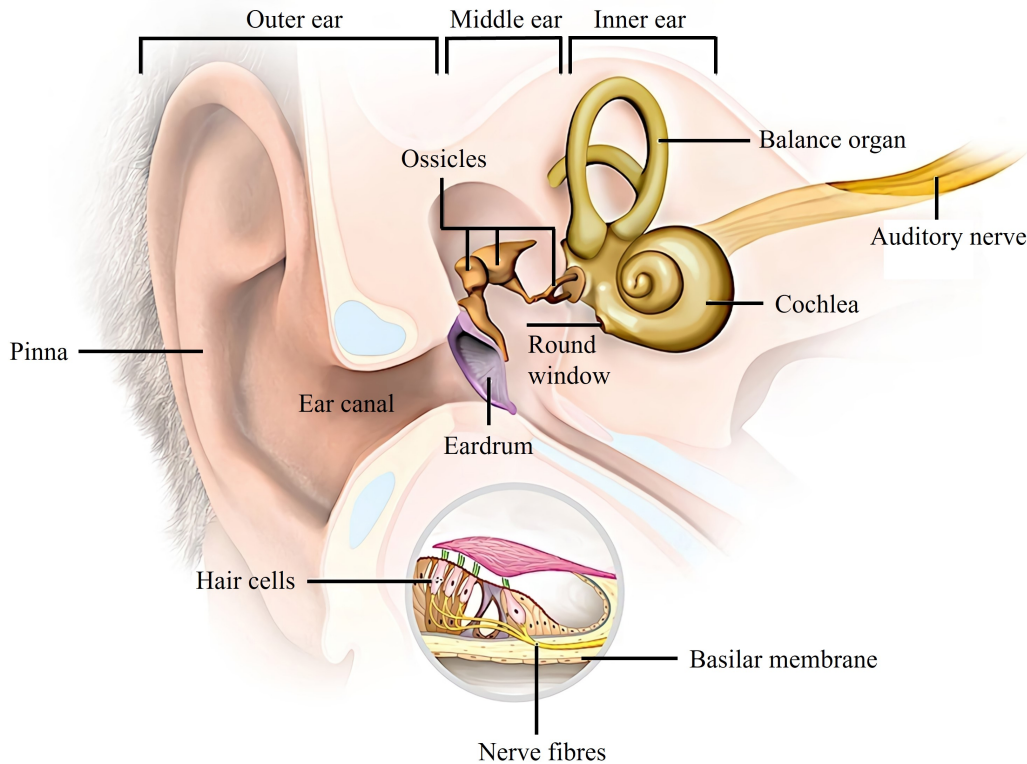


Figure 1.2: Basic anatomy of a normal-hearing ear. Adapted with permission from MED-EL, Copyright © 2022 MED-EL Medical Electronics.

dysfunctions, and 3) central hearing loss, which results from dysfunctions of the central auditory pathway or the auditory cortex (Zahnert, 2011).

As the name *conductive hearing loss* already implies, the conduction of sound to the inner is disrupted, either at the outer ear, the middle ear, or both. Ear wax (Cerumen) that is blocking the ear canal is a common cause of conductive hearing loss. Further typical reasons are infections of the auditory canal (Otitis externa), a middle ear infection (Otitis media), which often is the reason for conductive hearing loss in children, a middle ear effusion (accumulating fluid in the middle ear and eustachian tube, putting strain on the tympanic membrane; Searight et al. (2022)), a cholesteatoma (accumulations of keratinised squamous epithelium in the middle ear; Kennedy and Singh (2017)), otosclerosis (abnormal bone formation), or - very rarely - a tumour (Isaacson and Vora, 2003). Hearing aids can often help in conductive hearing loss if the hearing loss is not too severe. However, for people with severe to profound hearing loss, CIs are the only way to regain meaningful auditory information (Macherey and Carlyon, 2014).

Sensory hearing loss, which originates in the inner ear, is frequently caused by the degeneration or loss of sensory hair cells. Consequently, mechanical sound waves can no longer be converted into electrical signals (Müller, 2005). The most common causes of sensory hearing loss are presbycusis (age-related hearing loss), heredity (genetic mutations), extensive noise exposure, therapeutic drugs with ototoxic side effects, sudden hearing loss, and chronic conditions (Cunningham and Tucci, 2017). A

CI is the optimal hearing system in the case of sensory hearing loss or when patients do not benefit enough from a conventional hearing aid (Deep et al., 2019). In the case of congenitally deaf infants, CI implantation can aid in the acquisition of spoken language, with the best results obtained when the CI is implanted between the ages of 1 and 2 years (Tajudeen et al., 2010). An important prerequisite for enabling a CI implantation is that the auditory nerve is intact, which can be ensured by using high-resolution computed tomography (CT) and magnetic resonance tomography (MRT) (Deep et al., 2019; Dazert et al., 2020).

Neural hearing loss occurs when the auditory nerve is damaged or absent. It is frequently caused by tumours or infections of the petrous bone. *Central hearing loss* is diagnosed when further parts of the auditory pathway or the auditory cortex are affected, such that nerve signals arriving from the inner ear can no longer be correctly interpreted. Trauma, inflammations, infections, infarction, and space-occupying lesions (tumours, haemorrhage) can cause central hearing loss in adults (Zahnert, 2011). Thus, a CI cannot be used in patients with neural or central hearing loss and hearing rehabilitation becomes very demanding. A possible solution to restore at least some hearing sensations would be to use an auditory brainstem implant (ABI), which is a prosthesis that can electrically stimulate the cochlear nucleus in the brainstem (Colletti et al., 2012). Around the world, there are approximately a thousand deaf adults and children who have received an ABI so far (Wong et al., 2019). There is also a potential solution for patients who are unable to benefit from a CI or an ABI and have hearing impairments located within more central auditory structures. This solution is an auditory midbrain implant (AMI), which is a prosthesis that stimulates the inferior colliculus within the midbrain. AMIs, however, are still being developed, and only very few people have received it so far (Lim and Lenarz, 2015).

1.2 HEARING WITH A COCHLEAR IMPLANT

WHAT IS A COCHLEAR IMPLANT AND HOW DOES IT WORK? The CI is a small bionic prosthesis that compensates for damaged or missing cochlear hair cells in the inner ear by direct electrical stimulation of the auditory nerve (Carlson, 2020). Hence, an intact auditory nerve is a crucial prerequisite to make CIs work (Lenarz, 2018). A CI currently is the only clinically established prosthesis which can substitute a human sense (Macherey and Carlyon, 2014). Even though the first direct stimulation of the auditory nerve dates back to the year 1937, it was in the year 1957 when the notation "cochlear implant" was used for the first time. A single copper wire was inserted inside the cochlea of a man, which was then stimulated (DeSaSouza, 2022; Wilson and Dorman, 2008). A few years later, in 1961 the Americans William House and John Doyle implanted the first CI, turning CIs into a clinical reality (Deep et al., 2019; Wilson and Dorman, 2008). Approximately 20 years ago, the CI was a single-electrode device used primarily to improve lip-reading and to provide sound awareness (Zeng, 2004). Now, CI implantation is a standard clinical surgery, with approximately 750,000 patients worldwide that have received this prosthesis (Cullington et al., 2022), including approximately 50,000 in Germany (Lenarz, 2018).

The essential parts of a CI system are shown in Figure 1.3. They comprise a digital speech processor with an integrated microphone, a coil, an implant, and an electrode

array. While the external components of a CI, the speech processor and coil, are typically worn behind the ear, the internal components, the implant and electrode array, must be inserted surgically. The implant is placed behind the ear, beneath the scalp. The electrode array, which contains up to 22 individual electrodes, is inserted into the cochlea, close to the auditory nerve. The microphone registers the incoming sounds (e.g. speech, environmental sounds) which the speech processor transforms into a set of digitally coded signals that can be interpreted by the implant (Macherey and Carlyon, 2014; Carlson, 2020; Wilson and Dorman, 2008; Lenarz, 2018). In specific, the frequency, intensity, and time course of the acoustic signal, as well as characteristic (speech) features, must be analysed in the speech processor to allow for the most detailed speech perception possible. This is accomplished through a complex process of digitisation, frequency band separation (corresponding to the number of electrode contacts in the electrode array), and selection of the relevant information for speech perception, to finally decode the digital signal into an analog signal (Adunka and Kiefer, 2005). The external coil ("transmitter") transmits the information through transcutaneous radio-frequency coupling to the internal implant ("receiver"), and is held in place across the skin by magnets. The internal implant delivers this information through the electrode array, which substitutes the damaged or absent hair cells of the cochlea and directly electrically stimulates the auditory nerve fibres, allowing the central nervous system to interpret this information as meaningful acoustic events, such as speech (Macherey and Carlyon, 2014; Lenarz, 2018; Carlson, 2020; Wilson and Dorman, 2008). Taken together, the CI takes over the function of the inner and the middle ear, as it bypasses the sound transmission through the eardrum and ossicular chain (Müller, 2005).

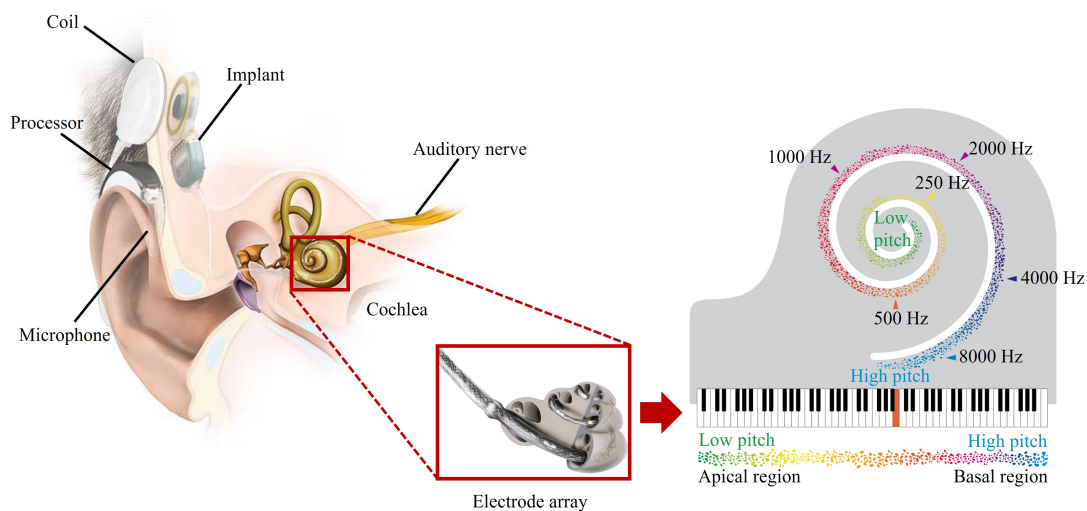


Figure 1.3: Structure and functionality of a CI. The individual components of a CI are shown on the left side. The schematic representation of inserting the electrode inside the cochlea is shown in the middle of the image. The tonotopic organisation of the cochlea is depicted on the right side. The apical region of the cochlea is in charge of analysing low-pitched sounds, while the basal region is responsible for analysing high-pitched sounds. Adapted with permission of MED-EL, Copyright © 2022 MED-EL Medical Electronics.

SOUND REPRESENTATION WITH A COCHLEAR IMPLANT As mentioned in the previous section, sounds are conducted to the cochlea via the outer and middle ear, where they are converted into electrical signals along the basilar membrane. The basilar membrane, which is an important part of the cochlea, is a soft-tissue membrane which is spatially structured such that each part of it is responsible for sound vibrations of a different frequency. Therefore, one can say that it acts as a spatial cochlear frequency map, which is called cochlear tonotopy (Li et al., 2021). Sounds in the high-frequency range cause vibrations at the base of the cochlea, while sounds in the low-frequency range cause vibrations at the apex (Macherey and Carlyon, 2014). The principle of tonotopic organisation of the cochlea is illustrated in Figure 1.3. In the case of a CI, a tonotopic stimulation of the auditory nerve and auditory perceptions of different frequencies can be obtained by placing the electrode contacts of the electrode array at different tonotopic locations in the cochlea. Hence, it replaces the function of the inner hair cells, that act as sensory receptors, and enables a direct electrical stimulation of the auditory nerve (Müller, 2005).

LIMITATIONS OF A COCHLEAR IMPLANT Given that roughly 3,000 inner hair cells are usually involved in producing a successful and perfect hearing sensation within frequencies ranging from 20 to 20,000 Hz (1,400 individual frequency steps) (Zeng, 2004; Zeng et al., 2014), it appears logical that up to 22 electrodes cannot fully compensate for this process as the covered frequency range is much narrower (approximately 200 – 8500 Hz), and the frequency resolution is much coarser (only as many frequencies as electrodes) (Limb and Roy, 2014). Consequently, electrical stimulation via a CI is very different compared to normal hearing in terms of psychophysical performance. In particular, there are crucial limitations in intensive, temporal, and spectral processing (Zeng, 2004). Nonetheless, a good speech understanding in quiet (optimal) listening environments is frequently achieved just a few months after implantation (Blamey et al., 2013; Holden et al., 2013). The most challenging hearing situations for CI users are those with additional noise or multiple concurrent speakers (Wilson and Dorman, 2008; Müller-Deile et al., 1995), in which it is difficult to extract the relevant speech information in such situations. In addition, music perception remains a major limitation of CIs (Drennan and Rubinstein, 2008) as the frequency representation of a CI is not fine-grained enough. However, within the past 35 years, the CI technology has improved significantly thanks to the close scientific cooperation between the manufacturers of CIs and clinicians (Dhanasingh and DeSaSouza, 2022). As a result, the features of CIs are gradually improving, for example, by developing noise reduction algorithms which improve hearing in noisy environments (Caldwell et al., 2017; Wolfe et al., 2015; Müller-Deile et al., 2008).

REHABILITATION AFTER IMPLANTATION The aim of a successful rehabilitation is to (re)gain the ability to hear and to understand speech. Ideally, it is attempted that the patient gets back the abilities to understand speech without the use of lip-reading, to use the telephone, and to understand speech in noisy environments (Dazert et al., 2020). The implantation of a CI alone does not always lead to better speech understanding. This entails a dynamic learning process that can last months or even years (Hoppe et al., 2017). The prerequisite for good hearing with a CI is

the individual fitting of the speech processor. This is typically performed by CI audiologists, who determine the individual stimulation parameters as well as the electrical stimulation current intensity (Hoppe et al., 2017; Dazert et al., 2020). The initial fitting of the speech processor takes place after the healing phase, about four weeks after the surgery. This is when the CI is switched on for the first time, and the CI users experience their first auditory impressions with the implant. These are often not perceived as speech by many CI users, but rather as noise (Deep et al., 2019). Important basic parameters, which are determined for each electrode separately are the C (comfort)-level, which is the volume at which sounds are perceived as comfortably loud, and the T (threshold)-level, which determines the hearing threshold (Hoppe et al., 2017). As part of the rehabilitation process, numerous subsequent fittings take place in the weeks following the initial fitting. In parallel, standardised clinical audiometric tests, medical examinations, and speech and language therapy must be performed by a large collaborative interdisciplinary group of specialists (Diller, 2009). After the rehabilitation phase, the patients have to come regularly to follow-up checks, ideally once a year. During these checks, hearing abilities are examined, speech processor settings are optimised, technical issues are resolved, and patient concerns are addressed.

FACTORS INFLUENCING THE OUTCOME Despite intensive hearing and speech training as part of rehabilitation and the efforts of finding the best CI settings, speech comprehension varies greatly between individuals and is affected by a variety of factors. Unfortunately, there is no perfectly developed theoretical account that explains the variability of speech outcomes after implantation. Therefore, it remains challenging to predict why some patients will have perfect outcomes with good speech intelligibility, while others will have only minimal or no benefits from using a CI (Pisoni et al., 2017; Boisvert et al., 2020). However, there are many studies showing that specific factors do have an influence on the outcome. Frequently reported factors are the aetiology of hearing loss, the duration of hearing loss, the duration of hearing aid use before implantation, the age of the patient at implantation, the clinical hearing scores before implantation, and surgical factors (Zhao et al., 2020; Blamey et al., 2013, 1992; Lazard et al., 2012a). Surgical factors are the placement of the electrode (Holden et al., 2013), the depth of insertion (James et al., 2019), the brand of the implant, and the amount of active electrodes during stimulation (Lazard et al., 2012c). While it is clear that there is a relationship between these factors and hearing outcome, the degree of evidence for many of these associations varies greatly between studies (Zhao et al., 2020). A recent study examined the largest cohort of adult CI implantees to date ($N = 2,735$) from three different clinics in three countries by including 21 factors, but they only could explain a modest percentage of the variance (Goudey et al., 2021). The authors suggest that including a large number of patients in a predictive analysis may not be the only solution, and that longitudinal studies may be more effective in identifying predictive variables that have an effect on the outcome.

Even though the previously mentioned factors, such as patient-related, aetiological, and surgical factors, have been shown to contribute to the clinical outcome with a CI, it is still poorly understood why exactly they are important. A largely supported view is that these factors are linked to neuroplasticity (Stropahl et al., 2017a). Neuroplasticity

is likely to appear in CI users as a result of auditory deprivation prior to implantation and the limited auditory input provided by the CI after implantation (Lazard et al., 2012a,b,c). Indeed, previous studies confirmed both correlations between longer periods of hearing deprivation and experience with a CI with a higher degree of cortical reorganisation (Giraud et al., 2001c; Green et al., 2005; Sandmann et al., 2015; Oh et al., 2003). Given its critical role, cortical reorganisation is discussed in greater depth in the following section.

1.3 NEUROPLASTICITY IN COCHLEAR IMPLANT USERS

Hearing function rehabilitation with a CI is impossible without neuroplasticity. Given that a CI user's auditory system has to learn to interpret the spectrally and temporally limited CI input as meaningful information, the brain must undergo structural and functional changes to accomplish this (Giraud et al., 2001a; Sandmann et al., 2015). This learning procedure is an example of neuroplasticity, as the nervous system has to adapt to a changing environment or lesion (Glennon et al., 2020; Merzenich et al., 2014; Hummel and Cohen, 2005). In CI research, two types of neuroplasticity are reported: intra-modal plasticity and cross-modal plasticity (e.g. Heimler et al., 2018; Stropahl et al., 2017a). Intra-modal plasticity takes place when neuronal changes occur within a specific cortical area as a consequence of increased or decreased input to that sensory system (Glick and Sharma, 2017). Hence, intra-modal plasticity is referred to the previously mentioned changes of the central auditory system in response to the CI. Cross-modal plasticity refers to the ability of one sensory system to activate the central processing areas of another sensory system when one sensory system is not fully functional, such as in the case of blindness or deafness (Glennon et al., 2020). Cross-modal and intra-modal changes have been observed in CI patients for both the auditory and the visual modality (for review see Heimler et al., 2014; Stropahl et al., 2017a). These mechanisms vary greatly depending on the onset of deafness as well as the time when hearing is restored. The following paragraphs expand on and review these two concepts of neuroplasticity.

INTRA-MODAL PLASTICITY Following CI implantation, the auditory system undergoes remarkable adaptive changes in order to understand the provided limited CI input. These adaptive changes within the auditory cortex in response to auditory sounds are an example of intra-modal plasticity. Electrophysiological studies with prelingual children (children who received their implant before acquiring natural speech, less than 3.5 years old) analysed auditory event-related potentials (AEPs) and a control group of age-matched NH children. They demonstrated a significant reduction in AEP latency and achieved latencies comparable to those of NH children in a short period of 3-6 months. However, in the case of children implanted later in life (after the age of 7 years), these positive effects of latency reduction were not observed, indicating that CI implantation should occur at the earliest possible age (Sharma et al., 2005, 2007). Thus, when children receive their CI early enough, the auditory system is able of catching up with the maturation, that was lost during the period of deprivation, even though there was no normal development with early acoustic experience (Kral et al., 2001, 2006). These children adapt quickly and easily to the new artificial input

provided by the CI, as evidenced by good speech recognition results (O'Donoghue et al., 2000; Hammes et al., 2002; Fryauf-Bertschy et al., 1997). In contrast, the speech outcomes of late-implanted children are significantly worse (Kral and O'Donoghue, 2010; Kral and Sharma, 2012; Niparko et al., 2010). The reason for these remarkable effects in early-implanted children is neuronal plasticity in the auditory cortex, which is greatest early in life (e.g. Kral et al., 2006). Thus, normal auditory system development necessitates auditory experience within a specific time window, which demonstrates the existence of a sensitive period for CI implantation (Niparko et al., 2010; Sharma et al., 2005, 2007). However, this does not necessarily imply that neuronal plasticity is no longer possible after the sensitive period has passed, but rather that the extent of plasticity is limited (e.g. Kral et al., 2016). Adults implanted later in life, on the other hand, cannot be easily compared to early implanted children, given that the auditory systems of those groups have substantially different prerequisites.

The temporal dynamics of cortical changes following implantation in postlingual adults (who became deaf after the development of language) are less clear. The best way to study these is to conduct longitudinal studies which track these changes. Despite the fact that there are not many of such studies, they discovered evidence of adaptive or intra-modal plasticity within the auditory cortex. For example, an electrophysiological study of Sandmann et al. (2015) demonstrated that within few days after switching on the implant, postlingually deaf CI users improved significantly in their speech perception ability. This improvement became even more apparent after a few weeks had passed, such that the CI users could get more experience with their new hearing input (Sandmann et al., 2015; Oh et al., 2003; Krueger et al., 2008). This comparably rapid adaptation (within a year) appears to occur even in patients who have had a long period of hearing loss. Further evidence for intra-modal changes comes from studies using positron-emission tomography (PET), that confirmed cortical activation within the primary auditory cortex, the secondary auditory cortex and auditory association areas in response to sounds after initial CI switch on (Naito et al., 1997). Giraud et al. (2001c) provided further evidence of increased activity within primary and non-primary auditory areas in the first week after implant switch-on. After one year of CI experience, cortical activation to sounds was found to be mostly centred on primary auditory areas, indicating a less effortful processing of the limited CI input. Moreover, presenting speech sounds caused increased activations within typical left-hemispheric language-related regions (Giraud et al., 2001c). A further longitudinal study that used Magnetoencephalography (MEG) by Pantev et al. (2006) showed that the greatest adaptation to the CI signal occurred within the first six months after switching on the CI. Accordingly, speech recognition performance improved during the first six months and then remained stable. Based on these results, the authors proposed that an intensive training should be implemented within the first 6 months to achieve the best possible results. The auditory system's ability to perform such rapid adaptations to unfamiliar auditory input demonstrates that, despite the period of auditory deprivation, it retains its high plasticity.

Moreover, these positive changes within auditory cortical areas are supported by positive correlations between auditory cortex activation and improved speech perception (Green et al., 2005; Olds et al., 2016), with better performing CI users showing cortical responses comparable to those of NH adults. Thus, good and bad CI

performers seem to be distinguishable by means of cortical responses (for review see Stropahl et al., 2017a). Nevertheless, it has been observed that increasing auditory-cortex activation as well as average speech intelligibility scores reach a plateau and remain stable after about one year post-implantation. This stable level can, in very special cases, get close to a NH listener performance (Debener et al., 2008), but is usually below the one of NH controls (Sandmann et al., 2015; Viola et al., 2012). However, being closer to NH listeners in terms of cortical activity seems to be a good sign for successful hearing rehabilitation with a CI, which further supports the importance of plasticity for the CI outcomes (Stropahl et al., 2017a).

In addition, it has been observed that deafness and subsequent implantation of a CI in postlingual CI users do not only induce changes in auditory processing, but also in visual processing (Stropahl et al., 2017a). EEG studies using visual-evoked potentials (VEPs) identified increased responses over occipital electrodes compared to NH controls for good CI performers and more anterior occipito-temporally distributed activation for poor CI performers, as shown by correlations with speech performance (e.g. Doucet et al., 2006; Kim et al., 2016). Similarly, using PET, Strelnikov et al. (2013) found that a stronger visual-cortex recruitment for visual stimuli measured right after the CI implantation resulted in improved speech performance after half a year of CI experience compared to CI users who had shown poorer visual-cortex responses.

Taken together, the brain of CI users undergoes plastic changes during the period of hearing deprivation and also after restoring hearing with a CI. These changes cause intra-modal adaptation to the CI input, which is reflected by enhanced auditory-cortex activation to sounds. As a results, postlingual deaf patients can, in theory, use their fully developed auditory system after CI implantation. However, due to their hearing loss prior to implantation, their auditory system may have degenerated or undergone cross-modal plasticity, as discussed in the following paragraph.

CROSS-MODAL PLASTICITY Cross-modal plasticity is a phenomenon in which intact sensory cortical regions (e.g. those responsible for perceiving touch or vision) take over functions of a deprived sensory cortical region (e.g. the auditory cortex), as in the case of hearing loss (Heimler et al., 2018). The first studies investigating this issue primarily focused on visual-evoked auditory-cortex responses or auditory-evoked visual-cortex responses. Depending on the type of cross-modal recruitment, it has been proposed that cross-modal cortical recruitment can be beneficial for hearing rehabilitation or not. For visually evoked potentials in the auditory cortex, several EEG studies have shown a negative correlation with speech performance in postlingual CI users (Sandmann et al., 2012; Doucet et al., 2006; Kim et al., 2016). Another study by Lee et al. (2003) discovered a negative correlation between the duration of deafness and the recovery of auditory responses, indicating that a greater cross-modal change, caused by a longer duration of deafness, is disadvantageous to auditory rehabilitation. This is in line with a PET study by Rouger et al. (2012), who found a negative relationship between the amount of cross-modal reorganisation and the outcome with the CI.

Other studies, using PET or functional near infrared spectroscopy (fNIRS), reported beneficial effects of cross-modal recruitment for speech intelligibility performance, explained by better audiovisual integration (Strelnikov et al., 2013; Anderson et al.,

2017; Rouger et al., 2012). Hence, cross-modal plasticity does not necessarily imply poor speech outcomes with a CI, it rather seems that the nature of cross-modal plasticity plays a crucial role. Although visual activation of the auditory cortex is potentially maladaptive, *auditory* activation of the *visual* cortex may improve the performance with a CI (Chen et al., 2016; Giraud et al., 2001a,b,c; Strelnikov et al., 2013). Initial supporting evidence came from various studies by Giraud and colleagues (Giraud et al., 2001a,b,c) who provided first reports of an enhanced recruitment of visual cortical areas in CI users compared to NH controls for a variety of auditory stimuli (words, syllables and environmental sounds). Interestingly, the enhanced visual-cortex activation appeared only in response to meaningful auditory input and not in response to white noise (Giraud et al., 2001b,c). Furthermore, better performing CI users showed more pronounced cross-modal activation of the visual cortex, which was also associated with speech perception and lip-reading abilities. Similar observations on cross-modal recruitment of the visual cortex have been made using fNIRS, showing that auditory-induced visual cortex activation was associated with good performers, whereas visual-induced auditory cortex activation was associated with bad performers (Chen et al., 2016). Thus, cross-modal reorganisation after CI implantation may be a mechanism that emerges during the post-implantation rehabilitation phase to compensate for the plastic changes that occurred before implantation due to deafness (Stropahl et al., 2017a).

Although the studies on cross-modal activation of the auditory cortex in response to visual tasks appeared to have negative effects on auditory perception, the reorganisation may have other positive effects on daily life communication by particularly focusing the visual aspect of natural speech, such as lip movements (Desai et al., 2008; Kawase et al., 2015). Accordingly, cross-modal activation in the auditory cortex for complex visual stimuli, such as faces, positively correlated with lip-reading skills (Stropahl et al., 2015). Similarly, cross-modal activation in the visual cortex was found to be positively related to lip-reading abilities (Giraud et al., 2001b). The fact that both types of cross-modal reorganisation improve lip-reading abilities adds to the theory that the development of various cortical processing patterns occurs with the overarching goal of making daily communication easier (Stropahl et al., 2017a).

In summary, postlingual CI users demonstrate both intra-modal and cross-modal changes both within the auditory and visual cortices. The precise function of cross-modal reorganisation remains unknown. During the period of hearing deprivation, cross-modal changes may enhance the remaining sensory modalities to make up for the missing or limited auditory input. On the other hand, cross-modal changes may improve hearing performance in CI patients after hearing is restored via a CI by evolving a stronger binding between the auditory and visual systems. This is why current studies have shifted their focus from analysing purely auditory and visual experimental tasks to including audiovisual conditions, as audiovisual conditions more closely capture real-life listening situations. The following section discusses the significance of audiovisual information in general, as well as the differences between NH listeners and CI users.

1.4 AUDIOVISUAL INTERACTIONS

Spoken language communication is an inherently audiovisual task, given that both auditory and visual input (lip movements, gestures, facial expressions), contribute to successful speech perception (Drijvers and Özyürek, 2017; Grant et al., 1998; Sumby and Pollack, 1954). This integration of audiovisual information is not essential in NH individuals in quiet surroundings, even though improvements in behavioural performance are observable even under such optimal situations (Arnold and Hill, 2001). However, during challenging hearing situations (for example, speech in background noise), the visual modality's supportive role in speech comprehension can be critical (Sumby and Pollack, 1954; Ross et al., 2007). Given the importance of audiovisual information for NH individuals, it stands to reason that audiovisual input is even more critical for CI implant users. In CI patients, only limited spectral and temporal auditory input is provided via the CI (Drennan and Rubinstein, 2008). Despite these pitfalls, CI users can reach a satisfying speech recognition levels in quiet environments (Hey et al., 2016; Holden et al., 2013), but speech recognition in noisy environments can be very frustrating (Müller-Deile et al., 1995; Hochmair-Desoyer et al., 1997). Therefore, employing multisensory listening strategies may be extremely beneficial for CI users. Additional sensory information (e.g. vision, touch) may considerably complement the degraded auditory input. And, as described in the previous sections, CI users experience auditory deprivation before implantation and after implantation they have to adapt to the novel acoustic information delivered through the CI (Sandmann et al., 2015).

To compensate for the artificial CI input, CI patients are expected to rely more heavily on vision (Kaiser et al., 2003), which can lead to the development of pronounced lip-reading abilities (Rouger et al., 2007; Stropahl et al., 2015). Indeed, it has been shown that CI users exhibit a stronger visual influence on auditory perception (Desai et al., 2008; Butera et al., 2018) and enhanced audiovisual interactions (Stevenson et al., 2017) both on the behavioural as well as the cortical level, which is assumed to emerge because of the limited CI input. Not only has it been shown that CI users are better than NH individuals at integrating audiovisual speech (Bavelier et al., 2006; Mitchell and Maslin, 2007; Kaiser et al., 2003; Rouger et al., 2007), they furthermore exhibit stronger bindings between the auditory and visual cortices when processing audiovisual speech (Giraud et al., 2001c; Strelnikov et al., 2015). The duration of CI experience seems to affect the degree of multisensory integration. A longer CI experience appears to be related to stronger audiovisual integrations in post-lingually deafened CI patients (Desai et al., 2008). However, most studies on altered processing of audiovisual speech conditions used PET, which is a technique with a good spatial resolution, but a poor temporal resolution. However, EEG studies are superior in that they provide excellent temporal resolution and allow analysing precise cortical processing steps (Biasiucci et al., 2019; Michel and Murray, 2012). The majority of previous EEG studies with CI users reported reduced and delayed cortical responses to *auditory* stimuli (Finke et al., 2016a; Henkin et al., 2014; Sandmann et al., 2009; Timm et al., 2012). Electrophysiological results on *audiovisual* conditions, especially linguistic stimuli are less often reported. A study by Schierholz et al. (2015) included audiovisual non-linguistic sinusoidal tones and white discs and showed enhanced

visual modulations of auditory cortical responses in CI users compared to NH controls. Moreover, a recent study by Radecke et al. (2022) reported differences in audiovisual processing and an audiovisual benefit for CI patients if complementing (congruent) visual information is available.

Given that previous studies using EEG reported in CI users difficulties in the processing of auditory information and enhancements in the processing of audiovisual information compared to NH individuals, one principle aim of this thesis is to provide further electrophysiological evidence for this issue by using audiovisual speech stimuli. Using electrical neuroimaging (topographic and source analyses), which are methods for differentiating the effects of cortical response latency, strength, and topographic differences (Michel et al., 2009), we aim to further understand the timecourse of audiovisual processing in CI users. These results are reported in Chapter 2, which covers the publication "The timecourse of multisensory speech processing in unilaterally stimulated CI implant users revealed by ERPs".

1.5 DIFFERENCE BETWEEN COCHLEAR IMPLANT USER GROUPS

As described in the previous sections, multisensory information seems to play an important role in CI users, given that they rely on visual input complementing the limited auditory CI input. However, the vast majority of previous studies on auditory and audiovisual processing in CI users included CI users with bilateral hearing loss (Finke et al., 2016a; Sandmann et al., 2015; Schierholz et al., 2015; Radecke et al., 2022). This means that the CI patients are provided with a CI bilaterally (CI + CI on contralateral side) or bimodally (CI + hearing aid on contralateral side). However, in recent years, the clinical margins for CI indication have been expanded to include unilateral hearing loss, allowing for the implantation of patients who suffer from single-sided deafness (SSD). These patients are supplied with a CI and are NH on the contralateral ear (Arndt et al., 2011, 2017; Buechner et al., 2010; Zeitler and Dorman, 2019). This CI user group is very special because the signal quality of the auditory input differs greatly between the two ears, resulting in a maximally asymmetric auditory processing (Gordon et al., 2013; Kral et al., 2013). Nevertheless, it has been shown that CI implantation is highly advantageous in these patients. Particularly because the restoration of binaural hearing leads to a better sound localisation, speech-in-noise intelligibility, and quality of life (Kitterick et al., 2015). But, it is unclear whether the aforementioned benefits in SSD CI users are dependent on the side of implantation. In line with insights from CI users with bilateral hearing loss (Mosnier et al., 2014), initial reports indicate a right-ear advantage for recognising speech in SSD CI patients (Wettstein and Probst, 2018). The hypothesis that the side of implantation plays an important role is further supported by a recent study by Han et al. (2021), who reported distinct cortical plasticity depending on the side of deafness in SSD CI users. During passive listening, right-sided deafness resulted in greater deafness-induced adaptive plasticity in the left hemisphere (Han et al., 2021). Furthermore, it is unknown whether SSD CI patients have the same speech processing abilities with their NH ear as NH individuals (Arndt et al., 2019; Maslin et al., 2015). A further aim of this thesis is to extend the limited number of studies on SSD CI patients by addressing these two scientific questions (side-of-implantation and speech processing

differences between CI and NH ears). These are covered in Chapter 3 which covers the second publication "Side-of-Implantation Effect on Functional Asymmetry in the Auditory Cortex of Single-Sided Deaf Cochlear-Implant Users".

Given that SSD CI users have an intact ear on the contralateral side, it stands to reason that, despite the benefits of the CI, the NH ear acts as the primary communication channel (Kitterick et al., 2015; Ludwig et al., 2021). This is consistent with previous studies that compared auditory stimuli processing between the CI ear and the NH ear in SSD CI users, which revealed delayed cortical responses in the CI ear (Finke et al., 2016b; Bönitz et al., 2018). However, studies on SSD CI users are scarce, and as far as we are aware, they only cover auditory processing and not audiovisual processing. Currently, it is unknown, whether the enhanced visual influence on auditory speech processing (e.g. Schierholz et al., 2015; Radecke et al., 2022), as well as the pronounced lip-reading skills (e.g. Rouger et al., 2007; Stropahl et al., 2015) which have been shown for CI users with bilateral hearing loss, are also present in SSD CI users. Systematic comparisons between CI users with bilateral and unilateral hearing loss are rare. The few studies found differences in performance between SSD CI users and bimodal or bilateral CI users in speech-in-noise conditions (Williges et al., 2019) and in conditions including multiple concurrent speakers (Bernstein et al., 2016). Another aim of this thesis is to extend previous results (see Chapter 2) to SSD CI users and to systematically compare the timecourse of auditory and audiovisual speech processing, as well as the lip-reading skills between CI users with bilateral hearing loss, unilateral hearing loss (SSD CI users), and NH controls. This study is described in Chapter 4 and is named "Electrophysiological differences and similarities in audiovisual speech processing in CI users with unilateral and bilateral hearing loss".

1.6 THE INFLUENCE OF BACKGROUND NOISE

As mentioned before, CI users are capable of reaching good levels of speech intelligibility in optimal (quiet) listening environments, already a few months after CI switch on (Fetterman and Domico, 2002; Hey et al., 2016; Holden et al., 2013; Sandmann et al., 2015). However, hearing in demanding conditions, for example when competing background noise is present, remains a challenge for CI users. And these hearing problems increase with increasing background noise intensity levels (Müller-Deile et al., 1995; Wilson and Dorman, 2008). NH individuals already have limited speech intelligibility in the presence of background noise (Billings et al., 2009), but CI users are far more severely affected, even after many years of use (Davidson et al., 2010; Fu and Nogaki, 2005; Radecke et al., 2022).

The negative effects of competing background noise have been confirmed by electrophysiological studies that showed delayed cortical responses in the presence of background noise in NH individuals (Billings et al., 2009), mild to moderate hearing impairment (Gillis et al., 2022), and CI users (Finke et al., 2016b; Han and Dimitrijevic, 2020). In addition to the objective EEG results, it is frequently reported that competing background noise affects the subjective listening effort, which increases as a function of background noise intensity level (Han and Dimitrijevic, 2020; Hughes and Galvin, 2013; Winn, 2016). Listening effort is defined as the amount of cognitive resources and attention necessary to comprehend speech (e.g. Gosselin and Gagne, 2011; Downs,

1982) and can, for instance, be assessed by means of subjective rating scales. Background noise, reduces the ability to understand speech, because it can interfere with the speech signal, both physically (energetic masking) and perceptually (informational masking due to competing talkers) (Pollack, 1975). This competition between noise signals and speech makes it more difficult to extract speech information (Helfer and Vargo, 2009; Romei et al., 2011) and necessitates leveraging higher cognitive processes, such as attention and memory, to comprehend the meaning of a spoken message (Kent, 1992). As a result, hearing-impaired individuals frequently complain about hearing restrictions and excessive listening effort, especially in noisy situations (for review see McCormack and Fortnum, 2013; Zhao et al., 1997).

To counteract this problem, CI manufacturers are constantly developing and improving noise reduction algorithms, which have been shown to be effective in specific situations with interfering background noise (Mauger et al., 2014; Müller-Deile et al., 2008; Wolfe et al., 2015). "ForwardFocus" (Cochlear Limited) is one example of such a noise reduction algorithm, which has been introduced recently (Hey et al., 2019). Another aim of this thesis is to objectively assess by means of electrophysiological measures, whether "ForwardFocus" is effective in specific noise situations. This is of interest, because this algorithm has only been tested by means of behavioural measures in auditory-only situations so far (Hey et al., 2019, 2021). Apart from extending previous observations by objective electrophysiological measures, we included both auditory and audiovisual stimuli. When NH individuals hear in quiet listening conditions, hearing naturally is the dominant sense because speech can be easily identified based solely on auditory information (Gatehouse and Gordon, 1990; Shannon et al., 1995). In contrast, understanding speech based solely on visual information is extremely difficult, as NH individuals are not trained in lip-reading (Bernstein and Liebenthal, 2014). However, when the information from one modality is distorted, the sensory dominance of one sense shifts to the one that is perceived more clearly. Thus, in situations with competing noise, people tend to rely more on the visual cues, which has been shown by Ma et al. (2009), who found that a visual benefit for audiovisual words was present at signal-to-noise (SNR) ratios of -8 dB or more. This is in accordance with the principle of inverse effectiveness (Meredith and Stein, 1986), which predicts enhanced audiovisual integration when one sense is distorted. Sensory degradation occurs in CI users, as their auditory input is distorted. As a consequence, it has been shown that CI users benefit more strongly from additional visual input and are better at integrating audiovisual information compared to NH individuals (Rouger et al., 2007; Stropahl et al., 2017b). Given the importance of visual cues for CI users in general (Rouger et al., 2007; Stevenson et al., 2017; Strelnikov et al., 2013), these visual cues may become even more essential in the presence of noise, as auditory information becomes more challenging to comprehend (van de Rijt et al., 2019). Thus, this thesis aims to address the question whether background noise has an impact on audiovisual interactions in CI users (Schierholz et al., 2015; Radecke et al., 2022) and whether a noise reduction algorithm, "ForwardFocus" in specific, can influence this process. These questions are addressed in Chapter 5 which discusses the submitted manuscript "The influence of background noise and noise reduction by 'ForwardFocus' on audiovisual speech perception in bimodal CI users: an ERP study".

1.7 METHODS

In this thesis, we perform a variety of behavioural analyses in conjunction with electrophysiological measures. These behavioural measures comprise hit rates (percentage of correct responses) and response times in the EEG paradigms, clinical tests measuring hearing abilities, demographic information, various questionnaires, cognitive tests, rating scales, and other supplementary tests for answering the scientific questions of the projects. The majority of behavioural measures are self-explanatory and are further described in the original publications. The electrophysiological methods, on the other hand, are more complicated, which is why the background for those will be presented in the following sections of this chapter. It is worth noting that there exist additional analytical methods for analysing ERP data, such as time-frequency analysis. Time-frequency analysis can be used to analyse alterations in spectral power and phase consistency across single trials that are time locked to events within the experiment (Makeig et al., 2004). However, this method is not described further because it was not used in the projects included in this thesis.

1.7.1 *Electroencephalography*

Electroencephalography (EEG) is a non-invasive method of recording the electrical activity of the brain. Electrodes placed on the head's surface measure voltage potentials caused by current flow in and around neurons. EEG is a method that is nearly a century old, which is widely used in clinical diagnostics and experimental neuroscience (Biasiucci et al., 2019). It was Hans Berger, a German psychiatrist and neurologist, who reported the first human EEG recording in 1929 (Berger, 1929; Biasiucci et al., 2019). He could show that it is possible to record signals via electrodes placed on the skull. Berger and his innovation were initially dismissed and questioned by the scientific community. Only after the British physiologists Edgar Adrian and Bryan Matthews replicated Berger's discoveries in 1934, EEG was accepted as a non-invasive measure of the brain's electric fields (Biasiucci et al., 2019; Adrian and Matthews, 1934). At that time, EEG measurements were primarily used to study brain oscillations, which is the rhythmic and repetitive electrical activity of the brain (Başar, 2022). Later, Davis (1939) reported so-called "event-related potentials" (ERPs), observable by means of EEG. ERPs are time-locked voltage changes generated by the brain in response to internal or external sensory, cognitive and motor events (Blackwood and Muir, 1990). In other words, event-related potentials are electrical potentials that are related to a specific event (Luck, 2014). They represent the sum of postsynaptic potentials generated when a large number of similarly oriented cortical pyramidal neurons (thousands to millions) fire in together while processing information (Peterson et al., 1995). ERPs are acquired by averaging numerous trials of the continuous EEG data, since they are relatively small and thus difficult to detect in raw signals. This procedure enables to identify clear and distinct components of the ERP signal. Figure 1.4 schematically depicts the basic pre-processing steps which are needed to obtain average ERPs. In 1964, the first reported ERP component was called "contingent negative variation" (CNV; Walter et al., 1964). Several more components were found in the following decades. These include the N1 (for review see Näätänen and Picton, 1987), the P3 (for

review see Polich, 2007) and the mismatch-negativity (MMN; for review see Näätänen et al., 2007). Usually, ERP components are named based on their polarity and order of appearance (e.g. N1; first negative response after the onset of a stimulus) or latency (e.g. N100; negative response occurring 100 ms after the onset of a stimulus) (Davis, 1939; Winkler et al., 2013). But, this is not always the case, as there are some exceptions, such as the MMN or the CNV.

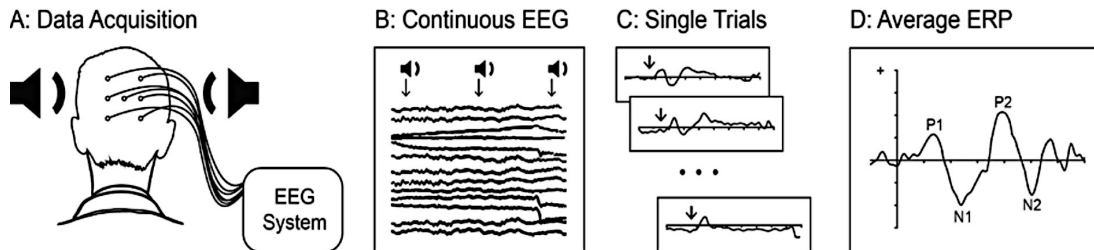


Figure 1.4: Schematic representation of how auditory ERPs are obtained from a continuous EEG. A) An EEG is recorded via electrodes placed on the scalp, while auditory stimuli are presented. B) Previously defined triggers mark the beginning of a trial and are saved within the recorded continuous EEG measurements. C) Based on the predefined triggers, individual segments (single trials) can be extracted from the continuous EEG signal. They contain a short pre-stimulus baseline as well as the period of interest after the stimulus onset. D) Based on the segments, an averaged ERP waveform that reflects time- and phase-locked neural activity associated with stimulus processing can be generated. Adapted with permission from Key (2016), Copyright © AIMS Press.

NEURAL ORIGINS OF ERPS Action potentials and postsynaptic potentials are the two main types of electrical activity produced by neurons. An action potential, also known as a spike, is an electrical signal that travels across the cell membrane of a neuron, beginning at the cell body and ending at the axon terminals, where neurotransmitters are released. Postsynaptic potentials occur when neurotransmitters bind to receptors on the membrane of a postsynaptic cell, altering the flow of ions across the cell membrane (Luck, 2014). Given the short duration of action potentials (less than 2 ms), individual action potentials do not contribute significantly to scalp-recorded EEG signals (Buzsáki et al., 2003). Postsynaptic potentials are the primary generators of EEG fields measured on the brain's surface or at the scalp (Buzsáki et al., 2003), as they can last up to tens or hundreds of milliseconds (Luck, 2014). The only exception when ERPs rather reflect action potentials than postsynaptic potentials are wave I and II of the brainstem auditory evoked response (see section "Auditory event-related potentials"). They reflect precisely synchronised action potentials generated within the cochlea and passing through the auditory nerve within a few milliseconds after a sudden sound onset (Luck, 2014; Luck and Kappenman, 2011).

When a postsynaptic potential is generated within a single neuron, a tiny electrical dipole (an oriented flow of current) is created. ERPs can only be recorded at the head's surface when the dipoles of thousands of similarly oriented neurons add up, creating a strong electric field. If the orientations of the neurons in a specific region are not the same, the dipoles can partially cancel out, making detection at a distant EEG electrode impossible (Luck, 2014). Pyramidal cells (located within the cerebral cortex) are the

principal neurons that are commonly assumed to be the primary generators of an ERP (Schaul, 1998). These cells are oriented perpendicular to the cortical surface (Picton, 2010). As mentioned before, the activity of a single pyramidal cell cannot be registered by an EEG electrode. To obtain an interpretable signal recorded from an electrode, several conditions must be met (see Luck, 2014): 1) A large amount of neurons has to be simultaneously activated. 2) The orientation of them has to be similar. 3) The majority of the neurons' postsynaptic potentials must originate from the same part of the neuron. 4) To prevent cancellation, most of the neurons must have the same current flow direction. If these conditions are fulfilled, the dipoles of each neuron can add up, generating a strong electric field. The brain then acts as a volume conductor, allowing EEG to be recorded on its surface (Schaul, 1998).

ADVANTAGES OF EEG COMPARED TO OTHER IMAGING TECHNIQUES There exist various other techniques, like functional Magnetic Resonance Imaging (fMRI), PET, and MEG for studying cortical activity. The great advantage of EEG compared to other methods is its non-invasive nature, its excellent temporal resolution, and its comparably low cost. Consequently, major disadvantages of other tools compared to the EEG are that they are invasive (PET) or expensive (fMRI, PET, MEG). But, not all techniques measure the same mechanisms. EEG and MEG basically measure the immediate correlate of cortical activity, given their excellent temporal resolution (1 ms or even better), which is impossible with fMRI or PET. These are haemodynamic (dynamics of blood flow) measures that can indirectly measure neural activity. PET and fMRI consequently have a temporal resolution of many hundred milliseconds, but they have the advantage of providing a good spatial resolution (within the millimetre range) (Luck, 2014). Functional near-infrared spectroscopy (fNIRS), which is an optical imaging technique, is also an option, given its satisfactory spatial and temporal resolution (Quaresima and Ferrari, 2019; Wilcox and Biondi, 2015). Even though fNIRS has a lower spatial resolution than other functional imaging techniques and it cannot measure activity deeper than about 1 cm below the surface of the brain, it has a higher temporal resolution than other functional imaging techniques and is relatively low-cost, like EEG (Wilcox and Biondi, 2015). Combining EEG and fNIRS is a good and feasible alternative for leveraging the benefits of both techniques (Chen et al., 2015).

When conducting a study with CI users, not every neuroimaging tool can be used. For instance, fMRI cannot be used in CI patients because images can be degraded by artefacts caused by the implant's magnet (Kim et al., 2015). Furthermore, the magnetic field can cause a device displacement or heating of the implant (Giraud et al., 2001c). Modern CI models are MRI-compatible up to 3 Tesla. However, this does not completely eliminate the possibility of serious accidents, which is why fMRI is usually not used for research purposes (Bawazeer et al., 2019). Other techniques, such as PET also have significant drawbacks, which make conduction of studies with CI users a challenge. The number of consecutive measurements is limited as radioactive tracers for imaging active brain regions are used. The doses of these radioactive tracers are highly restricted (Giraud et al., 2001c). As a result, PET cannot be used in children with CIs and is inapplicable in longitudinal studies as they require repeated measurements within a relatively short time, such as those used to evaluate

the rehabilitation process. MEG is expensive and the CI causes strong artefacts in the signal. Nevertheless, it has been shown that MEG can be used in CI users under certain conditions (Pantev et al., 2006), for which the authors developed a unique set-up in their laboratory. fNIRS has been shown to be suitable for investigating CI users (Chen et al., 2016), and it can be used in combination with EEG (Chen et al., 2015). Finally, many studies have shown that using an EEG in the context of auditory processing in CI users is a highly effective tool (e.g. Sandmann et al., 2015; Schierholz et al., 2015; Sharma et al., 2002). The electrical artefact caused by the CI can be effectively removed by means of an independent component analysis (ICA; Debener et al., 2008; Viola et al., 2012). Given the EEG's excellent temporal resolution, single cortical processing steps can be tracked. This allows for the investigation of fast brain dynamics or the tracking of the timecourse of auditory and (by extension) audiovisual speech processing steps (Biasiucci et al., 2019; Michel and Murray, 2012). By using "electrical neuroimaging", which comprises topographic and source analyses, information on the spatial distribution and the likely origin of an EEG signal can be obtained (Michel and Murray, 2012). When used correctly, this is a powerful approach for studying multisensory interactions (Stevenson et al., 2014). For that reason, the projects included in this thesis make use of analytical approaches coming from electrical neuroimaging, in order to provide spatio-temporal information on the data. The precise definitions and explanations, as well as how these methods can be applied to an ERP dataset, will be provided later in this chapter (section "ERP data analysis").

1.7.2 Auditory event-related potentials

Auditory event-related potentials (AEPs) are brain responses to an auditory stimulus that can be classified based on the time point of their appearance. Auditory responses are divided into three categories: early, middle, and late (Picton, 2010).

EARLY AUDITORY RESPONSES Early AEPs can be further divided into electrocochleographic responses (ECochG), which occur around 2-3 ms after stimulus onset (Ferraro, 2010), and auditory brainstem responses (ABR), which occur around 10 ms after an acoustic event (Winkler et al., 2013). To measure ECochGs, either the external ear canal or the tympanic membrane are used, and they are helpful for diagnostics (inner ear and auditory nerve disorders) and during surgeries for checking the functionality of the cochlea and the auditory nerve (Ferraro, 2010). ABRs originate from subcortical brain structures and they are recorded from electrodes that are placed on the skull. ABRs are labelled by Roman numerals arranged in the order of their elicitation (Winkler et al., 2013). ABRs are primarily measured in clinical settings due to their automatic nature. Therefore, they can be used in diagnostics for estimating peripheral hearing loss (Stone et al., 2009). As they can be recorded automatically, ABRs have been used in newborn hearing screening for several years, along with other measures such as otoacoustic emissions (Gravel et al., 2005).

MIDDLE-LATENCY RESPONSES Middle-latency responses (MLR) are typically recorded from the vertex, with a mastoid or neck electrode as a reference and can be

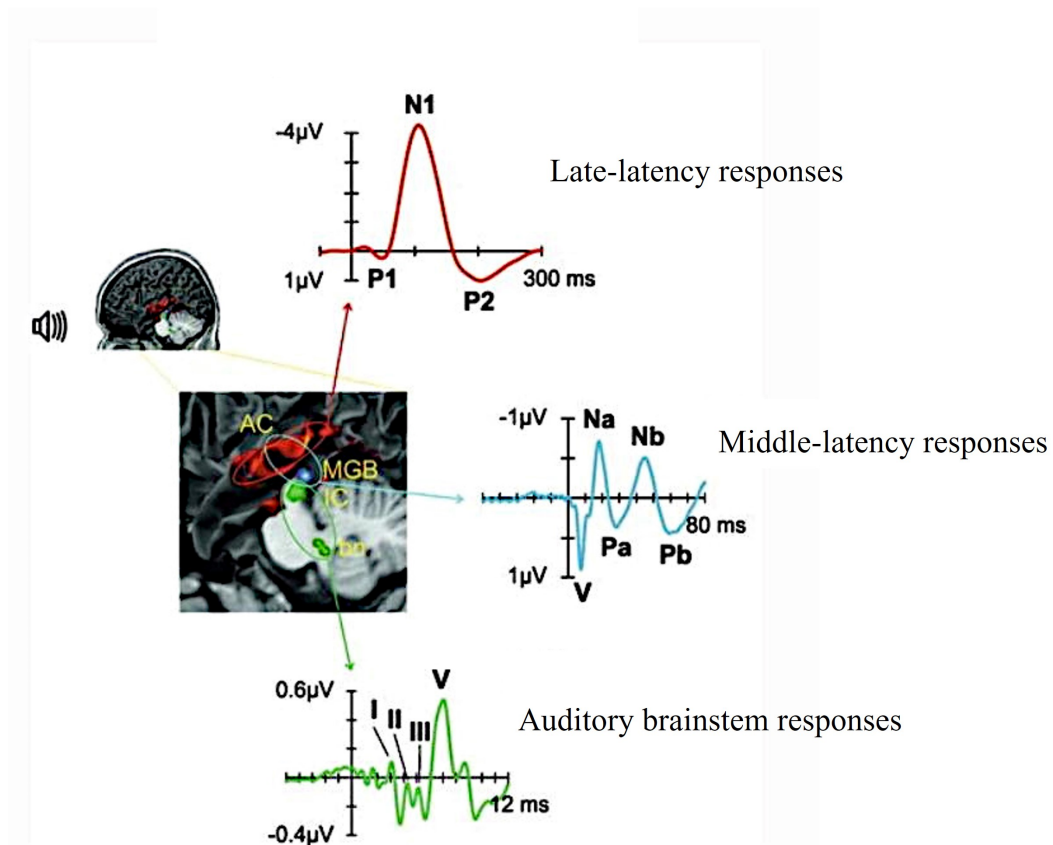


Figure 1.5: Auditory event-related responses: main waveforms and neuronal generators. There are three types of waveforms elicited by auditory events, which are grouped by their latency ranges. The auditory brainstem response belongs to the early latency responses and is displayed in green. Their neuronal origin comprises the brainstem nuclei (bn) and the inferior colliculus (IC). The middle-latency responses, coloured in blue, originate from the medial geniculate body in the thalamus (MGB). The late-latency responses (in red) primarily originate from the auditory cortex. Adapted with permission from Winkler et al. (2013), Copyright © 2013 Springer Science+Business Media New York.

measured at latencies within 12-50 ms (Winkler et al., 2013). MLRs originate from the thalamus and parts of the auditory cortex and are labelled by their polarity (capital letters P or N) and appearance, e.g. Na (20 ms), Pa (30 ms), and Nb (around 40 ms) (Winkler et al., 2013; Picton, 2010). MLRs can be used for a variety of purposes, including the detection of low-frequency hearing thresholds, the evaluation of the function of CI implants, the evaluation of the proper function of the auditory pathway, and the localisation of auditory pathway lesions (Kraus and McGee, 1993).

LATE-LATENCY AUDITORY RESPONSES Late auditory responses, also referred to as cortical AEPs are typically the responses measured by an EEG within experimental studies. They occur at latencies starting from 50 ms after the onset of an acoustic event. Late auditory responses have two naming conventions: they either denote the serial order of the response, beginning with the first detected response, e.g. "N1," or they indicate the typical peak latency of the waveform, e.g. "N100" (Davis, 1939; Winkler

et al., 2013). Aside from the peak latency classification, there are two other common distinctions which can be made. They can either be exogenous (obligatory) if they are elicited by every event, regardless of its relationship to previous or concurrent events or the participant's task, motivations, knowledge, and so on. Or, they can be endogenous (cognitive) components, which are elicited only when there is a relationship between the event and other events or some aspect of the participant's mental state (Winkler et al., 2013; Kraus and Nicol, 2009). Because endogenous components are strongly influenced by attention, it is usually required to pay attention to an active cognitive task for eliciting endogenous components (Cone-Wesson and Wunderlich, 2003). Examples of endogenous components are the P₃, which reflects the evaluation and classification of incoming auditory events (for review see Polich, 2007), or the N₄₀₀, which is considered an indicator of lexical and semantic processing (for review see Lau et al., 2008). Exogenous components are primarily influenced by external factors, such as acoustic stimulus characteristics. They can, however, be modulated by higher order cognitive processes, such as attention and memory (Picton et al., 1974; Tremblay et al., 2001).

The projects described in this thesis focus on the N₁ and P₂ components, as they represent the sensory encoding of auditory stimulus features (Näätänen and Picton, 1987). Regarding the origin of the N₁ (negative deflection approximately measured 100 ms after stimulus onset), it has been stated that they are generated within the primary and secondary auditory cortex and the auditory association cortices (Ahveninen et al., 2006; Bosnyak et al., 2004; Näätänen and Picton, 1987; Lütkenhöner and Steinsträter, 1998; Hari et al., 1987). Concerning the origin of the P₂ (positive deflection approximately measured 200 ms after stimulus onset), previous studies using MEG or fMRI have reported that the auditory P₂ is generated in the anterior part of the auditory cortex, in particular the lateral part of the Heschel's gyrus (Bosnyak et al., 2004; Ross and Tremblay, 2009; Tremblay et al., 2014; Hari et al., 1987; Pantev et al., 1996). In addition, the planum temporale has been reported as a generator for the P₂ (Crowley and Colrain, 2004; Godey et al., 2001; Hari et al., 1987). However, current auditory signal propagation models consider the existence of an underlying anatomy with a semi-hierarchical and highly parallel organisation (Kaas and Hackett, 2000). This suggests that the prominent auditory N₁ and P₂ components include generators not only within auditory cortices but also within a distributed network along superior temporal cortices, fronto-parietal structures, and even visual cortices. Importantly, top-down effects, such as attention direction or the prediction of auditory events originating in the frontal cortical areas can modulate these auditory processes (Dürschmid et al., 2019).

1.7.3 ERP data analysis

There exist various approaches to analyse ERP data recorded from the continuous EEG. These range from traditional peak amplitude and latency measures, to topographic and source analyses and time-frequency and connectivity methods. The following paragraphs explain the methods used in the projects included in this thesis.

SENSORY DATA ANALYSES, GLOBAL FIELD POWER Traditionally, the EEG is analysed by grouping certain electrodes of interest and studying the peak voltage within a specific time window (Michel and Murray, 2012). The precise time of the peak (latency) and its magnitude (amplitude) are of particular interest. Previous studies have shown that certain neural processes are most strongly reflected at specific electrode sites on the scalp and within a specific time window after stimulus onset. For example, the auditory N1 usually has its peak between 75-150 ms and the P2 peaks between 150-250 ms in NH individuals, which has its largest peak at the vertex electrodes compared to occipital, lateral or frontal ones (Näätänen and Picton, 1987; Crowley and Colrain, 2004). Therefore, when investigating the N1 many researchers (here in the case of CI research) choose a set of electrodes around the vertex (Cz electrode) (e.g. Finke et al., 2016a) or even only the Cz electrode (e.g. Schierholz et al., 2015).

A better or more objective approach is to study the global field power (GFP), which is the standard deviation of the activity across all electrodes at a specific point in time (Lehmann and Skrandies, 1980; Murray et al., 2008). The GFP has the great advantage of being a reference-independent measure of response strength that does not require an a priori selection of a subset of electrodes (Murray et al., 2008). The GFP has been investigated in studies with CI users previously (e.g. Sandmann et al., 2015; Radecke et al., 2022). The method of using a priori defined channels of interest was applied in project 2 and project 4, and the GFP was chosen for project 2 and project 3 (see following chapters).

TOPOGRAPHIC ANALYSIS, GLOBAL MAP DISSIMILARITY Apart from analysing specific peaks, it is also possible to analyse the distribution of the voltage on the scalp, by looking at the so-called topography or scalp potential map (Michel and Murray, 2012). A new topographic map is generated within the millisecond range, and a change indicates that the orientation and/or the distribution of dipoles within the brain has changed (Vaughan Jr, 1982; Lehmann, 1987). Topographies do not randomly vary across time, but rather remain stable for a set period of time before changing to another topography, which are defined as microstates (Michel and Koenig, 2018). Distinct topographies correspond to different neural generator configurations. This means that topographic analyses can be used to test for topographic differences between conditions or groups, because they directly permit the conclusion that the underlying neural generators are different (Michel and Murray, 2012). To do so, one can analyse the global map dissimilarity (GMD; Lehmann and Skrandies, 1980), which quantifies differences in topographies independently of the signal strength (Murray et al., 2008). The software CARTOOL (Brunet et al., 2011) was employed to analyse the GMD by calculating a so-called "topographic" ANOVA (TANOVA; Murray et al., 2008). A more detailed description can be found in publication 1 and 3 of this thesis.

If the results of the GMD indeed show that distinct topographies are present for different groups or experimental conditions, this can be seen as indication for distinct neuronal generators (Vaughan Jr, 1982). The reason for this difference in generators cannot be explained solely on the basis of the GMD. Possible reasons for differences in topographies between groups or conditions, can be different neuronal generators or a latency shift. To disentangle these effects, one can apply a hierarchical clustering

analysis with group and/or condition averaged data to identify the most prominent topographies within the time windows of interest. Again, the software CARTOOL is a suitable tool for conducting this analysis. Hierarchical clustering identifies the minimal number of topographic maps that can explain the greatest variance in a dataset (Murray et al., 2008). After obtaining predominant topographic maps in a given dataset, one can explore how these are distributed on a single-subject level by applying a single-subject fitting analysis (Murray et al., 2008). Possible interesting variables could be the *first onset of topographic maps*, which corresponds to the latency, or the *map presence*, which indicates how dominantly or frequently specific topographic maps are represented. This method was used in a subset of the projects in this thesis, and it is described in greater detail in publications 1 and 3.

SOURCE ANALYSIS Given the excellent temporal resolution of EEG signals, ERPs recorded from the scalp (sensor ERPs) can be used to analyse neuronal dynamics, particularly the timecourse of cognitive and perceptual processes. However, complex neuronal processes are characterised by a spatio-temporal interaction of different cortical regions. Even for very simple sounds, several cortical areas are engaged, and information from various cortical areas must be integrated within a millisecond range (Shahin et al., 2007). The signal recorded from a single electrode is generated by a combination of a variety of distinct neural sources, and how strongly each source contributes to the measured signal is unknown and may be even distant from one another (da Silva, 2013). Therefore, when only sensor level data are considered, differences between conditions or individuals cannot be easily interpreted in terms of their spatial origin (Stropahl et al., 2018). Source modelling allows to approximate the location and timing of cortical processes (da Silva, 2013; Michel et al., 2004). As described in the paragraph on topographic analysis, different topographies across experimental conditions or groups are caused by different underlying neural sources. If this is confirmed by hierarchical clustering and single-subject fitting analysis, source analysis can help to find these different sources (Murray et al., 2008; Michel and Murray, 2012).

However, it is important to bear in mind that a source analysis cannot provide a completely accurate picture of the neuronal origins of the ERP signals, unlike for fMRI. Given the locations and orientation of the cortical dipoles, as well as a structural MRI scan, it is possible to calculate the voltage distribution on the scalp. This is known as the "forward problem" (Luck, 2014). In general, forward modelling is the prediction of observations from a model with a given set of parameters (Baillet, 2013). In contrast, an observed voltage distribution on the scalp does not uniquely determine its underlying dipole configuration. This is referred to as the "inverse modelling problem": A particular voltage distribution can be produced by a large number of different dipole configurations, and there is no way of inferring which one is the right one based on scalp-recordings (Luck, 2014). Thus, "inverse modelling" is the reciprocal problem in which observations are used to approximate some unknown model parameters (Baillet, 2013). To be more specific, the inverse problem consists of inferring the location, orientation, and amplitude of the dipole sources given the electrode potentials measured on the scalp. In general, this is a non-linear and, as mentioned before, an underconstrained problem. To proceed, there exist two major

classes of approaches: *non-parametric distributed source models*, for which the positions and possibly the directions of a potentially large number of dipoles are assumed, such that only the amplitudes need to be recovered, which makes the problem linearly solvable; and *parametric concentrated source models*, for which the full problem of determining the optimal positions, orientations, and amplitudes is solved for just a small number of dipoles (Grech et al., 2008; Sorrentino and Piana, 2017).

As a result, a significant amount of work has been done on practical methods for MEG/EEG forward and inverse modelling, which has resulted in a handful of approaches that are now well defined (Salmelin and Baillet, 2009). EEG with a large number of electrodes in conjunction with source modelling is regarded an electrical brain imaging tool (Michel and Murray, 2012; Salmelin and Baillet, 2009), which can be used to investigate the likely spatial origin of the EEG signals recorded from the scalp.

In this thesis, we chose to use the software Brainstorm (Tadel et al., 2011) for source analysis, as previous studies successfully conducted source analysis with CI users and deaf individuals using this software (Stropahl et al., 2015; Prince et al., 2021; Bottari et al., 2020; Stropahl et al., 2018). The exact procedure applied in the projects which are part of this thesis is described in each publication in detail. In project 1 and 3 we applied source analysis in order to test whether the topographic differences observed in the sensor data can be explained by distinct neuronal generators. In project 2 and 4 we used source analysis to further strengthen the differences observed in the sensory ERPs.

1.8 AIM OF THIS THESIS

The general aim of this thesis is to gain a broader understanding of auditory and especially audiovisual speech processing in CI users in four ERP studies.

The first project included in this thesis "The timecourse of multisensory speech processing in unilaterally stimulated cochlear implant users revealed by ERPs" (Chapter 2) is an ERP study investigating how audiovisual speech stimuli are processed by CI users. Previous studies have shown that CI users show both intra-modal and cross-modal experience-related cortical reorganisation before and after implantation (for review see Stropahl et al., 2017a). Cross-modal activity in the visual cortex has yet only been demonstrated in PET and fNIRS studies (Giraud et al., 2001c; Chen et al., 2016), however the precise underlying temporal dynamics are still unknown. Furthermore, previous research has indicated that CI users have a stronger visual influence on auditory speech perception (Desai et al., 2008; Butera et al., 2018), an enhanced binding between the auditory and visual systems (Giraud et al., 2001c; Strelnikov et al., 2015), and more pronounced lip-reading abilities than NH listeners (Rouger et al., 2007; Stropahl et al., 2015). ERP studies, however, were rather focused on purely auditory stimuli (Finke et al., 2016a; Sandmann et al., 2009), despite the fact that audiovisual stimuli better reflect CI users' real-life listening situations (Stevenson et al., 2017). By including simple, non-linguistic audiovisual stimuli, a recent ERP study could show that CI users demonstrate an enhanced visual influence on auditory processing at N1 latency compared to NH controls (Schierholz et al., 2015). However, similar ERP results on audiovisual speech stimuli have not been reported yet. By

administering methods from electrical neuroimaging (Michel et al., 2009), including topographic and source analysis, our study provides a further spatio-temporal perspective. Specifically, we addressed the research question: Do CI users show an altered timecourse of audiovisual speech processing and differences in audiovisual interactions compared to NH controls?

The second project "Side-of-implantation effect on functional asymmetry in the auditory cortex of single-sided deaf cochlear-implant users" (Chapter 3) focuses on auditory processing of speech stimuli in SSD CI users. Previous studies assessing auditory and audiovisual processing in CI users mainly examined CI users with bilateral hearing loss (Beynon et al., 2005; Finke et al., 2016a), neglecting the group of SSD CI users. A few years ago, it was common practise not to implant CIs in patients who still had an intact contralateral ear (Arndt et al., 2011; Buechner et al., 2010; Zeitler and Dorman, 2019). Recent studies, however, have provided evidence demonstrating the beneficial effects of implantation as reflected in improved sound localisation, speech-in-noise intelligibility, and quality of life as a result of binaural hearing restoration (Kitterick et al., 2015). Because this patient group is relatively new, research on SSD CI users is still rather scarce. However, it has been shown previously that processing via the CI ear of SSD CI users is delayed when compared to processing via the NH ear (Bönitz et al., 2018; Finke et al., 2016b). In contrast, it is unknown whether SSD CI users' NH ears process auditory stimuli in the same way that NH listeners do. Furthermore, it is unknown whether the side of implantation has an impact on the benefits of implantation. A previous ERP study suggested an altered auditory-cortex asymmetry in CI users compared to NH listeners, however the results were based on CI users with bilateral hearing loss (Sandmann et al., 2009). With the second project, we concentrated on the research questions: 1) Do SSD CI users show differences in auditory speech processing between the CI ear and the NH ear? 2) Does the side of implantation affect auditory speech processing with the CI ear of SSD CI users? 3) Does the side of stimulation affect auditory speech processing with the NH ear of SSD CI users, and is this pattern different in NH controls?

The third project "Electrophysiological differences and similarities in audiovisual speech processing in cochlear implant users with unilateral and bilateral hearing loss" (Chapter 4) addresses the question whether there are any differences in auditory and audiovisual speech processing, as well as lip-reading abilities, between CI users with bilateral hearing loss, unilateral hearing loss, and NH controls. As mentioned before, the number of behavioural and ERP studies on auditory processing focusing on SSD CI users is very limited, with even fewer studies comparing SSD CI users to CI users with bilateral hearing loss. The few existing studies reported behavioural differences in speech-in-noise performance (Williges et al., 2019) and in situations with multiple concurrent speakers (Bernstein et al., 2016). Furthermore, as far as we are aware, no ERP study on audiovisual processing or lip-reading abilities in SSD CI patients has been conducted yet. For this reason, another goal of this thesis is to extend previous results on CI users with bilateral hearing loss to SSD CI users, given the previously reported altered audiovisual interactions in CI users with bilateral hearing loss compared to NH controls. To this end, we conducted a follow-up study of Layer et al. (2022), including a group of SSD CI users, allowing for a systematic comparison of different CI patient groups.

Finally, this thesis discusses the issue of the impact of background noise on auditory and audiovisual speech processing in CI users, which is covered in the fourth project "The influence of background noise and noise reduction by 'ForwardFocus' on audiovisual speech perception in bimodal CI users: an ERP study" (Chapter 5). Background noise in general has a significant impact on CI users' hearing ability, delaying auditory ERPs (Finke et al., 2016a; Han and Dimitrijevic, 2020) and increasing subjective listening effort (Hughes and Galvin, 2013; Winn, 2016). However, noise reduction algorithms have been shown to be effective aids in noisy environments (Mauger et al., 2014; Wolfe et al., 2015). A recently introduced algorithm ("ForwardFocus") was tested in purely auditory behavioural settings and was found to be useful in situations with background noise (Hey et al., 2019). One aim of the fourth project is to provide objective evidence of beneficial effects of employing "ForwardFocus" in noise by using ERPs. In addition, as already mentioned, audiovisual information is highly beneficial for CI users, and when the auditory input is further hampered by noise, the visual aspect of speech may become even more crucial. It is unknown, whether background noise in general can modulate audiovisual interactions in CI users and whether a noise reduction algorithm, such as "ForwardFocus", has an effect on auditory and audiovisual speech processing. To address these issues, this project covers the following research questions: 1) How exactly does background noise affect behaviour as well as auditory and audiovisual speech processing in CI users? 2) Can potential benefits of using "ForwardFocus" be objectively evaluated by means of ERPs, and does "ForwardFocus" affect audiovisual interactions?

To sum up, this thesis investigates a variety of research questions within the framework of auditory and especially audiovisual speech processing in different CI user groups and in background noise. The following chapters contain the original publications as well as a detailed summary of each project. The final chapter of this thesis is dedicated to a broad discussion of the projects.

"THE TIMECOURSE OF MULTISENSORY SPEECH PROCESSING
IN UNILATERALLY STIMULATED COCHLEAR IMPLANT USERS
REVEALED BY ERPS"

This chapter is based on the following publication:

Layer, N., Weglage, A., Müller, V., Meister, H., Lang-Roth, R., Walger, M., Murray, M. M., & Sandmann, P. (2022). The timecourse of multisensory speech processing in unilaterally stimulated CI implant users revealed by ERPs. *NeuroImage: Clinical*, 34, 102982. DOI: 10.1016/j.nicl.2022.102982.

Author contributions:

Natalie Layer: Conceptualisation; literature review; creation, editing and calibration of stimuli; programming of the experiment; pilot measurements; participant recruitment; data acquisition; data analysis; interpretation and presentation of results to research group; discussion of findings and further procedures; writing initial manuscript draft; incorporating suggestions of co-authors; submission to journal; presentation of results at national and international conferences. **Anna Weglage:** Assistance in data acquisition; assistance in data analysis; group discussions; proof-reading of manuscript. **Verena Müller:** Advisory role (CI technician); initial assistance in participant enrolment; proofreading of manuscript. **Hartmut Meister:** Funding acquisition; conceptualisation; provision of software for stimuli generation; discussion; suggestions; proofreading of manuscript. **Ruth Lang-Roth:** Conceptualisation; proof-reading of manuscript. **Martin Walger:** Conceptualisation; proof-reading of manuscript. **Micah M. Murray:** Presentation of analytical method (topographic analysis); validation of analysis; discussions; suggestions; proofreading of manuscript. **Pascale Sandmann:** Supervisor of Natalie Layer; central role in project conceptualisation; funding acquisition; suggestions; discussions; assistance in writing manuscript and proofreading of manuscript.

SYNOPSIS

OBJECTIVES A CI can only transmit limited spectral and temporal information (Drennan and Rubinstein, 2008). Previous research suggests that CI patients compensate for this by an enhanced interaction between the auditory and visual system (Stevenson et al., 2017). Moreover, previous PET and fNIRS studies revealed that CI users recruit the visual cortex for purely auditory tasks (Chen et al., 2016; Giraud

et al., 2001c) and that they are better lip-readers than NH controls (Rouger et al., 2007). Event-related potential (ERP) studies confirmed enhanced multisensory interactions for CI users for rudimentary non-linguistic stimuli (Schierholz et al., 2015). The aim of the presented study is to extend previous results to audiovisual speech stimuli and to explore the timecourse of audiovisual speech processing by employing methods from electrical neuroimaging (Michel et al., 2009), including topographic and source analyses.

METHODS We tested postlingual CI users, who became hearing impaired after language acquisition, and a group of matched NH controls. The CI users had bilateral hearing loss and were fitted with either two CIs or a CI plus a hearing aid on the contralateral ear. We recorded an EEG with 64 electrodes while the participants performed a speeded response task. The stimuli were videos of the syllables /ki/ and /ka/ which were presented in auditory (A), visual (V) and audiovisual (AV) conditions. Each button on a computer mouse corresponded to one of the two syllables. The participants' task was to press the corresponding button as quickly and as accurately as possible when a syllable was recognised. Moreover, we asked the participants to rate the difficulty of performing the task after each experimental block, and we administered a lip-reading test with monosyllabic words after the EEG experiment.

ANALYSIS During the EEG task, we obtained both behavioural and ERP measures. For behavioural measures, we analysed the response times (how often the decision was correct) and the hit rates (how accurately a decision was made). The ERP data were analysed by employing an additive model (Barth et al., 1995), since we aimed to investigate audiovisual interactions. To do so, we compared auditory responses (A) with visually modulated auditory responses (AV-V). Non-linear multisensory interactions were assumed if these responses were unequal (e.g. Vroomen and Stekelenburg, 2010). First, we examined the ERP data at the sensor level by calculating the global field power (the standard deviation of voltages across all electrodes on the scalp) during the time windows of the N1 and P2 peaks. Specifically, we compared the amplitudes and latencies of those peaks across groups (NH, CI) and conditions (A, AV-V). Further, we used hierarchical clustering to identify dominant topographic maps in the data, which we then used to perform a subsequent single-subject fitting analysis. We concentrated on the map presence and the first onset of the maps within the N1 and P2 time windows and examined whether there are any statistically significant group and condition differences. Finally, we conducted a source analysis, focusing on the visual and auditory cortices. Again, we concentrated on the N1 and P2 time windows, examined the activity peaks and latencies, and compared groups and conditions statistically.

RESULTS The behavioural results revealed multisensory integration for both groups, as evidenced by shorter audiovisual response times compared to the two unisensory conditions. The topographic and source analyses of the N1 and P2 ERPs showed distinct P2 patterns and smaller amplitudes for AV-V compared to A. These observations confirmed a multisensory effect for both groups and revealed a cortical auditory response that was modulated by the simultaneous processing of the visual stimulus.

Nonetheless, CI users revealed a different pattern of N1 topography, indicating a strong visual impact on auditory speech processing. Aside from these condition-specific effects, the results also revealed ERP differences between the two groups, not only in N1/P2 ERP topographies, but also in the cortical source configuration. In contrast to NH listeners, the CI users additionally recruited the visual cortex within the N1 time window, which was positively correlated with the experience with the CI. In addition, CI users showed a delayed and reversed right-lateralised activation within the auditory cortex compared to NH listeners. Finally, CI users identified more correct monosyllabic words, demonstrating superior performance in the lip-reading test compared to NH listeners.

DISCUSSION Overall, the behavioural and ERP results demonstrate a clear audio-visual benefit for both groups and a CI-specific pattern in cortical activation at N1 latency. This CI-specific pattern could be a strategy to compensate for the limited CI input, allowing CI users to develop enhanced lip-reading skills and get closer to the behavioural performance of NH listeners in audiovisual speech conditions. Our findings are clinically relevant, as they emphasise the necessity of evaluating the CI outcome not only in auditory-only, but also in audiovisual speech conditions.



The timecourse of multisensory speech processing in unilaterally stimulated cochlear implant users revealed by ERPs

Natalie Layer^{a,*}, Anna Weglage^a, Verena Müller^a, Hartmut Meister^b, Ruth Lang-Roth^a, Martin Walger^{a,b}, Micah M. Murray^{c,d,e,f}, Pascale Sandmann^a

^a University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Otorhinolaryngology, Head and Neck Surgery, Audiology and Pediatric Audiology, Cochlear Implant Center, Germany

^b Jean-Uhrmacher-Institute for Clinical ENT Research, University of Cologne, Germany

^c The Sense Innovation and Research Center, Lausanne and Sion, Switzerland

^d The LINE (The Laboratory for Investigative Neurophysiology), Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

^e CIBM Center for Biomedical Imaging of Lausanne and Geneva, Lausanne, Switzerland

^f Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

ARTICLE INFO

Keywords:

Cochlear implant
Event-related potential
Cortical plasticity
Multisensory integration
Audiovisual interaction
Audiovisual speech perception

ABSTRACT

A cochlear implant (CI) is an auditory prosthesis which can partially restore the auditory function in patients with severe to profound hearing loss. However, this bionic device provides only limited auditory information, and CI patients may compensate for this limitation by means of a stronger interaction between the auditory and visual system. To better understand the electrophysiological correlates of audiovisual speech perception, the present study used electroencephalography (EEG) and a redundant target paradigm. Postlingually deafened CI users and normal-hearing (NH) listeners were compared in auditory, visual and audiovisual speech conditions. The behavioural results revealed multisensory integration for both groups, as indicated by shortened response times for the audiovisual as compared to the two unisensory conditions. The analysis of the N1 and P2 event-related potentials (ERPs), including topographic and source analyses, confirmed a multisensory effect for both groups and showed a cortical auditory response which was modulated by the simultaneous processing of the visual stimulus. Nevertheless, the CI users in particular revealed a distinct pattern of N1 topography, pointing to a strong visual impact on auditory speech processing. Apart from these condition effects, the results revealed ERP differences between CI users and NH listeners, not only in N1/P2 ERP topographies, but also in the cortical source configuration. When compared to the NH listeners, the CI users showed an additional activation in the visual cortex at N1 latency, which was positively correlated with CI experience, and a delayed auditory-cortex activation with a reversed, rightward functional lateralisation. In sum, our behavioural and ERP findings demonstrate a clear audiovisual benefit for both groups, and a CI-specific alteration in cortical activation at N1 latency when auditory and visual input is combined. These cortical alterations may reflect a compensatory strategy to overcome the limited CI input, which allows the CI users to improve the lip-reading skills and to approximate the behavioural performance of NH listeners in audiovisual speech conditions. Our results are clinically relevant, as they highlight the importance of assessing the CI outcome not only in auditory-only, but also in audiovisual speech conditions.

1. Introduction

A cochlear implant (CI) is a neuroprosthesis that can help regaining hearing abilities and communication in patients with sensorineural hearing loss. However, the electrical CI signal transmits only limited

spectral and temporal information compared to natural acoustic hearing (Drennan and Rubinstein, 2008). Consequently, after implantation, the central auditory system has to learn to recognise the new, artificial input as meaningful sounds (Giraud et al., 2001a; Sandmann et al., 2015). This learning is an example of neural plasticity, wherein the nervous system

* Corresponding author.

E-mail address: natalie.layer@uk-koeln.de (N. Layer).

<https://doi.org/10.1016/j.nicl.2022.102982>

Received 3 December 2021; Received in revised form 1 March 2022; Accepted 2 March 2022

Available online 4 March 2022

2213-1582/© 2022 The Authors.

Published by Elsevier Inc.

This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

adapts to the changing environment (Glennon et al., 2020; Merzenich et al., 2014).

Neural plasticity in CI users has been examined in several studies, reporting an increasing activation in the auditory cortex to auditory stimuli over the first months after implantation (Giraud et al., 2001c; Green et al., 2005; Sandmann et al., 2015). These experience-related cortical changes may even extend to the visual cortex, as indicated by the finding that CI patients engage not only the auditory, but also the visual cortex in purely auditory speech tasks (Chen et al., 2016; Giraud et al., 2001c). However, the latter results are based on positron emission tomography (PET) neuroimaging, which provides better spatial than temporal resolution. The temporal properties of the cross-modal response in the visual cortex of CI users have not yet been characterised.

As both, auditory and visual input (articulatory movements, gestures), contribute to speech perception, spoken language communication is an inherently audiovisual task (Drijvers and Özyürek, 2017; Grant et al., 1998; Sumbly and Pollack, 1954). Especially in difficult hearing situations (for instance speech in background noise) the supportive role of the visual modality can be crucial for understanding speech (Sumbly and Pollack, 1954). This has been characterised by the “principle of inverse effectiveness”, which states that if unisensory signals are poorly perceptible, there is a remarkable perception enhancement for multisensory signals (Stein and Meredith, 1993). The principle of inverse effectiveness has been assessed for audiovisual speech, in particular words, and has shown that if the visual or auditory presentation alone of a word was difficult, there was a higher benefit to understand audiovisual words (van de Rijt et al., 2019). It is therefore not surprising that CI users show an enhanced visual influence on auditory perception (Desai et al., 2008) and stronger audiovisual interactions (Stevenson et al., 2017) due to the limited CI input, on both behavioural and cortical levels. CI users are not only better at integrating audiovisual speech information when compared to normal-hearing (NH) listeners (Bavelier et al., 2006; Kaiser et al., 2003; Mitchell and Maslin, 2007; Rouger et al., 2007). They also show a remarkably strong binding between the activation in the visual and auditory cortices during visual and audiovisual speech processing (Giraud et al., 2001c; Strelnikov et al., 2015). Given that previous findings on altered audiovisual speech processing in CI users are based on PET imaging, it is currently not well understood, whether these cortical alterations are already present at initial (sensory) or only at later cognitive processing stages.

Event-related potentials (ERPs) derived from continuous electroencephalography (EEG) represent a suitable tool to study cortical alterations in CI users (Sandmann et al., 2009, 2015; Sharma et al., 2002; Viola et al., 2012). ERPs have a high temporal resolution, which allows tracking single cortical processing steps (Biasiucci et al., 2019; Michel and Murray, 2012). In particular, the auditory N1 (negative potential around 100 ms after stimulus onset) and the P2 (positive potential around 200 ms after stimulus onset) are elicited for auditory stimuli and are generated at least partly in the primary and secondary auditory cortex (Ahveninen et al., 2006; Bosnyak et al., 2004; Näätänen and Picton, 1987). Current frameworks suggest an automatic and sequential mode of cortical auditory processing, and a parallel processing of auditory information of distinct stimulus features within the supratemporal plane (De Santis et al., 2007; Inui et al., 2006). These auditory processes however can be influenced by top-down effects, in particular attention or predictions of incoming auditory events mediated by the frontal cortex (Dürschmid et al., 2019). Most of the previous ERP studies with CI users have focused on the N1 ERP to *unisensory auditory stimuli*, reporting that the N1 in CI users is reduced in amplitude and/or prolonged in latency compared to NH listeners (Finke et al., 2016a; Henkin et al., 2014; Sandmann et al., 2009; Timm et al., 2012). Recent studies, however, used more rudimentary, non-linguistic audiovisual stimuli (sinusoidal tones and white discs) to show an enhanced *visual* modulation of the *auditory* N1 ERP in CI users when compared to NH listeners (Schierholz et al., 2015, 2017). These findings in CI users suggest not only difficulties in the cortical sensory processing of auditory stimuli,

but also enhanced audiovisual interactions at sensory processing stages, particularly in conditions involving basic audiovisual stimuli.

The present EEG study extends previous research (Schierholz et al., 2015, 2017) by examining whether initial sensory cortical processing is also altered in CI patients in more complex audiovisual stimulus situations, especially in speech conditions. We used auditory syllables in combination with a computer animation of a talking head, providing the visual component of the stimuli (Fagel and Clemens, 2004; Schreit-müller et al., 2018). These types of speech stimuli are advantageous as they are highly controllable, reproducible, speaker-independent and perfectly timed. We compared the cortical processing of auditory-only and audiovisual syllables between CI users and NH listeners by means of electrical neuroimaging (Michel et al., 2009), including topographic and ERP source analysis to explore the timecourse of cortical processing. Compared to the traditional analysis of ERP data, which is based on waveform morphology at certain electrode positions, electrical neuroimaging is reference-independent and considers the spatial characteristics and the temporal dynamics of the global electric field to distinguish between effects of response strength, latency and distinct topographies (Michel et al., 2009).

Based on previous results with simple audiovisual stimuli (Schierholz et al., 2015, 2017), we predicted a multisensory effect in CI users and NH listeners for the more complex audiovisual speech stimuli as well, both on the behavioural and on the cortical level. However, due to the limited CI input and experience-related cortical changes after implantation, we expected group differences in voltage topographies and cortical source configurations, which would point to altered cortical activation in CI users during auditory and audiovisual speech processing.

2. Material and methods

2.1. Participants

In total, forty participants took part in the present study. Among these participants, six had to be excluded from further analyses due to the following reasons: bad EEG signal quality ($n = 2$), low number of correct trials ($n = 2$), one CI patient was hearing too well on the contralateral ear (rather single-sided deaf), and one NH participant was excluded due to the use of psychotropic drugs. Thus, for further analyses the total number of participants was thirty-four, with seventeen CI users (11 female, mean age: 59.3 years, range: 27–75 ± 12.1 years) and seventeen NH listeners (11 female, mean age: 59.9 years, range: 34–79

± 11.1 years). NH controls were matched by gender, age, handedness, stimulated ear and years of education. The CI users were post-lingually deafened, implanted either unilaterally ($n = 3$, all left-implanted using a hearing-aid on the contralateral ear) or bilaterally ($n = 14$). All CI users had been using their CI continuously for at least 12 months prior to the experiment. Therefore, in the case of bilateral implantation, either the ear which met that criterion or the ‘better’ ear (the ear showing the higher speech perception scores in the Freiburg Monosyllabic test) was used as stimulation side. Details on the implant system and the demographic variables can be found in Table 1.

To verify age-appropriate cognitive abilities, the DemTect Ear test battery was used (Brünecke et al., 2018). DemTect Ear is a version of the conventional DemTect (Kalbe et al., 2004) especially adjusted for patients with limited hearing abilities. This version enables to test cognitive skills independently of hearing and prevents disadvantages caused by hearing loss. It consists of various subtests including a word list, a number transcoding task, a word fluency task, digit span reverse, and delayed recall of the word list. These tests measure attention, memory and word fluency skills. All participants achieved total scores within the normal, age-appropriate range (13–18 points). Additionally, speech recognition abilities were measured by means of the German Freiburg monosyllabic speech test (Hahlbrock, 1970), using a sound intensity level of 65 dB SPL (see Table 4). None of the participants had a history of

Table 1
Demographic information on the CI participants; HA = hearing aid.

ID	Sex	Age	Handedness	Fitting	Stimulated ear	Etiology	Age at onset of hearing impairment (years)	CI use of the stimulated ear (months)	CI manufacturer
1	m	61	right	CI + CI	left	unknown	41	15	MedEl
2	f	62	right	CI + CI	left	unknown	10	36	Cochlear
3	f	51	right	CI + CI	right	amalgam poisoning	20	54	Cochlear
4	m	65	right	CI + CI	left	hereditary	28	66	MedEl
5	m	53	right	CI + CI	left	Cogan syndrome	43	99	Cochlear
6	f	75	right	CI + CI	left	hereditary	57	30	Advanced Bionics
7	f	39	right	CI + CI	right	otosclerosis	24	17	Advanced Bionics
8	f	70	right	CI + CI	left	unknown	37	56	MedEl
9	f	70	right	CI + CI	left	meningitis	69	20	MedEl
10	m	59	right	CI + HA	left	unknown	49	33	Advanced Bionics
11	f	63	right	CI + CI	left	meningitis	20	106	Advanced Bionics
12	f	64	right	CI + CI	left	whooping cough	9	78	Cochlear
13	m	53	right	CI + CI	left	unknown	30	235	Cochlear
14	f	58	right	CI + HA	left	unknown	49	18	Advanced Bionics
15	f	71	right	CI + CI	left	otitis media	21	63	Cochlear
16	f	27	right	CI + HA	left	sudden hearing loss	6	236	MedEl
17	m	56	right	CI + CI	right	hereditary	19	63	MedEl

psychiatric illness and all participants were native German speakers. All participants had normal or corrected-to-normal vision, as measured by the Landolt test (Landolt C) according to the DIN-norm given by Wesemann et al. (2010) and all participants were right-handed (assessed by the Edinburgh inventory; range: 80–100%; Oldfield, 1971).

All participants gave written informed consent prior to data collection and were reimbursed. The study was conducted in accordance with the Declaration of Helsinki (2013) and was approved by the Ethics Committee of the medical faculty of the University of Cologne (application number: 18–257).

2.2. Stimuli

The stimuli used in this experiment were presented in three different conditions: visual-only (V), auditory-only (A) and audiovisual (AV). Additionally, there were trials with a black screen only ('nostim'), to which the participants were instructed to not respond. We used the Presentation software (Neurobehavioral Systems, version 21.1) and a computer in combination with a duplicated monitor (69 in.) to control stimulus delivery. The stimuli were the two syllables /ki/ and /ka/ and they were taken from the Oldenburg logatome speech corpus (OLLO; Wesker et al., 2005). The syllables were cut out of the available logatomes from one speaker (female speaker 1, V6 'normal spelling style', no dialect). These two syllables differed in their phonetic distinctive features (vowel place and height of articulation) in the vowel contrast (/a/ vs. /i/; Micco et al., 1995). These German vowels are different in terms of central frequencies of the first (F1) and second formant (F2). They represent the highest contrast between German vowels (e.g. Obleser et al., 2003), which makes them highly distinguishable for CI patients. These two syllables not only highly differ in terms of auditory (phoneme) realisation, but also in their visual articulatory (viseme) realisation. A viseme is the visual equivalent of the phoneme: a static image of a person articulating a phoneme (Dong et al., 2003). There are some phonemes that share identical visemes (Cappelletta and Harte, 2012; Lucey et al., 2004; Mahavidyalaya, 2014), but for the vowels of the syllables used in this study, the visemes are clearly distinguishable (see illustrations in Jachimski et al., 2018), which is of importance given that we present visual-only trials as well. The syllables were edited using Audacity (version 3.0.2) in order to be cut and adjusted to the same duration of 400 ms. The intensity level of the syllables was adjusted in

Adobe Audition CS6 (version 5.0.2) to ensure an equal intensity level of the stimuli (adjusted to the maximal amplitude).

The visualisation of the syllables was created with MASSY (Modular Audiovisual Speech SYNthesizer; Fagel and Clemens, 2004), which is a computer-based video animation of a talking head. This talking head has been previously validated for CI patients and is an objective tool to create visual and audiovisual speech stimuli (Massaro and Light, 2004; Meister et al., 2016; Schreitmüller et al., 2018). To be able to create articulatory movements corresponding to the auditory speech sounds, it is necessary to create files that transform the previously transcribed sounds into a probabilistic pronunciation model providing the segmentation and the timing of every single phoneme. This can be achieved by means of the web-based tool MAUS (Munich Automatic Segmentation; Schiel, 1999). To create a video file of the MASSY output, the screen recorder Bandicam (version 4.1.6) was used. The final editing of the stimuli was done with Pinnacle Studio 22 (version 22.3.0.377), creating video files of each stimulus in each condition: 1) Audiovisual (AV): lip-movements with corresponding speech sounds, 2) Auditory-only (A): black screen (video track turned off) combined with speech sounds, 3) Visual-only (V): lip-movements without speech sounds (audio track turned off). Each trial consisted of a static face (500 ms) and the video, which had a duration of 800 ms (20 ms initiation of lip movements + 400 ms auditory syllable + 380 ms completion of lip movements). For further analyses, we focused on the processing of the moving face (starting 500 ms post-stimulus onset/after the static face), as the responses to static faces comparing NH listeners and CI users have been investigated previously (Stropahl et al., 2015).

For CI users, the stimuli were presented via a loudspeaker (Audiometer-Box, type: LAB 501, Westra Electronic GmbH) placed in front of the participant. Since the CI users were supplied with a hearing aid or a second CI at the contralateral side, the device was removed during the recording. Additionally, the contralateral ear was closed with an ear-plug to ensure that only one ear was stimulated. For the NH participants, the stimuli were presented monaurally via insert earphones (3 M E-A-RTONE 3A) at the same side as in the matched CI users, and the contralateral ear was covered with an ear-plug as well. Stimulus presentation via insert earphones is advantageous when compared to free-field presentations because insert earphones avoid a confounding stimulation of the second ear. Similar to the CI users, three of the NH listeners were stimulated on the right ear, and fourteen on the left ear. The

syllables were presented at 65 dB SPL. Both the CI users and the NH listeners rated the perceived loudness of the speech sounds by means of a seven-point loudness rating scale (Sandmann et al., 2009, 2010), which allowed to adjust the sound intensity to a moderate level of 60–70 dB (Allen et al., 1990).

2.3. Procedure

The participants were seated comfortably in an electromagnetically shielded and dimly lit room. The distance from head to screen was approximately 175 cm. The task was to discriminate the syllables /ki/ and /ka/ regardless of their modality (AV, A, V) by pressing a corresponding button as fast as possible. Each syllable was assigned to one of the two buttons of a mouse. The sides of the assigned buttons were counterbalanced across participants to avoid a potential bias caused by the used finger.

In all conditions (AV, A, V, 'nostim'), 90 trials each were presented per syllable, resulting in a total number of 630 trials (90 repetitions $\times$ 3 conditions (AV, A, V) $\times$ 2 syllables (/ki/, /ka/) + 90 'nostim'-trials). Each trial started with a static face of the talking head (500 ms) followed by a visual-only, auditory-only, audiovisual syllable or 'nostim'. Right after, a fixation cross was shown until a response has been given by the participant. The trials were pseudo-randomised such that no trial of the same condition and syllable appeared twice in a row. In total, there were five blocks which lasted for five minutes each, resulting in a total experimental time of approximately 25 minutes. After each block, the participants were given a short break. Before the experiment, the participants were given a practice block consisting of five trials per condition to ensure that the task was understood. An overview of the paradigm is displayed in Fig. 1A.

2.4. Additional behavioural measures

Apart from hit rates and response times recorded during the EEG task, we obtained additional behavioural measures. After each experimental block in the EEG task, the participants were asked to indicate on a scale, how effortful it was to perform the task. For that purpose, we used the 'Borg Rating of Perceived Exertion'-scale (Borg RPE-scale; Williams, 2017). Further, we examined the lip-reading abilities of the participants by using a behavioural lip-reading test, which has been applied in previous studies (Stropahl et al., 2015; Stropahl and Debener, 2017).

It consists of monosyllabic words taken from the German Freiburg monosyllabic speech test (Hahlbrock, 1970) and is video-taped with various professional speakers. The task of the participants was to watch the video of the visual performance of a word and to report, which word was lip-read. Additionally, the CI patients had to fill in the Nijmegen Cochlear Implant Questionnaire (NICQ; Hinderink et al., 2000) assessing the quality of life in CI users.

2.5. EEG recording

EEG data were recorded with 64 AG/AgCl ActiCap slim electrodes using a BrainAmp system (BrainProducts, Gilching, Germany) and a customised electrode cap with an electrode layout (Easycap, Herrsching, Germany) according to the 10–10 system. Two of the 64 electrodes were placed below (vertical eye movements) and beside (horizontal eye movements) the left eye to record an electrooculogram (EOG). All channels were recorded against a nose-tip reference, with a midline ground electrode, placed slightly anterior to Fz. Data were recorded with a sampling rate of 1000 Hz and were online analogically filtered between 0.02 and 250 Hz. Electrode impedances were kept below 10 k Ω throughout the recording.

2.6. Data analysis

Data analysis was performed using MATLAB 9.8.0.1323502 (R2020a; Mathworks, Natick, MA) and R (version 3.6.3; R Core Team (2020), Vienna, Austria). Topographic analyses were conducted in CARTOOL (version 3.91; Brunet et al., 2011). For source analysis, Brainstorm (Tadel et al., 2011) was used. The following R packages have been utilised: ggplot2 (version 2.3.3) for creating plots; dplyr (version 1.0.4), tidyverse (version 1.3.0) and tidyr (version 1.1.3) for data formatting; ggpubr (version 0.4.0) and rstatix (version 0.7.0) for statistical computations.

2.6.1. Behavioural data

First, trials with missing or false alarm responses were removed from the dataset. For each participant, outlier trials with reaction times (RTs) exceeding the individual mean by more than three standard deviations were removed for each condition separately. Afterwards, the mean hit rates and RTs were calculated for each condition (AV, A, V). We focused on the averages of hit rates and RTs computed across the two syllables

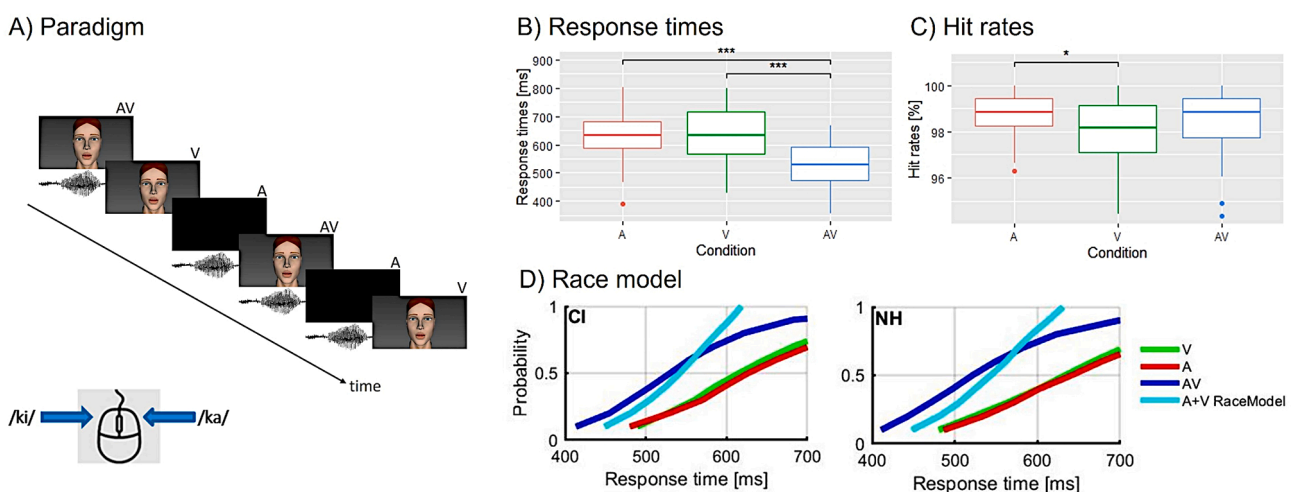


Fig. 1. Behavioural results. A) Simplified illustration of the paradigm. B) Mean response times for auditory (red), visual (green) and audiovisual (blue) syllables averaged over both groups, illustrating shorter response times for audiovisual syllables compared to auditory-only and visual-only RTs. C) Mean hit rates for auditory (red), visual (green) and audiovisual (blue) syllables averaged over both groups, exhibiting higher hit rates for auditory compared to visual syllables. D) Cumulative distribution functions for CI and NH. For both groups, the race model is violated, since both groups show that the probability of faster response times is higher for audiovisual stimuli (blue line) compared to the ones estimated by the race model (cyan line). Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(/ki/, /ka/), given that supplementary analyses revealed no group differences and no group-specific effects with regards to the syllable conditions (Supplementary Fig. 1). To compare the performance for each group and for each condition, the hit rates and RTs were separately entered into a 3×2 mixed ANOVA, with condition (AV, A, V) as the within-subjects factor and group (NH, CI) as the between-subjects factor. In case of violation of the sphericity assumption, a Greenhouse-Geisser correction was applied. Moreover, post-hoc t-tests were performed and corrected for multiple comparisons using a Bonferroni correction, if there were significant main effects or interactions ($p \leq .05$).

The effect of facilitation in RTs for audiovisual input (i.e. faster RTs for AV) compared to unimodal stimuli (A, V) is known as the redundant signals effect (Miller, 1982). There are two models proposing an explanation for this issue: the race model and the coactivation model. The race model postulates that no neural integration is required to get a redundant signals effect (Raab, 1962). There is a competition of the independent unimodal stimuli (A, V) which are leading to a 'race' that determines whether one or the other unimodal stimulus determines the RTs. Therefore, the RTs of redundant signals (AV) can be assumed to be faster due to statistical facilitation, since the probability of either of the stimuli (A and V) to show a fast RT is higher than from one stimulus (A or V) alone. Whereas the coactivation model (Miller, 1982) claims that there is an interaction of the neural responses of the single sensory stimuli of a pair resulting in a combination and creation of a new product before a motor response is started. This process leads to faster RTs. To examine whether faster RTs for audiovisual syllables resulted from statistical facilitation or were based on multisensory processes (coactivation), the race model inequality (Miller, 1982) was applied. This approach is widely used in the area of multisensory research. The assumption is that violation of the race model is evidence for the presence of multisensory interactions (Ulrich et al., 2007) and evidence against independent processing of each stimulus of a pair. According to the race model inequality, the cumulative distribution function (CDF) of the RTs in the audiovisual condition can never be larger than the sum of the CDFs of the unisensory (A, V) conditions:

$$P(RT_{AV} \leq t) \leq P(RT_A \leq t) + P(RT_V \leq t), \text{ for all } t \leq 0,$$

where $P(RT_x \leq t)$ is the probability of a condition $x \in \{AV, A, V\}$ to be lower than an arbitrary value t . For any given value of t , violation of the race model points to the existence of multisensory interactions (see also Ulrich et al. (2007) for further detailed information). We used the RMITest software by Ulrich et al. (2007) to prove violation of the race model inequality. For each participant, the CDFs of the RT distributions for each condition (AV, A, V) and for the sum of the modality-specific conditions (A + V) were estimated. To determine percentile values, the individual RTs were rank ordered for each condition (Ratcliff, 1979). Next, the CDFs for the modality-specific sum (A + V) and for the redundant signals conditions (AV) were compared for the five fastest deciles (bin width: 10 %) for each group separately (NH, CI). We used one-tailed paired t-tests with subsequent Bonferroni correction to account for multiple comparisons. Significance at any decile bin pointed to a violation of the race model suggesting multisensory interactions to occur.

Concerning the other behavioural measures, in particular the lip-reading test and a subjective rating of the listening effort during the EEG task, we calculated two-sample t-tests to assess differences between CI users and NH listeners.

2.6.2. EEG preprocessing

EEG data were analysed with EEGLAB (version v2019.1; Delorme and Makeig, 2004) which is running in the MATLAB environment (Mathwork, Natick, MA). The raw datasets were downsampled to 500 Hz and filtered using a FIR-filter with a high pass cut-off frequency of 0.5 Hz and a maximum possible transition bandwidth of 1 Hz (cut-off frequency multiplied by two). Additionally, we applied a low pass cut-

off frequency of 40 Hz and a transition bandwidth of 2 Hz. The Kaiser-window (Kaiser- $\beta = 5.653$, max. stopband attenuation = -60 dB, max. passband deviation = 0.001) was used for both filters (Widmann et al., 2015). This procedure maximises the energy concentration in the main lobe by averaging out noise in the spectrum and minimising information loss at the margins of the window. For CI users, the channels located around the region of the speech processor and transmitter coil were removed (mean: 3.2 electrodes; standard error of the mean: 1.0, range: 1–4). Next, the datasets were epoched into 2 s dummy epoch segments, and pruned of unique, non-stereotype artefacts using an amplitude threshold criterion of four standard deviations. An independent component analysis (ICA) was computed (Bell and Sejnowski, 1995) and the resulting ICA weights were applied to the epoched original data (1–40 Hz, -200 to 1220 ms relative to the stimulus onset (including the static and moving face). Independent components exhibiting horizontal and vertical eye movements, electrical heartbeat activity, as well as other sources of non-cerebral activity were removed (Jung et al., 2000). Similar to the procedures used in previous EEG studies with CI users (Debener et al., 2008; Sandmann et al., 2009; Viola et al., 2012), independent components accounting for the electrical CI artefact were identified and removed by means of ICA. The identification of these CI-related components was based on the stimulation side and the time course of the component activity, showing a pedestal artefact around 700 ms after the auditory stimulus onset (520 ms). Afterwards, the previously removed channels in CI users located around the speech processor and the transmitter coil were interpolated using a spherical spline interpolation (Perrin et al., 1989). This interpolation procedure enables a good dipole source localisation of auditory event-related potentials (ERPs) in CI participants (Debener et al., 2008; Sandmann et al., 2009). Only trials yielding correct behaviour (NH: 92.3 % $\pm$ 4.27 %; CI: 88.96 % $\pm$ 4.0 %) were retained for further analyses.

2.6.3. EEG data analysis

Event-related potentials (ERPs) of all conditions (AV, A, V) were compared between CI and NH participants. Similar to the behavioural data analysis, we focused on the ERP averages computed across the two syllables (/ki/, /ka/), given that supplementary analyses revealed no group differences and no group-specific effects with regards to the syllable conditions (Supplementary Fig. 1). To examine multisensory interactions, we used the additive model which is expressed by the equation $AV = A + V$ (Barth et al., 1995). If the processing of multisensory (AV) stimuli is the sum of the processing of unisensory (A, V) stimuli, the model is satisfied and suggests independent processing steps or interactions that are fully linear. By contrast, if the model is not satisfied, it is assumed that there are non-linear interactions between the sensory modalities (Barth et al., 1995). As a next step, the equation was rearranged to $A = AV - V$. This allowed to compare the recorded ERP response to auditory-only stimuli (A) and the term $[AV - V]$, reflecting the ERP difference wave that is calculated by subtracting the visual ERP from the audiovisual ERP response. The term $[AV - V]$ represents a visually-modulated auditory ERP response, which is the estimation of the auditory response in a multisensory context. If there is no interaction between the auditory and the visual modalities, the auditory (A) and the modulated auditory (AV-V) ERPs should be the same. By contrast, if the ERPs for A and AV-V are not equal, this can be interpreted as indication for non-linear multisensory interactions (Besle et al., 2004; Stekelenburg and Vroomen, 2007; Stevenson et al., 2014; Teder-Sälejärvi et al., 2005; Vroomen and Stekelenburg, 2010). The interpretation of such non-linear effects as either sub-additive ($A > AV - V$) or super-additive ($A < AV - V$) is not straightforward and requires measurements that are reference-independent, measurements of power, or measurements of source estimates (e.g. Cappe et al., 2010). To ensure that an equal number of epochs for each condition contributed to the difference wave (AV-V), we reduced the number of epochs according to the condition with the lowest number of epochs for each subject individually. This procedure reduced the number of the initial epochs to 88.04 % ($\pm$ 4.63

%). After reducing the epochs to the same amount for each condition, we computed the difference waves (AV-V).

In this study, we used an electrical neuroimaging (Michel et al., 2009) analysis framework, including topographic and ERP source analysis to compare auditory (A) and modulated (AV-V) ERPs between and within groups (NH, CI). We explored these ERP differences by analysing the global field power (GFP) and the global map dissimilarity (GMD) to quantify ERP differences in response strength and response topography, respectively (Murray et al., 2008). In a first step, we analysed the GFP, at the time window of the N1 and the P2 (N1: 80–200 ms; P2: 200–370 ms). The time windows were chosen based on visual inspection of the GFP computed for the grand average ERP. The GFP, which was first introduced by Lehmann and Skrandies (1980), equals the root mean square (RMS) across the average-referenced electrode values at a given instant in time or, more simply put, the spatial standard deviation of all electrodes at a given time (Murray et al., 2008). We deliberately chose the GFP in this study to analyse the cortical response strength without having the disadvantage that arbitrary channel selection may bias the results of the peak detection. Fig. 2A displays the ERPs for A and AV-V for CI users and NH listeners. The individual GFP peak mean amplitudes and latencies were detected (for each condition (A, AV-V) and time window (N1 and P2)) and were statistically analysed by using a 2×2 mixed ANOVA with group (NH, CI) as the between-subjects factor and condition (A, AV-V) as the within-subjects factor. This was done separately for each peak (N1, P2).

In a second step, we looked at the GMD (Lehmann and Skrandies, 1980) which quantifies topographic differences (and by extension, distinct configurations of neural sources; Vaughan Jr, 1982) between groups or experimental conditions, independently of the signal strength (Murray et al., 2008). The GMD was analysed in CARTOOL (Brunet et al., 2011) by computing a so-called 'topographic ANOVA' (TANOVA; Murray et al., 2008) to explore topographic differences between groups for each condition (CI(A) vs. NH(A) and CI(AV-V) vs. NH(AV-V)). Importantly, this is no analysis of variance, but a non-parametric

randomisation test. This randomisation test was conducted by using 5,000 permutations and by computing sample-by-sample p-values. An FDR-correction was applied to control for multiple comparisons (FDR = false discovery rate; Benjamini and Hochberg, 1995). The minimal significant duration was set to 15 consecutive time frames, given that previous observations suggest that ERP topographies do not continuously change as time elapses, but stay stable for a certain period of time before changing to another topography (Michel and Koenig, 2018) and to account for temporal autocorrelation.

2.6.4. Hierarchical clustering and single-subject fitting analysis

Our results revealed differences across groups for the GMD for each condition (see section 'ERP results on the sensor level: GMD' for more details), which is an indication for different underlying neural generators (e.g. Vaughan Jr, 1982). But, if this is the case, it is interesting to know, how these differences are caused. On the one hand, a GMD may be explained by the fact that CI users and NH listeners reveal a fundamentally different configuration of neural generators during audiovisual speech processing. On the other hand, a difference in topography can also result from a latency shift which causes similar topographic maps to be shifted in time (Murray et al., 2008). To distinguish between the two exploratory approaches, we performed a hierarchical topographic clustering analysis with group-averaged data (NH(A), NH(AV-A), CI(A), CI(AV-V)) to identify template topographies in the time windows of interest (N1, P2). This analysis was done in CARTOOL (Brunet et al., 2011). Specifically, we applied the atomize and agglomerate hierarchical clustering (AAHC) which has been devised for EEG-data by Murray et al. (2008). It includes the global explained variance of a cluster and hinders blind combinations (or agglomerations) of clusters with short durations. Hence, the topographic clustering finds the minimal number of topographies explaining the greatest variance in a given dataset (here CI(A), CI(AV-V), NH(A), NH(AV-V)). An ERP topography does not show a random variation across time, but rather stays stable for some time before changing to another topography. This empirical

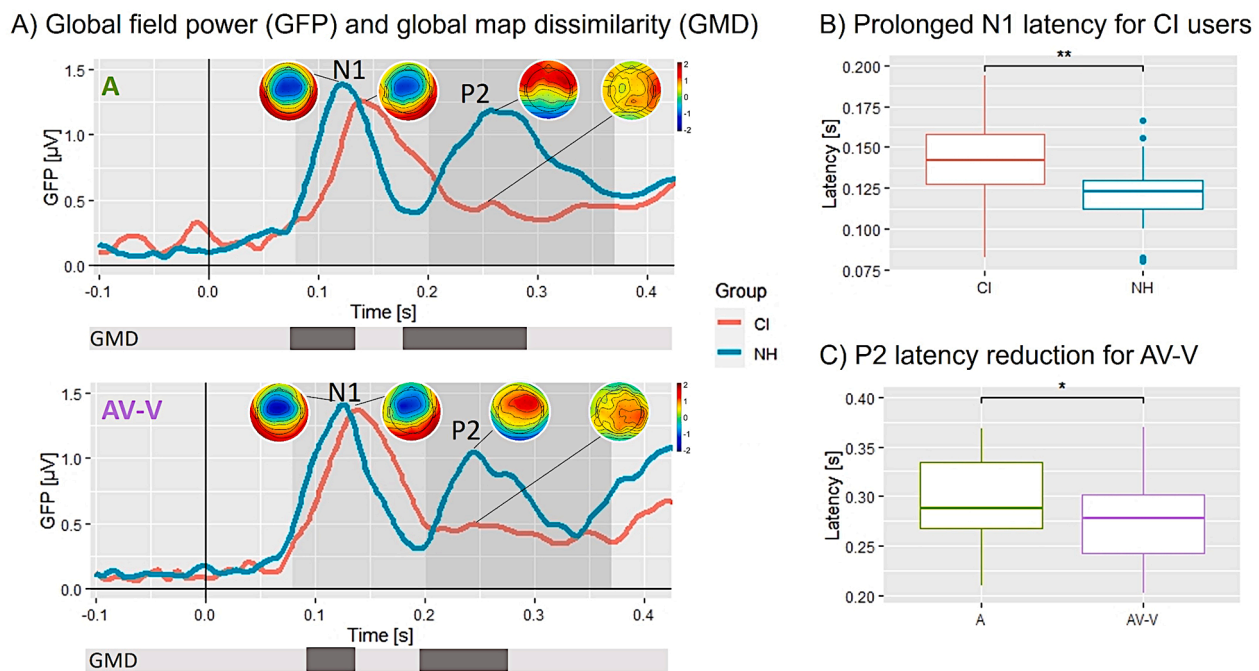


Fig. 2. Sensor ERP results. A) GFP of the conditions A and AV-V for NH listeners (blue) and CI users (red). Note that the GFP provides only positive values, given that it represents the standard deviation of all electrodes separately for each time point. The ERP topographies at the GFP peaks (N1, P2) are given separately for each group. Grey-shaded areas indicate the N1 and P2 time windows for peak and latency detection. Grey bars below mark the time window of significant GMDs between the two groups. B) N1 group effect (independent of the condition): the N1 latency is prolonged for CI users (red) compared to NH listeners (blue). C) P2 condition effect (independent of the group): the P2 latency is shorter for AV-V compared to A. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

observation has been referred to as microstates (Michel and Koenig, 2018). A more detailed description of this approach can be found in Murray et al. (2008).

Next, the template maps identified by the AAHC were submitted to a single-subject fitting (Murray et al., 2008) to see, how specific templates are distributed on a single-subject level. This was achieved by computing sample-wise correlations for each subject and condition between each template topography and the observed voltage topographies. Each sample was matched to the template map with the highest spatial correlation. We conducted a statistical analysis on the output, which was the *first onset of maps* (latency) and the *map presence* (number of samples in time frames) being assigned to a specific template topography. Specifically, we performed a mixed ANOVA with group (NH, CI) as the between-subjects factor and condition (A, AV-V) and template map as within-subject factors, separately for each time window (N1, P2). In the case of significant three-way interactions, group-wise mixed ANOVAs (condition $\times$ template map) were computed. In case of violation of the sphericity assumption, a Greenhouse-Geisser correction was applied. Follow-up t-tests were computed and corrected for multiple comparisons using a Bonferroni correction.

2.6.5. Source analysis

We computed an ERP source analysis for each group and condition by using Brainstorm (Tadel et al., 2011). This was done to assess the question, whether the observed topographic group differences at N1 latency can be explained by a fundamentally different configuration of neural generators during audiovisual speech processing. The source analysis was conducted according to the step-by-step tutorial provided by Stropahl et al. (2018). There are various proposed options to estimate source activities, but we chose the method of dynamic statistical parametric mapping (dSPM; Dale et al. 2000). This method has been successfully applied in previous studies with CI patients (Bottari et al., 2020; Stropahl et al., 2015; Stropahl and Debener, 2017). The dSPM method is able to localise deeper sources more precisely than standard norm methods, but the spatial resolution tends to remain low (Lin et al., 2006). dSPM uses the minimum-norm inverse maps with constrained dipole orientations to estimate the locations of the scalp-recorded electrical activity. It works well for localising auditory cortex sources, even though it is a relatively small cortical area (Stropahl et al., 2018). Single-trial pre-stimulus onset baseline intervals (−50 to 0 ms) were taken to calculate individual noise covariances to get an estimation of single-subject based noise standard deviations at each location (Hansen et al., 2010). As forward solution, the boundary element method (BEM) which is implemented in OpenMEEG was used as head model. The BEM gives three realistic layers and representative anatomical information (Gramfort et al., 2010). The activity data is shown as absolute values with arbitrary units based on the normalisation within the dSPM algorithm.

The inbuilt function to perform t-tests against zero (Bonferroni corrected, $p = .001$) was applied to the averaged dataset (averaged over conditions and groups) to obtain regions of interest (ROIs) based on the maximal source activity around the visual and auditory cortices in the time windows of interest (N1, P2). The predefined regions implemented in Brainstorm of the Destrieux-atlas (Destrieux et al., 2010; Tadel et al., 2011) were chosen based on the t-test as they were lying within this area. The *auditory* ROI was defined as a combination of four smaller regions from the Destrieux-atlas (Destrieux: G_temp_sup-G_T_transv, S_temporal_transverse, G_temp_sup-Plan_tempo and Lat_Fis-post) of which the first three regions have been used in e.g. Stropahl and Debener (2017) or Stropahl et al. (2015). Similarly, previous studies have reported activity in the auditory cortex during unisensory auditory processing in CI patients in Brodmann area 41 and 42 (Sandmann et al., 2009; Stropahl et al., 2018) which can be approximated by the chosen ROIs. As *visual* ROIs, three regions again from the Destrieux-atlas were chosen (Destrieux: G_cuneus, S_calcarine, S_parieto_occipital) that matched the results of the t-test against zero and which comprise the

secondary visual cortex. Similar visual regions have been identified in previous studies with CI users (Giraud et al., 2001a; Giraud et al., 2001b; Prince et al., 2021). The selected ROIs can be inspected in the boxes of Fig. 4A.

Single-subject source activities for each ROI, condition and group were exported from Brainstorm, and they were statistically analysed in the software R. First, the peak means and latencies of the peaks were calculated. We performed a mixed ANOVA separately for each time window of interest (N1: 80–200 ms, P2: 200–370 ms), with group (NH, CI) as between-subject factor and condition (A, AV-V), ROI (auditory, visual) and hemisphere (left, right) as within-subjects factors. If the sphericity assumption was violated, a Greenhouse-Geisser correction was applied. In the case of significant interactions or main effects, follow-up t-tests were computed and corrected for multiple comparisons using a Bonferroni correction.

3. Results

3.1. Behavioural results

In general, the participants achieved a very high-performance level (hit rates $\geq 97.9\%$) in all conditions, and the mean of the RT was between 522 ms and 640 ms in both groups (Table 2). Importantly, the CI users with a hearing-aid on the contralateral ear (implanted unilaterally) were no behavioural outliers in comparison to the bilaterally implanted CI users (mean of response times $\pm$ one standard deviation: unilaterally implanted: 589 ms $\pm$ 152 ms, bilaterally implanted: 597 ms $\pm$ 96.2 ms). The 3×2 mixed ANOVA with condition (AV, A, V) as within-subject factor and group (CI, NH) as between-subject factor showed no main effect of group ($F_{1,32} = 0.00039$; $p = .98$, $\eta^2 = 0.00001$), but a main effect of condition for RTs ($F_{1,41} = 69.18$; $p \leq .001$, $\eta^2 = 0.22$). Subsequent post-hoc t-tests revealed that RTs to redundant signals (AV) were significantly faster when compared to V ($t(33) = 15.4$; $p \leq .001$) or A ($t(33) = 10.9$; $p \leq .001$). There was no difference in RTs between the unisensory stimuli A and V ($t(33) = -0.2$; $p = 0.87$). These results are illustrated in Fig. 1B.

Concerning the hit rates, the 3×2 mixed ANOVA with condition (AV, A, V) as within-subject factor and group (CI, NH) as between-subject factor revealed a significant main effect of condition ($F_{2,64} = 3.8$; $p \leq .05$, $\eta^2 = 0.05$). Follow-up t-tests showed a difference in hit rates between A and V ($t(33) = 2.74$; $p \leq .05$), with A having significantly more correct hits than V. The other comparisons (AV vs. A and AV vs. V) were not significant (AV vs. A ($t(33) = 0.29$; $p = 0.88$) and AV vs. V ($t(33) = -1.89$; $p = 0.2$)). These results are illustrated in Fig. 1C.

To examine the violation of the race model, the race model inequality was used. The one sample t-tests were significant in at least one decile for each group (Table 3). This means that the probability of faster response times is higher for audiovisual conditions compared to the ones estimated by the race model. This can also be observed in Fig. 1D. In sum, the violation of the race model in CI users and NH listeners points out the presence of multisensory integration for both groups.

3.2. Other behavioural results

We calculated a two-sample t-test to assess differences in *auditory word recognition ability* and *(visual) lip-reading abilities* between CI users and NH listeners. The results revealed poorer auditory performance ($p \leq .001$, *Cohen's* $d = 2.48$), but better lip-reading skills for CI users compared to NH listeners ($p \leq .001$, *Cohen's* $d = 1.42$). Concerning the *subjective listening effort* measured during the EEG task, the t-test did not show a difference between CI users and NH listeners ($p \geq .05$, *Cohen's* $d = 0.19$). Therefore, none of the two groups found the task more difficult than the other group. The scores for these measures are displayed in Table 4.

A) 80–200 ms: Condition effect for CI users

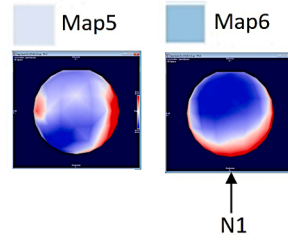
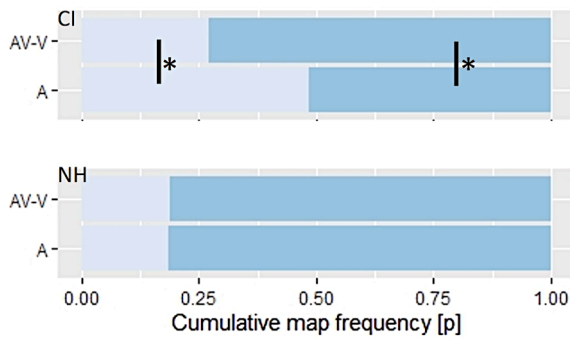
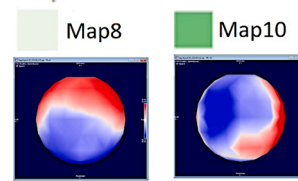
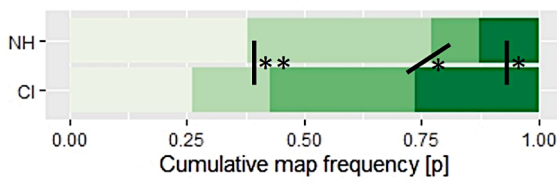
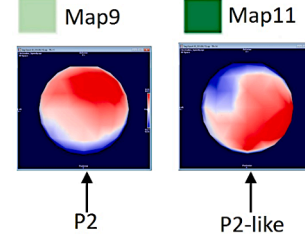
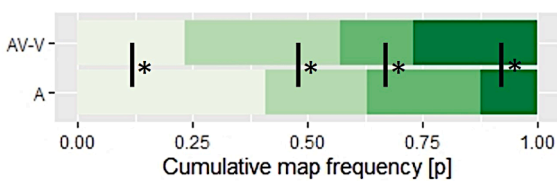


Fig. 3. Hierarchical clustering and single-subject fitting results. A) Cumulative map frequency of the N1 maps: the CI users, but not the NH listeners show a conditions effect, with more frequent N1 map presences for AV-V compared to A. The corresponding map topographies are displayed on the right side, with map 6 being referred to as the N1 topography. B) Cumulative map frequency of the P2 group effect: NH listeners reveal a more frequent presence of a P2 topography (map 9) compared to CI users. C) Cumulative map frequency of the P2 condition effect: there is an increase in the presence for P2 (map 9) and P2-like (map 11) topographies for AV-V compared to A. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

B) 200–370 ms: Group effect



C) 200–370 ms: Condition effect



3.3. ERP results on the sensor level: GFP

Fig. 2A displays the GFP of the grand averaged auditory ERPs for each group, in particular the unisensory auditory (A) and the visually modulated auditory (AV-V) ERP responses. The first prominent peak is visible for both groups approximately around 120 ms, which corresponds to the time window of the N1. The second peak around 220 ms is less prominent for CI users than for NH listeners and is referred to as the P2 ERP. For the N1, a 2×2 mixed ANOVA with group (NH, CI) as between-subject variable and condition (A, AV-V) as within-subject factor was calculated for the GFP amplitude peak mean and the GFP peak latency. Concerning the peak mean, there were no statistically significant interactions or main effects. But, for peak latency, there was a significant main effect of group ($F_{1,32} = 5.07$; $p \leq .05$, $\eta^2 = 0.13$) for both A and AV-V. Subsequent t-tests revealed a prolonged N1 latency in CI users than NH listeners ($t(33) = 2.78$; $p \leq .05$; Fig. 2B).

Similar to the N1 ERP analysis, a 2×2 mixed ANOVA has been applied for the P2 GFP peak mean and latency. Regarding the P2 peak mean, the mixed ANOVA showed no significant interactions or main effects. For the P2 latency, the mixed ANOVA revealed a significant condition main effect ($F_{1,32} = 8.75$; $p \leq .05$, $\eta^2 = 0.04$). Follow-up t-tests showed a significant difference between A and AV-V ($t(33) = 3.0$; $p \leq .05$) with shorter latencies for AV-V, indicating multisensory interaction effects to occur, independent of the factor group (Fig. 2C).

3.4. ERP results on the sensor level: GMD

Regarding the GMD, we computed sample-by-sample p-values to quantify differences in ERP topographies between groups and conditions. Comparing the CI users and NH listeners *separately for each condition* (CI(A) vs. NH(A) and CI(AV-V) vs. NH(AV-V)), the results revealed

significant group differences for each condition (A and AV-V), both at the N1 (A = 80–136 ms; AV-V = 92–136 ms) and the P2 (A = 184–290 ms; AV-V = 198–278 ms) latency range (Fig. 2A, grey bars beneath the GFP plots). Additionally, we compared the ERP topographies of the groups with regards to the *condition difference*, which was obtained by subtracting A from AV-V (CI(AV-V)-A vs. NH(AV-V)-A). This analysis revealed a significant group difference at the time window 156–174 ms, pointing towards group-specific audiovisual interaction processes.

3.5. ERP results on the sensor level: Hierarchical clustering and single-subject fitting results

To explore the origin of the underlying topographic difference between the two groups, we performed a hierarchical topographic clustering analysis with group-averaged data (CI(A); CI(AV-V); NH(A); NH(AV-V)) to identify template topographies in the time windows of interest (N1, P2). We applied the atomize and agglomerate hierarchical clustering (AAHC) to find the minimal number of topographies explaining the greatest variance in our dataset (i.e. the group-averaged ERPs). This procedure identified 12 template maps in 8 clusters that collectively explained 87.8% of these concatenated data. In particular, we observed two prominent maps for the N1 (map 5 and map 6) and four for the P2 (map 8, 9, 10, 11). These template maps were submitted to a single-subject fitting (Murray et al., 2008) to see, how the specific template maps are distributed on a single-subject level. In the following, *template map 6* will be referred to as N1 topography, since it matches the topography of the N1 peak from previous studies (Finke et al., 2016; Sandmann et al., 2015) and from our ERP results in both groups (Fig. 2A). The *template map 9* matches the P2 topography from previous studies (Finke et al., 2016a; Schierholz et al., 2021) and from NH

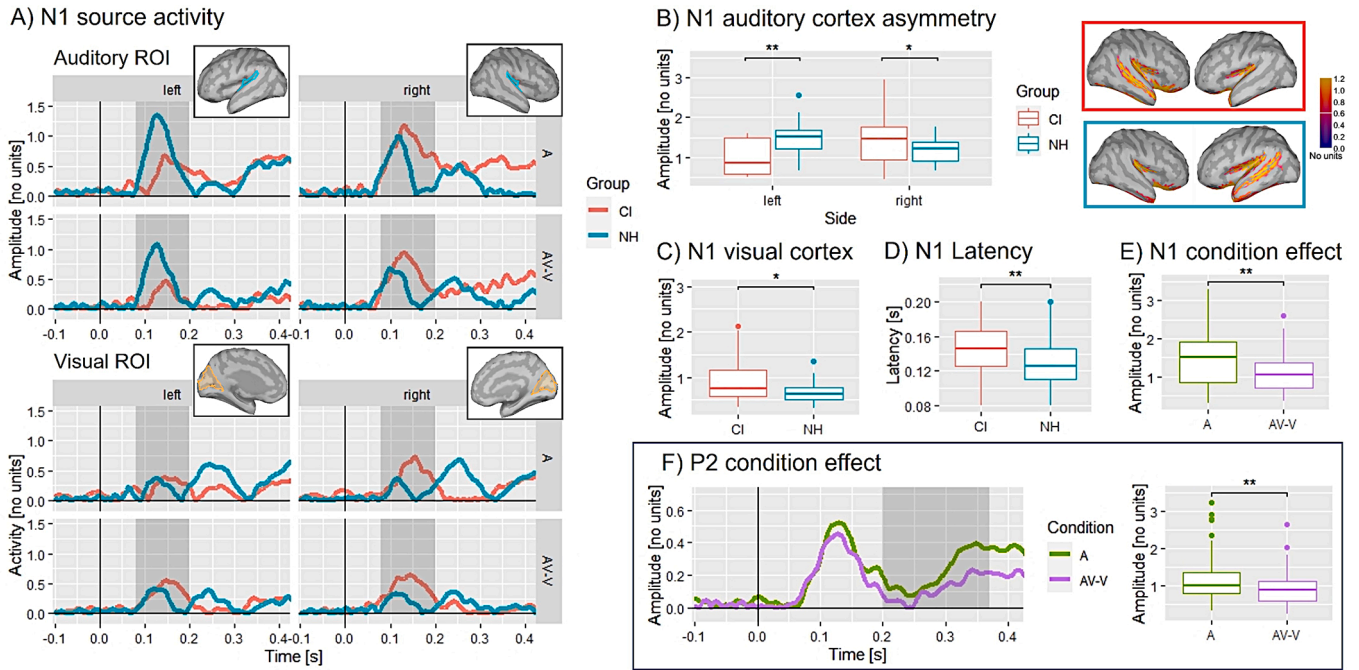


Fig. 4. Source analysis results. A) N1 source activity for CI (red) and NH (blue) for each ROI separated by hemisphere. The source activity is shown as absolute values and has arbitrary units based on the normalisation within the dSPM algorithm in Brainstorm. The grey shaded areas mark the N1 time window. The location of the defined ROIs is illustrated in the boxes with auditory ROIs in blue and visual ROIs in yellow. B) Hemisphere effect of the N1 peak mean: NH listeners (blue) show more activity in the left auditory cortex and CI users (red) show more activity in the right auditory cortex regardless of the condition. C) Group effect of the N1 peak mean: CI users show more activity in the visual cortex compared to NH listeners, regardless of the condition. D) N1 latency effect: CI users show a prolonged N1 latency compared to NH listeners, regardless of the condition. E) Condition effect of the N1 peak mean: there is a difference in activity between A and AV-V in the auditory cortex, indicating multisensory processes in both groups. F) P2 condition effect for the P2 peak mean: there is a significantly reduced auditory-cortex activation in AV-V compared to A, pointing towards cortical audiovisual interactions in both groups. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2

Mean hit rates (in %) and mean response times (in ms).

Condition	Hit rates		Response times	
	NH	CI	NH	CI
A	99 ± 0.8	98.5 ± 1	627 ± 81.1	632 ± 99.1
V	98.1 ± 1.3	97.9 ± 1.5	640 ± 102	623 ± 108
AV	98.5 ± 1.5	98.3 ± 1.4	522 ± 86.2	532 ± 86.3

Table 3

Redundant signals and modality-specific sum in each decile. AV corresponds to the redundant signals condition. A + V corresponds to the modality-specific sum. Paired-samples one-tailed t-tests with Bonferroni correction for multiple comparisons were conducted for each group. An asterisk indicates a significant result ($p \leq .05/5$).

Decile	NH			CI		
	AV	A + V	p	AV	A + V	p
0.10	411	448	0.000*	413	449	0.000*
0.20	443	481	0.000*	455	481	0.003*
0.30	471	505	0.000*	480	505	0.007*
0.40	497	524	0.002*	506	525	0.030
0.50	521	543	0.011	530	542	0.108

listeners in the current results (Fig. 2A), which is why this template map is henceforth referred to as the P2 topography. Accordingly, the *template map 11* is comparable to the P2 topography observed in our CI users (Fig. 2A below). Given that the CI users in the present study showed only a weakly pronounced P2 ERP, the template map 11 is called P2-like topography in the next section.

The observed topographic differences between groups and conditions (as described in section 'ERP results on the sensor level: GMD') could be due to a latency shift of ERPs and/or due to a fundamentally

Table 4

Other behavioural measures for CI users and NH listeners. A score of 100% means that all words have been repeated correctly both in the Freiburg monosyllabic test and the lip-reading test. Concerning the Nijmegen Cochlear Implant Questionnaire (NCIQ), a higher percentage value corresponds to a better quality of life. A higher value for the listening effort corresponds to more effort to perform the task.

Group	Freiburg test (%)	Lip-reading test (%)	NCIQ total (%)	Listening effort
NH	96.2 ± 5.7	15.8 ± 10.2	not applicable	11.8 ± 2
CI	73.8 ± 11.4	35.5 ± 16.4	69.7 ± 15.1	12.2 ± 1.9

different configuration of neural generators during speech processing. To distinguish between these two explanatory approaches, we focused on two output parameters in the hierarchical clustering analysis, in particular the *first onset of maps* as well as the *map presence*.

In a first step, we analysed the *first onset of maps* to address the question whether the observed differences in GMD can be explained by a latency shift. For this purpose, we performed a mixed-model ANOVA with group (NH, CI) as the between-subjects factor and condition (A, AV-V) and template map as the within-subject factors, separately for the N1 and P2 time windows. For the N1 maps, the three-way mixed ANOVA showed a significant group × map interaction ($F_{1,21} = 6.37$; $p \leq .05$, $\eta^2 = 0.1$). Post-hoc t-tests exhibited a difference between CI and NH for map 6 ($t(28) = 3.79$; $p \leq .001$). Given that template map 6 corresponds to a conventional N1 topography according to the distribution of the potentials on the scalp, these results confirm that the N1 of the CI users is generated later than the one of the NH listeners. For the first onset of the P2 maps, there were no significant main effects or interactions.

In a second step, we analysed the *map presence* (the number of time frames a corresponding template map is best-correlated to the single-subject data) to explore whether the topographic differences can be explained by group-specific patterns of ERP maps, which would point to distinct underlying neural generators between the two groups. Again, we performed a mixed-model ANOVA with group (NH, CI) as the between-subjects factor and condition (A, AV-V) and template map as the within-subjects factors, separately for the N1 and P2 time window.

The results revealed a group $\times$ map $\times$ condition interaction ($F_{1,32} = 9.1$; $p \leq .005$, $\eta^2 = 0.05$). Post-hoc t-tests revealed in particular for the group of CI users a significant difference of map presence between A and AV-V for the template map 5 ($t(16) = 3.42$; $p \leq .01$) and the template map 6 ($t(16) = -3.42$; $p \leq .01$). In CI users, the map 6 was similarly frequent to the map 5 in the auditory-only condition (A; number of samples 28.9 ± 20.6 vs. 27.1 ± 20.6), whereas the map 6 was more frequent compared to map 5 in the modulated condition (AV-V; number of samples 40.8 ± 15.6 vs. 15.2 ± 15.8). Thus, in CI users, the template map 6 better characterised the modulated response (AV-V) than the unisensory response (A). By contrast, the NH listeners in general showed a greater presence of map 6, irrespective of the condition (A: number of samples 45.6 ± 7.81 ; AV-V: number of samples 45.4 ± 8.81), and these individuals did not show a difference in the N1 ERP map presence between the modulated (AV-V) and the unisensory (A) responses ($t(16) = 0.12$; $p = .9$). Given that the template map 6 depicts a conventional N1 topography, the results suggest that specifically the CI users generate a N1 ERP map for the modulated response (AV-V) more frequently compared to the unisensory (A) condition. The comparison of groups by means of post-hoc t-tests revealed only a difference between CI users and NH listeners for the unisensory (A) ($t(16) = -2.79$; $p \leq .05$), but not for the modulated response ($t(16) = -1.42$; $p = .16$).

In sum, our results about the *first onset of maps* and the *map presence* at N1 latency suggest that the group differences in N1 topography originate from two reasons, in particular 1) generally slowed cortical N1 ERPs in CI users (regardless of condition), and 2) the distinct pattern of ERP maps in CI users as compared to NH listeners. The observation that the N1 topographies of the CI users differ from NH listeners especially in the auditory-only condition and that the two groups approximate in conditions with additional visual information, suggests that the CI users have a particularly strong visual impact on auditory speech processing. This visual impact on auditory speech processing is a manifestation of multisensory processes that remain intact in CI users despite the limited auditory signal provided by the CI.

Regarding the P2, the three-way mixed ANOVA revealed a significant group $\times$ map ($F_{2,77} = 4.35$; $p \leq .05$, $\eta^2 = 0.09$) and a condition $\times$ map ($F_{3,96} = 6.79$; $p \leq .001$, $\eta^2 = 0.06$) interaction. For the group $\times$ condition interaction, post-hoc t-tests revealed significant differences between NH and CI for three maps, in particular map 9 ($t(33) = -3.48$; $p \leq .001$), map 10 ($t(33) = 3.2$; $p \leq .01$) and map 11 ($t(33) = 2.34$; $p \leq .05$). NH listeners showed a conventional P2 topography (map 9) according to the distribution of the potentials on the scalp more often than CI users regardless of the condition (see Fig. 3B).

Following the condition $\times$ map interaction, post-hoc t-tests revealed significant differences between A and AV-V for all maps (map 8 ($t(33) = 3.11$; $p \leq .01$), map 9 ($t(33) = -2.1$; $p \leq .05$), map 10 ($t(33) = 2.27$; $p \leq .05$), map 11 ($t(33) = -2.69$; $p \leq .05$)). This result shows that P2 topographies (map 9) and P2-like topographies (map 11) are generated more often for modulated responses (AV-V) compared to unmodulated responses (A), which is illustrated in Fig. 3C.

In sum, our results about the *first onset of maps* and the *map presence* at P2 latency suggest group-specific topographic differences at P2 latency, with a stronger presence of a conventional P2 topography (map 9) in NH listeners, and a stronger presence of the slightly different P2-like topography (map 11) in CI users. Both of these maps 9 and 11 are more frequent in the modulated than in the auditory-only condition, pointing to changes in the cortical activation due to the presence of additional

visual information in the speech signal.

3.6. Results from ERP source analysis

A source analysis was conducted to further explore the observed group differences and to evaluate the visual modulation of auditory speech processing in different cortical areas. Based on previous findings of experience-related cortical changes in CI users (Chen et al., 2016; Giraud et al., 2001b; Giraud et al., 2001c; Green et al., 2005; Sandmann et al., 2015), we focused on the auditory and visual cortex activity in both hemispheres. Single-subject source activities for each ROI, condition and group were exported from Brainstorm and were statistically analysed. The source waveforms for the N1 are displayed in Fig. 4A, showing the response in the auditory cortex (CI mean = $141 \text{ ms} \pm 29 \text{ ms}$; NH mean = $125 \text{ ms} \pm 25 \text{ ms}$) and the visual cortex (CI mean = $147 \text{ ms} \pm 30 \text{ ms}$; NH mean = $137 \text{ ms} \pm 34 \text{ ms}$) for both groups. First, the peak means and latencies of the peaks were calculated, which were the dependent variables for the subsequent ANOVA. We performed a mixed-model ANOVA with group (NH, CI) as the between-subjects factor and condition (A, AV-V), ROI (auditory, visual) and hemisphere (left, right) as the within-subject factors for each time window of interest (N1, P2) separately.

For the N1 *peak mean*, the mixed ANOVA revealed a group $\times$ ROI $\times$ hemisphere interaction ($F_{1,32} = 8.44$; $p \leq .01$, $\eta^2 = 0.03$) and a ROI $\times$ condition interaction ($F_{1,32} = 16.81$; $p \leq .001$, $\eta^2 = 0.02$). For the group $\times$ ROI $\times$ hemisphere interaction, post-hoc t-tests confirmed a significant difference in the *left* auditory cortex activity between NH listeners and CI users ($t(28) = -4.04$; $p \leq .001$), with NH listeners showing stronger activation in the left auditory cortex than the CI users, regardless of condition. Whereas, post-hoc t-tests confirmed a significant difference in *right* auditory cortex activity between NH and CI ($t(28) = 2.03$; $p \leq .05$), with CI users showing stronger activation in the right auditory cortex than the NH listeners (Fig. 4B), regardless of the condition.

To resolve the ROI $\times$ condition interaction from the initial ANOVA, post-hoc t-tests confirmed a difference between A and AV-V for the *auditory cortex* ($t(67) = 5.71$; $p \leq .001$), with more activity for A compared to AV-V, pointing towards multisensory interaction processes in both groups (Fig. 4E), regardless of the hemisphere.

For the *visual cortex*, post-hoc t-tests showed a significant difference in peak mean between NH listeners and CI users ($t(67) = 3.6$; $p \leq .001$), with CI users exhibiting a stronger activation in the visual cortex than NH listeners (Fig. 4C), regardless of the condition and hemisphere.

Concerning the N1 *latency*, the four-way mixed ANOVA showed a main effect of group ($F_{1,32} = 7.39$; $p \leq .05$, $\eta^2 = 0.05$) and ROI ($F_{1,32} = 4.89$; $p \leq .05$, $\eta^2 = 0.02$). Post-hoc t-tests revealed a difference in N1 latency between NH listeners and CI users ($t(135) = 3.82$; $p \leq .001$) and a difference between the auditory and visual cortex ($t(135) = -2.53$; $p \leq .05$). Again, the results confirmed that the N1 is generated later in CI users compared to NH listeners regardless of the condition and hemisphere (Fig. 4D).

Similar to the N1 ERP, a group (NH, CI) $\times$ condition (A, AV-V) $\times$ ROI (auditory, visual) $\times$ hemisphere (left, right) ANOVA was computed on the peak mean and latency of the P2. Regarding the P2 *peak mean*, the mixed ANOVA revealed a main effect of ROI ($F_{1,32} = 18.31$; $p \leq .001$, $\eta^2 = 0.08$) and condition ($F_{1,32} = 18.97$; $p \leq .001$, $\eta^2 = 0.04$). Post-hoc t-tests showed a difference between the auditory and visual cortex ($t(136) = 5.98$; $p \leq .001$), with more activity for the auditory cortex compared to the visual cortex, and a significant difference in peak mean between A and AV-V ($t(136) = 5.25$; $p \leq .001$), indicating multisensory interaction in both groups. This multisensory effect at P2 latency is illustrated in Fig. 4F, providing the source waveforms (and the corresponding box plot) separately for the auditory-only (A) and the modulated (AV-V) response, averaged across the two groups.

Regarding the P2 *latency*, we found no statistically significant

interactions or main effects for the P2 latency.

3.7. Correlations

We performed targeted correlations matching findings from previous studies using the Pearson's correlation. The Benjamini-Hochberg (BH) procedure was applied to account for multiple comparisons (Benjamini and Hochberg, 1995). First, we verified that lip-reading abilities are related to the CI experience and the age at onset of hearing loss (Stropahl et al., 2015; Stropahl and Debener, 2017). The results revealed a positive relationship between lip-reading abilities and CI experience ($r = .51$; $p = 3.8e-2$; significant according to BH), as well as a negative relationship between lip-reading abilities and the age at onset of hearing loss ($r = -.73$; $p = 9.5e-4$; significant according to BH). This means that the longer the CI experience and the earlier the onset of hearing impairments, the more pronounced are the lip-reading abilities in CI users. Second, we explored whether there is a relationship between CI experience and the activation in the visual cortex (Giraud et al., 2001c). Indeed, there was a positive correlation between CI experience and visual cortex activation at N1 latency ($r = .68$; $p = 2.9e-3$; significant according to BH), indicating that a longer experience with the CI results in enhanced recruitment of the visual cortex during auditory speech processing. Moreover, previous studies reported a positive relationship between left-auditory cortex activation and speech comprehension, which however was not confirmed by the present results (Freiburg monosyllabic test: $r = .27$; $p = .29$; not significant according to BH). Finally, we correlated the NCIQ with the Freiburg monosyllabic test (Vasil et al., 2020) and found a strong positive relationship ($r = .75$; $p = 4.9e-4$; significant according to BH), demonstrating that the quality of life is rising with better hearing performance.

4. Discussion

The present study investigated audiovisual interactions in CI users and NH listeners by means of EEG and behavioural measures. On the behavioural level, the results showed multisensory integration for both groups, as indicated by shortened response times for the audiovisual as compared to the two unisensory conditions. A multisensory effect was confirmed by the ERP analyses for both groups, showing a reduced activation in the auditory cortex for the modulated (AV-V) compared to the auditory-only (A) response, both at the N1 and the P2 latency. Nevertheless, we found specifically in the group of CI users a change of N1 voltage topographies when visual information was presented in addition to auditory information, which resulted in an approximation of the N1 topographies of the two groups. Thus, our behavioural and ERP findings demonstrate a clear audiovisual benefit in both groups, with a particularly strong visual impact on auditory speech processing in the CI users.

Apart from these condition effects, we found differences between CI users and NH listeners, not only in ERP topographies (N1 and P2 latency) but also in auditory- and visual-cortex activation (N1 latency). Regarding the latter, the ERP source analyses revealed that the CI users, when compared to the NH listeners, generally have an enhanced recruitment of the *visual cortex* at N1 latency, which is positively correlated with the duration of CI experience. Further, the ERP source analyses revealed a delayed N1 response in CI users for the *auditory cortex*, and a group-specific pattern of activation, with a leftward hemispheric asymmetry in the NH listeners and a rightward hemispheric asymmetry in the CI users.

Taken together, these results suggest that the topographic group differences observed at N1 latency are caused by two different factors, notably the generally slowed cortical N1 response in CI users, and the group-specific differences in cortical source configuration.

4.1. Multisensory integration in both groups – behavioural level

Behavioural measures revealed that both the NH listeners and the CI users showed shorter reaction times for audiovisual syllables compared to unisensory (auditory-alone, visual-alone) syllables. There was no difference between auditory and visual conditions. Therefore, both groups show a clear redundancy effect for audiovisual syllables, suggesting that on the behavioural level, the benefit for cross-modal input is comparable between the CI users and the NH listeners (Laurienti et al., 2004; Schierholz et al., 2015), at least in conditions with short syllables that are combined with a computer animation of a talking head. Even though the CI provides only a limited input, the CI users' responses were not delayed on the behavioural level. In line with this, there was no difference in subjective ratings of listening effort between the two groups, indicating that both groups were able to perform the task with comparable listening effort.

Concerning purely visual syllables, we expected to find differences between CI users and NH listeners with shorter reaction times for CI users compared to NH listeners. This expectation is based on results from studies showing shorter reaction times for visual stimulation in congenitally deaf individuals (Bottari et al., 2014; Finney et al., 2003; Hauthal et al., 2014) and visually-induced activation in the auditory cortex, both in deaf individuals and CI users (Bavelier and Neville, 2002; Bottari et al., 2014; Heimler et al., 2014; Sandmann et al., 2012). This cross-modal reorganisation seems to be induced by auditory deprivation and might provide the neural substrate for enhanced superior visual abilities. CI users experience not only a time of deprivation before implantation, but they also perceive a limited input provided by the implant, which may force these individuals to develop compensatory strategies. Indeed, previous studies have revealed that CI users show specific enhanced behavioural visual abilities, in particular improvements in (visual) lip-reading skills when compared to NH listeners (Rouger et al., 2007; Schreitmüller et al., 2018; Stropahl et al., 2015; Stropahl and Debener, 2017).

Nevertheless, the behavioural results on *syllables* in the present study show comparable results between CI users and the NH listeners, which resembles previous results from a speeded response task with basic stimuli, presenting tones and white discs as auditory and visual stimuli, respectively (Schierholz et al., 2015). The authors argued that CI users do not have better reactivity to visual stimuli compared to NH listeners when performing a simple detection task. However, the present task required detection *and* discrimination of the syllables. But, since there was only a choice of two syllables that had to be discriminated, the task might have been too easy to detect differences between the two groups, leading to ceiling effects (hit rates $\geq 97.9\%$). This argument is supported by our finding that CI users did *not* report an enhanced listening effort when compared to NH listeners, which contradicts previous studies that used more complex speech stimuli and which reported an enhanced listening effort in CI users (Finke et al., 2016a; Finke et al., 2016b). Nevertheless, similar to our study, Stropahl and Debener (2017) used a syllable identification task to compare CI users and NH listeners. They did not find behavioural group differences for visual syllables as well, although CI users in general showed enhanced lip-reading abilities, as in the present study. It seems that behavioural improvements in CI users are stimulus- and task-selective, and may be observed only under specific conditions and might be more pronounced if speech stimuli include semantic information (Moody-Antonio et al., 2005; Rouger et al., 2008; Tremblay et al., 2010).

A benefit for cross-modal input, which is represented by a redundant signals effect for audiovisual stimuli, does not necessarily confirm multisensory integration. Faster reaction times for audiovisual stimuli might also result from statistical facilitation due to a competition between the unimodal auditory or visual stimuli, leading to a 'race' that determines the reaction times (Miller, 1982). To test whether statistical facilitation or coactivation were the cause for the redundant signals effect, we applied the race model inequality (Miller, 1982). Given that

the race model was violated for both the NH listeners and the CI users over the faster deciles of the RT distribution, we can conclude that audiovisual integration was contributing to the observed redundant signals effect in both groups.

Our results on syllable perception revealed that the audiovisual gain in CI users and NH listeners was comparable. However, Rouger et al. (2007) used a word recognition task and showed for the CI users an enhanced gain in word recognition performance when the visual and auditory input was combined. The authors concluded that CI users are 'better audiovisual integrators' when compared to NH listeners. The inconsistency to our results might originate from the difference of demands in the task. As argued before, the task in our experiment and in the previous one with tones (Schierholz et al., 2015) was probably too easy to detect differences between the groups on the behavioural level. Schierholz et al. (2015) suggested that audiovisual benefits in CI users might be more pronounced in more ecologically valid stimulus conditions, since CI users might need their compensatory strategies rather for more complex speech stimuli.

However, it seems that syllables are not yet complex enough to show a group difference in multisensory enhancement between CI users and NH listeners. As both groups performed with a very high hit rate, a difference was not detected due to ceiling effects. We suppose that experimental designs with even more complex speech stimuli, such as words or sentences, are necessary to reveal behavioural differences in audiovisual speech processing between CI users and NH listeners.

4.1.1. Multisensory interactions in both groups - cortical level

To analyse the influence of the visual input on the auditory cortical response, we compared the ERPs to the auditory condition (A) with the modulated ERP response (AV-V). If there is a difference between the auditory and the modulated response, this can be interpreted as indication for multisensory interactions (Besle et al., 2004; Stekelenburg and Vroomen, 2007; Teder-Sälejärvi et al., 2005; Vroomen and Stekelenburg, 2010). In that case, the difference can be either subadditive ($A > AV - V$) or superadditive ($A < AV - V$) (Stevenson et al., 2014). In multisensory research, both subadditive results (e.g. Cappe et al., 2010) and superadditive results (e.g. Schierholz et al., 2015) have been reported for ERPs. Regarding the current study, the source analyses of the N1 and P2 ERPs revealed a reduced cortical activation for the modulated (AV-V) compared to the auditory-only (A) ERPs. Thus, our results point to a subadditive affect in both the CI users and the NH listeners, which is present at different processing stages, specifically at N1 and P2 latency.

Further evidence of cortical multisensory interaction in both the CI users and the NH listeners comes from our analysis of the global field power (GFP) and the topographic clustering analysis. Regarding the GFP, we found that the peak in the P2 time window was shortened for the modulated (AV-V) compared to the auditory-alone (A) responses. This observation confirms previous conclusions that a reduction in latency can be a sign for multisensory integration processes (van Wassenhove et al., 2005). Regarding the topographic clustering analysis, we found an increased presence of the P2 and P2-like topographies for the modulated responses (AV-V) compared to auditory responses (A). This effect was not specific to one group but was observed for both the CI and the NH individuals. Thus, our findings point to a visual modulation of auditory ERPs in both the CI users and the NH listeners, which confirms our behavioural results of multisensory integration in both groups.

Similar to the present study, Schierholz et al. (2015) investigated audiovisual interactions by applying the additive model to ERP data. They showed a cortical audiovisual interaction as well, as indicated by a significant visual modulation of auditory cortex activation at N1 latency, specifically for elderly CI users as opposed to young CI users and elderly NH individuals. In our study however, the visual modulation in cortical activation was present in both the NH and the CI group. This discrepancy in results may result from the fact that we did not include two age groups in our study, as it has been done in Schierholz et al. (2015).

4.2. Group differences in auditory and audiovisual processing

There was no difference in global field power (GFP) between CI users and NH listeners, indicating that the strength of cortical responses was comparable between the two groups. However, there were group differences in ERP topographies as shown by the global map dissimilarity (GMD). In particular, GMD group differences were found not only for the auditory-only condition (A) and the modulated response (AV-V), but also for the difference wave ((AV-V)-A).

In sum, these results suggest that differences between groups did not result from differences in signal strength, but rather from differences in topography. A distinct electric field topography however can be caused by a latency shift of ERPs and/or by distinct configurations of the neural sources. We performed different analyses, including hierarchical clustering, single-subject fitting and source analysis, to distinguish between these two causes for topographic group differences. The following sections discuss the results of these analyses.

4.2.1. P2 time window

Hierarchical clustering and single-subject fitting analyses for the P2 time window revealed a condition effect (independent of group) and a group effect (independent of condition). The *condition* effect, showing a more frequent map presence for P2 and P2-like topographies has been discussed already in the section 4.1.1. Concerning the *group* effect, we found a more frequent presence for P2 topographies for NH listeners compared to CI users and more frequent P2-like topographies for CI users compared to NH listeners. But, as a whole, four template maps have been identified by the clustering analysis, showing variable distributions across groups and conditions. This leads to the conclusion that group differences are relatively probable but are not easy to depict due to the high variability in the data. Therefore, it is unlikely that a source analysis will find clear results to reveal group-specific cortical activation patterns. Indeed, the source analysis did not reveal group differences, neither in the auditory, nor the visual cortex activation. However, there was an amplitude reduction for AV-V compared to A over both groups, confirming audiovisual interactions but no differences between CI users and NH listeners (see section 4.1.1).

It may be argued that our selected auditory ROI was not appropriate enough for P2 responses, which could have prevented finding a group difference in auditory-cortex activation. To verify the ROI selection, we conducted a supplementary source analysis for the P2 ERP with an adjusted, smaller ROI that was selected by considering previous studies. Specifically, previous studies using MEG (magnetoencephalography) or fMRI (functional magnetic resonance imaging) have reported that the auditory P2 is generated in the anterior part of the auditory cortex, in particular the lateral part of the Heschl's gyrus (Bosnyak et al., 2004; Lütkenhöner and Steinsträter, 1998; Ross and Tremblay, 2009; Tremblay et al., 2014). This region is also implemented in the Destrieux-atlas in Brainstorm (G_temp_sup-G_T_transv), a subregion of our originally selected auditory ROI. Statistical analyses did not show any different results between the adjusted (i.e. smaller) and the original ROI. Therefore, we conclude that CI users and NH listeners show a comparable recruitment of the auditory cortex during auditory and audiovisual speech processing, at least at the P2 latency range.

4.2.2. N1 time window

Hierarchical clustering and single-subject fitting results for the N1 time window demonstrated a more frequent presence of N1 topographies for the modulated responses (AV-V) compared to purely auditory responses (A) for CI users, but not for NH listeners. Specifically for CI users, there was a difference between A and AV-V, showing a more frequent map 6 (N1 topography) and a less frequent map 5 presence (blurred N1 topography) for modulated responses and an equal map presence of map 5 and map 6 for purely auditory responses. By contrast, the NH listeners showed no differences in map presence between auditory and modulated responses, hinting towards similar neural processes

to evaluate these stimuli.

Regarding the CI users, the increased presence of a conventional N1 topography (map 6) specifically for the modulated response indicates alterations in audiovisual processing and a multisensory benefit for CI users, if additional visual information is provided. It is likely that these cortical alterations in CI users reflect compensatory strategies, which CI users develop to overcome the limited CI input. This compensatory processing strategy may allow the CI users to better use the visual input to reach a processing level closer to the one of the NH listeners. This is confirmed by direct group comparisons showing differences between NH listeners and CI users only for the auditory, but not for the modulated response. Overall, as suggested by [Stevenson et al. \(2014\)](#), hierarchical clustering with subsequent single-subject fitting is a powerful approach to study multisensory interactions. This was confirmed in this study, since this method allowed us to demonstrate a clear audiovisual benefit for CI users when they were provided with additional visual input. We suggest that further studies using more complex speech stimuli should apply this approach as well to study multisensory interactions in CI users. As with behavioural results ([Moody-Antonio et al., 2005](#); [Rouger et al., 2008](#); [Tremblay et al., 2010](#)), it may be speculated that cortical audiovisual interactions in CI users are more pronounced if speech stimuli with semantic information are used.

On the sensor level, ERP data showed a prolonged N1 latency for CI users compared to NH listeners, for both the modulated and the auditory-only responses. Hierarchical clustering and single-subject fitting analyses confirmed the delayed N1 topography in CI users compared to NH listeners by assessing the first onset of map 6 (N1 topography). On the source level, again, the source activity results confirmed a delayed N1 response for CI users compared to NH listeners for both auditory and visual cortices. In fact, one might assume that the time until the electrical signal reaches the auditory nerve with a CI is faster than in NH listeners, since the signal does not have to undergo the processing through the middle and inner ear due to the direct stimulation of auditory nerve fibres. Interestingly, in the study of [Schierholz et al. \(2017\)](#), delayed cortical responses have been found not only in patients with a CI but also in patients with an auditory brainstem (ABI; direct stimulation of the cochlear nucleus) and an auditory midbrain implant (AMI; direct stimulation of the inferior colliculus). Although central auditory implants bypass more structures than the CI, ABI and AMI patients revealed even more delayed cortical responses and poorer behavioral abilities when compared to CI patients ([Schierholz et al., 2017](#)). Given that ABI and AMI are not yet optimised in electrode array designs and processing strategies ([Lim and Lenarz, 2015](#); [McKay et al., 2013](#)), the remarkable delay of the cortical response seems to be particularly related to the insufficient input provided by the central auditory implants ([Schierholz et al., 2017](#)). Similarly, the auditory input provided by the CI is limited, and the delayed N1 ERP responses in CI users, when compared to NH listeners, may reflect enhanced difficulties to process speech sounds, resulting in more time to encode the stimuli. Indeed, previous studies with NH listeners have shown that difficult acoustic listening conditions, for instance speech in background noise, cause a delay in the N1 response ([Billings et al., 2011](#); [Finke et al., 2016a](#)). Moreover, our results of a slowed cortical N1 ERP in CI users are consistent with several other studies, which used purely auditory stimuli of different acoustic complexity in unilaterally stimulated CI users ([Beynon et al., 2005](#); [Finke et al., 2016a](#); [Sandmann et al., 2009, 2015](#)) and in bilaterally stimulated CI users ([Senkowski et al., 2014](#)).

Taken together, our results revealed group differences for the N1 ERP. First, the hierarchical clustering and single-subject fitting analyses revealed a *condition effect* specifically for the group of CI users, showing distinct patterns of N1 topography between the auditory-only (A) and the modulated auditory response (AV-V; [Fig. 3A](#)). By contrast, the NH listeners showed no differences in map presence between auditory and modulated responses, hinting towards similar neural processes to evaluate these stimuli ([Fig. 3A](#)). However, the CI-specific condition effect points to a particularly strong visual impact in CI users on auditory

speech processing, which may allow these individuals to compensate for the degraded CI input. Apart from this CI-specific condition effect, our results revealed a *map dissimilarity at N1 latency between the two groups*, independent of the condition. The topographic and source analyses revealed that these group differences at N1 latency can be explained by a delayed cortical N1 response in CI users when compared to NH listeners. However, as discussed in the next two sections (4.2.3 and 4.2.4), the source analyses also revealed that this map dissimilarity at N1 latency is additionally caused by a distinct configuration of neural sources. When compared to the NH listeners, the CI users showed an additional activation in the visual cortex and a delayed cortical activation with a reversed, rightward functional lateralisation in the auditory cortex (see sections 4.2.3 and 4.2.4 for more details).

4.2.3. Intra-modal plasticity in the auditory cortex

The source analysis of cortical activity revealed hemispheric differences between NH listeners and CI users, with NH listeners showing more activity in the left auditory cortex and CI users showing more activity in the right auditory cortex. There are two possible explanations for these results.

The first explanation is associated with the CI-related stimulus degradation, which affects the basic acoustic properties of the presented syllables. Indeed, many CI users reported that they did not perceive the syllables as syllables, but rather as distinct sounds. Therefore, it is conceivable that the CI users were able to discriminate the different stimuli, although the sounds were not identified as speech sounds. The two hemispheres of the brain have been found to be specialised for the processing of distinct basic acoustic properties ([Belin et al., 2000](#); [Poeppel, 2003](#); [Zatorre and Belin, 2001](#)). The left auditory cortex is involved in speech processing, as it preferentially processes fast temporal cues, which are largely contained in speech signals ([Boemio et al., 2005](#)). Whereas the right auditory cortex has been found to be more involved in the processing of spectral aspects of sounds and slowly modulated temporal acoustic cues, which is a property of musical stimuli ([Belin et al., 2000](#); [Poeppel, 2003](#); [Zatorre and Belin, 2001](#)). Thus, the hemispheric differences during speech and music processing can be attributed to the relative specialisation of the two hemispheres for basic acoustic stimulus properties, in particular fast temporal versus slow spectrotemporal acoustic cues. For this reason, one can assume that there must be more involvement of the left auditory cortex for processing syllables, which is the case for NH listeners. For CI users however, the present results show more recruitment of the right auditory cortex. It is likely that this hemispheric difference was caused by the CI-related syllable degradation, resulting in alterations of basic acoustic properties of the CI signal and leading to a relatively enhanced activation in the right auditory cortex. This interpretation is confirmed by our finding that CI users rather perceived the syllables more as meaningless sounds and not as linguistic input.

The second explanation for group differences in hemispheric asymmetries is related to cortical reorganisation, which might be caused by deafness before implantation and subsequent fitting with a CI. In the study of [Sandmann et al. \(2009\)](#) this claim has been investigated by comparing processing of NH listeners for normal and noise-vocoded stimuli to simulate the hearing experience of a CI patient. The results revealed no difference between the normal and vocoded responses in NH listeners, providing evidence against pure stimulus degradation effects. The authors rather suggest that functional differences arose due to a distinct hearing experience as a consequence of cortical plasticity. Our CI users experienced profound deafness, and through the supply of a CI, the hearing was restored but was still artificial and highly different compared to normal acoustic hearing. This process leads to experience-related changes in the normal pattern of response asymmetry. A previous study with unilaterally-implanted CI users ([Mortensen et al., 2006](#)) suggests that it is important to maintain the left auditory cortex for auditory processing to show more success with a CI. This study found associations for speech-induced left superior temporal gyrus activation

with high levels of speech comprehension in CI users. Accordingly, Giraud et al., (2001a) demonstrated enhanced activation in the left auditory cortex for phonology-specific stimuli for NH listeners compared to unilaterally-implanted CI users, which was interpreted as decreased left hemispheric specialisation for the processing of speech in CI users. This is in line with other studies on patients with profound hearing loss (Fujiki et al., 1998; Ponton et al., 2001) and studies with CI patients showing cortical reorganisation after cochlear implantation (Green et al., 2005; Pantev et al., 2006).

In the present study, we can report group differences in auditory cortex asymmetries as well, as our groups were carefully matched in stimulation side. Since all participants were right-handed, we can exclude lateralisation effects caused by handedness. However, we cannot exclude that differences happened due to the side of implantation, as we could not split up our groups into right-implanted and left-implanted patients due to the limited sample size. Further studies are required to systematically disentangle the effects of implantation side and cortical asymmetries in CI users for different kinds of stimulation (speech and non-speech).

4.2.4. Cross-modal plasticity in the visual cortex

The source analysis showed a stronger activation in the visual cortex for CI users compared to NH listeners, hinting towards cross-modal activation of the visual cortex. This confirms the results from previous studies, reporting that cortical changes in CI users are not restricted to the auditory cortex (Campbell and Sharma, 2013, 2016; Chen et al., 2016; Giraud et al., 2001a; Giraud et al., 2001b), but are also present in visual and even in inferior frontal areas (Rouger et al., 2012).

Previous results have suggested that CI users compensate the limited auditory input provided by the CI by additionally recruiting the visual cortex (Doucet et al., 2006; Giraud et al., 2001c; Strelnikov et al., 2010, 2013). Giraud et al. (2001c) reported in their PET (positron emission tomography) study greater auditory activation in the visual cortex for meaningful speech sounds in unilaterally-implanted CI users compared to NH listeners. This additional visual activity was associated with enhanced speech-reading abilities and longer CI experience (Giraud et al., 2001c). The authors suggest that the activation of the visual cortex for auditory processing is present due to an enhanced synergy of the auditory and visual modalities in CI users. It is conceivable that this audiovisual synergy helps recovering speech comprehension after cochlear implantation, when CI users need to learn to match the novel auditory speech inputs with the corresponding visual speech cues (Strelnikov et al., 2013). Indeed, the finding of a positive correlation between visual-cortex activation and the duration of CI experience indicates that the visual modality helps to restore auditory speech perception. This is in line with our results, demonstrating a positive relationship between CI experience and visual cortex activation, supporting experience-related plastic changes after implantation.

Up to now, the cross-modal recruitment of the visual cortex for processing auditory stimuli in CI users has been observed for syllables, words, environmental sounds, pure tones and reversed words (Chen et al., 2016; Giraud et al., 2001a; Giraud et al., 2001c). Thus, cross-modal activation of the visual cortex is not restricted to a semantic content of the auditory stimulus (Chen et al., 2016), but seems to be highly beneficial with regards to behavioural abilities, such as lip-reading (Rouger et al., 2007; Stropahl et al., 2015; Stropahl and Debener, 2017) and audiovisual word integration (Rouger et al., 2007).

Our study is the first EEG study showing visual recruitment for auditory (and audiovisual) stimuli and a positive correlation of CI experience with visual cortex activation in CI users, extending existing evidence in this direction. This confirms that EEG is a suitable tool to study cortical reorganisation in CI users with the advantage of the highly precise temporal resolution allowing to track single processing steps. Here, we found differences in cortical activation between CI users and NH listeners already at the initial cortical processing stage of the N1 time window (approximately 100 ms post stimulus onset).

4.3. Clinical implications

Given that life expectancy and the exposure to noise are rising, hearing problems will become increasingly prevalent in the ageing population (Sun et al., 2021). Therefore, treatment methods have to be optimised, as hearing loss is an important factor influencing the quality of life (confirmed here with the NCIQ). As both auditory and visual input (lip movements, gestures) contribute to speech perception, spoken language communication is an audiovisual task (Drijvers and Özyürek, 2017; Grant et al., 1998; Sumby and Pollack, 1954). Especially in difficult hearing situations (for instance speech in background noise) the supportive role of the visual modality can be crucial for understanding speech (Sumby and Pollack, 1954). Enhanced lip-reading and audiovisual integration abilities may support advantages in perceiving speech in ecologically valid listening situations comparable to real-life situations including both modalities (Rouger et al., 2007, 2012; Stropahl et al., 2015, 2017). Therefore, it is essential to not only investigate auditory or visual processing in CI users separately, but to focus on the interaction of the two modalities. Currently, the clinical assessment of speech perception in CI users is typically restricted to purely auditory methods. Given the importance of visual input in real-life situations, this approach is insufficient for gauging the actual capabilities of CI patients (Stevenson et al., 2017). To further optimise treatment options for CI patients, it is important to develop audiovisual, ecologically valid assessment methods, which requires an improved understanding of the relationship between multisensory interactions and hearing performance with a CI. The present study confirmed that CI users highly benefit from additional visual input and seem to use distinct processing strategies compared to NH listeners. In particular, their multisensory processes remain intact despite the insufficient auditory CI input. Our results are of clinical relevance, because strong audiovisual interactions could help optimising auditory rehabilitation after cochlear implantation. In the first weeks and months after implantation, the CI users need to learn to match the novel auditory inputs from the CI with the corresponding visual cues (Strelnikov et al., 2009). It may be speculated that audiovisual trainings can fasten and improve the auditory rehabilitation with a CI. Indeed, first results from NH listeners (with vocoded speech sounds) as well as results from late-implanted deafened ferrets have indicated that audiovisual training can enhance auditory perceptual learning (Bernstein et al., 2013; Bernstein et al., 2014; Isaiah et al., 2014; Kawase et al., 2009). Thus, future studies should examine whether these positive effects of audiovisual training on auditory-only perceptual learning can also be confirmed in CI users by including different groups (control group, training group) and by differentiating patients with short and long CI experience. This would allow to find out whether an audiovisual-based rehabilitation strategy can optimise rehabilitation and whether this approach is applicable even several months after CI switch-on.

5. Limitations

One limitation of this study is that we chose to stimulate only one ear, which might be unnatural for both the CI users and the NH listeners. This limits the conclusions regarding cross-modal reorganisation in CI users and makes it difficult to directly connect the experimental setting to the real world, as usually both ears are used. However, unilateral stimulation enabled us to control for the possible confounds caused by the contralateral ear (different hearing systems and hearing thresholds). Another limitation is that we used headphones for NH controls and loudspeakers for the CI users to present the stimuli. As we chose to stimulate one ear only, one cannot use loudspeakers for NH controls, as it is not possible to appropriately mute the second ear. Nevertheless, both the loudspeakers and the headphones were equally calibrated to an intensity level of 65 dB SPL, and a loudness rating scale was used to verify that the stimulus situation was as similar as possible across the two groups of participants. Moreover, one could argue that the different

approach of auditory presentation could have influenced audiovisual integration. We are not aware about studies that systematically examined differences in audiovisual speech, comparing the stimulation via headphones with the stimulation via loudspeakers. However, precise time-alignment of the auditory and visual stimulus – an important factor in multimodal processing – was given in both presentation schemes. Still, in order to better understand alterations in audiovisual processing after cochlear implantation, future studies are required to systematically compare the stimulation of the CI ear alone, the contralateral ear alone and both ears together. This should be done with ecologically valid stimuli, such as complex speech stimuli, and different task conditions (e. g. passive vs. active). Further, we suggest that future studies should compare different subgroups of CI patients (bimodal (CI + hearing aid), bilateral (CI + CI), single-sided deaf (CI + NH)) to explore potential differences in multisensory processing and cross-modal reorganisation across the different CI groups. Specifically, these studies need to assess whether the observed alterations in audiovisual speech processing in different groups of CI users can be used to fasten and improve the auditory rehabilitation after cochlear implantation.

6. Conclusion

This EEG study used electrical neuroimaging, including topographic and ERP source analysis, to examine cortical interactions during audiovisual syllable discrimination in CI users and NH listeners. Our results revealed a clear multisensory effect in both the CI users and the NH listeners, as indicated by shorter behavioural response times and reduced auditory-cortex activation at N1 and P2 latency. In addition, for the N1 ERP topography we found a condition effect specifically for the group of CI users, showing a distinct pattern of N1 topography between the auditory-only (A) and the modulated auditory response (AV-V). This observation points to a strong visual impact on auditory speech processing in CI users, which allows these individuals to compensate for the limited CI input. This visual impact on auditory speech processing is a manifestation of multisensory processes that remain intact in CI users despite the limited auditory signal provided by the CI. Finally, for the N1 ERP topography we found a group effect, which was independent of the condition (A, AV-V). Our source analyses suggest that this group difference is caused by two different reasons. First, the CI users show a delay in the cortical N1 response, pointing to difficulties in the cortical processing of the degraded signal. Second, the groups show distinct configurations of cortical sources at N1 latency, including not only a distinct directionality of auditory-cortex asymmetry (left-ward lateralisation in NH listeners, right-ward lateralisation in CI users), but also the additional recruitment of the visual cortex in the group of CI users, which was positively related to the duration of CI experience. These results confirm distinct multisensory processing between CI users and NH listeners, highlighting the need for audiovisual testing in clinical practice and for developing audiovisual methods in order to improve auditory rehabilitation.

Funding

This work was supported by the German Research Foundation (DFG) [SA 3615/1-1 (project number: 415896102) and SA 3615/2-1 (project number: 416867313)] and by the Köln Fortune Program/Faculty of Medicine, University of Cologne [278/2020] to PS. MMM is supported by the Swiss National Science Foundation (169206).

CRedit authorship contribution statement

Natalie Layer: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Project administration. **Anna Weglage:** Software, Validation, Formal analysis, Investigation, Writing – review & editing. **Verena Müller:** Conceptualization, Writing – review & editing.

Hartmut Meister: Conceptualization, Software, Resources, Writing – review & editing. **Ruth Lang-Roth:** Conceptualization, Writing – review & editing. **Martin Walger:** Conceptualization, Writing – review & editing. **Micah M. Murray:** Software, Formal analysis, Validation, Writing – review & editing. **Pascale Sandmann:** Conceptualization, Methodology, Validation, Resources, Writing – original draft, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (DFG) [SA 3615/1-1 (project number: 415896102) and SA 3615/2-1 (project number: 416867313)] and Köln Fortune [278/2020] for providing funding for this research. Additionally, we thank Pauline Burkhardt and Lina Wiesel for their help in data collection and Moritz Layer for his support with the implementation of the analysis. Finally, we thank all participants for their time and effort.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.nicl.2022.102982>.

References

- Ahveninen, J., Jääskeläinen, I.P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.-H., Sams, M., Shinn-Cunningham, B.G., Witzel, T., Belliveau, J. W., 2006. Task-modulated “what” and “where” pathways in human auditory cortex. *Proc. Natl. Acad. Sci.* 103 (39), 14608–14613.
- Allen, J.B., Hall, J.L., Jeng, P.S., 1990. Loudness growth in 1/2-octave bands (LGOB)—a procedure for the assessment of loudness. *J. Acoust. Soc. Am.* 88 (2), 745–753.
- Barth, D.S., Goldberg, N., Brett, B., Di, S., 1995. The spatiotemporal organization of auditory, visual, and auditory-visual evoked potentials in rat cortex. *Brain Res.* 678 (1–2), 177–190.
- Bavelier, D., Dye, M.W.G., Hauser, P.C., 2006. Do deaf individuals see better? *Trends Cognitive Sci.* 10 (11), 512–518.
- Bavelier, D., Neville, H.J., 2002. Cross-modal plasticity: Where and how? *Nat. Rev. Neurosci.* 3 (6), 443–452.
- Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. *Nature* 403 (6767), 309–312.
- Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7 (6), 1129–1159.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.: Ser. B (Methodol.)* 57 (1), 289–300.
- Bernstein, L.E., Auer, E.T., Eberhardt, S.P., Jiang, J., 2013. Auditory perceptual learning for speech perception can be enhanced by audiovisual training. *Front. Neurosci.* 7, 1–16.
- Bernstein, L.E., Eberhardt, S.P., Auer, E.T., 2014. Audiovisual spoken word training can promote or impede auditory-only perceptual learning: results from prelingually deafened adults with late-acquired cochlear implants and normal-hearing adults. *Front. Psychol.* 5, 1–20.
- Besle, J., Fort, A., Delpuech, C., Giard, M.-H., 2004. Bimodal speech: early suppressive visual effects in human auditory cortex. *Eur. J. Neurosci.* 20 (8), 2225–2234.
- Beynon, A.J., Snik, A.F.M., Stegeman, D.F., van den Broek, P., 2005. Discrimination of speech sound contrasts determined with behavioral tests and event-related potentials in cochlear implant recipients. *J. Am. Acad. Audiol.* 16 (01), 042–053.
- Biasiucci, A., Franceschiello, B., Murray, M.M., 2019. Electroencephalography. *Curr. Biol.* 29 (3), R80–R85.
- Billings, C.J., Bennett, K.O., Molis, M.R., Leek, M.R., 2011. Cortical encoding of signals in noise: effects of stimulus type and recording paradigm. *Ear Hear.* 32 (1), 53.
- Boemio, A., Fromm, S., Braun, A., Poeppel, D., 2005. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat. Neurosci.* 8 (3), 389–395.
- Bosnyak, D.J., Eaton, R.A., Roberts, L.E., 2004. Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cereb. Cortex* 14 (10), 1088–1099.
- Bottari, D., Bednaya, E., Dormal, G., Villwock, A., Dzhelelova, M., Grin, K., Pietrini, P., Ricciardi, E., Rossion, B., Röder, B., 2020. EEG frequency-tagging demonstrates increased left hemispheric involvement and crossmodal plasticity for face processing in congenitally deaf signers. *NeuroImage* 223, 117315.

- Bottari, D., Heimler, B., Caclin, A., Dalmolin, A., Giard, M.-H., Pavani, F., 2014. Visual change detection recruits auditory cortices in early deafness. *Neuroimage* 94, 172–184.
- Brüneck, I., Hölsken, S., Kessler, J., 2018. Der DemTect Eye+ Ear-Neues kognitives Screeningverfahren bei schwerhörigen Menschen mit Demenzverdacht. *Zeitschrift für Audiologie* 57 (3), 121.
- Brunet, D., Murray, M. M., ichel, C. M. (2011). *Spatiotemporal analysis of multichannel EEG: CARTOOL. Computational Intelligence and Neuroscience*, 2011.
- Campbell, J., Sharma, A., 2013. Compensatory changes in cortical resource allocation in adults with hearing loss. *Front. Syst. Neurosci.* 7, 71.
- Campbell, J., Sharma, A., 2016. Visual cross-modal re-organization in children with cochlear implants. *PLoS ONE* 11 (1), e0147793.
- Cappe, C., Thut, G., Romei, V., Murray, M.M., 2010. Auditory–visual multisensory interactions in humans: timing, topography, directionality, and sources. *J. Neurosci.* 30 (38), 12572–12580.
- Cappelletta, L., Harte, N., 2012. Phoneme-to-viseme mapping for visual speech recognition. *ICPRAM* 2, 322–329.
- Chen, L.-C., Sandmann, P., Thorne, J.D., Bleichner, M.G., Debener, S., 2016. Cross-modal functional reorganization of visual and auditory cortex in adult cochlear implant users identified with fNIRS. *Neural Plasticity* 2016, 1–13.
- Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26 (1), 55–67.
- De Santis, L., Clarke, S., Murray, M.M., 2007. Automatic and intrinsic auditory “what” and “where” processing in humans revealed by electrical neuroimaging. *Cereb. Cortex* 17 (1), 9–17.
- Debener, S., Hine, J., Bleck, S., Eyles, J., 2008. Source localization of auditory evoked potentials after cochlear implantation. *Psychophysiology* 45 (1), 20–24.
- Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134 (1), 9–21.
- Desai, S., Stickney, G., Zeng, F.-G., 2008. Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. *J. Acoust. Soc. Am.* 123 (1), 428–440.
- Destrieux, C., Fischl, B., Dale, A., Halgren, E., 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* 53 (1), 1–15.
- Dong, L., Foo, S. W., Lian, Y. (2003). Modeling continuous visual speech using boosted viseme models. *Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint*, 3, 1394–1398.
- Doucet, M., Bergeron, F., Lassonde, M., Ferron, P., Lepore, F., 2006. Cross-modal reorganization and speech perception in cochlear implant users. *Brain* 129 (12), 3376–3383.
- Drennan, W. R., & Rubinstein, J. T. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *J. Rehabilitation Res. Develop.*, 45(5), 779.
- Drijvers, L., Özyürek, A., 2017. Visual context enhanced: The joint contribution of iconic gestures and visible speech to degraded speech comprehension. *J. Speech Language Hearing Res.* 60 (1), 212–222.
- Dürschmid, S., Reichert, C., Hinrichs, H., Heinze, H.-J., Kirsch, H.E., Knight, R.T., Deouell, L.Y., 2019. Direct evidence for prediction signals in frontal cortex independent of prediction error. *Cereb. Cortex* 29 (11), 4530–4538.
- Fagel, S., Clemens, C., 2004. An articulation model for audiovisual speech synthesis—determination, adjustment, evaluation. *Speech Commun.* 44 (1–4), 141–154.
- Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., Sandmann, P., 2016a. On the relationship between auditory cognition and speech intelligibility in cochlear implant users: an ERP study. *Neuropsychologia* 87, 169–181.
- Finke, M., Sandmann, P., Bönitz, H., Kral, A., Büchner, A., 2016b. Consequences of stimulus type on higher-order processing in single-sided deaf cochlear implant users. *Audiol. Neurotol.* 21 (5), 305–315.
- Finney, E.M., Clementz, B.A., Hickok, G., Dobkins, K.R., 2003. Visual stimuli activate auditory cortex in deaf subjects: evidence from MEG. *NeuroReport* 14 (11), 1425–1427.
- Fujiki, N., Naito, Y., Nagamine, T., Shiomi, Y., Hirano, S., Honjo, I., Shibasaki, H., 1998. Influence of unilateral deafness on auditory evoked magnetic field. *NeuroReport* 9 (14), 3129–3133.
- Giraud, A.L., Price, C.J., Graham, J.M., Frackowiak, R.S., 2001a. Functional plasticity of language-related brain areas after cochlear implantation. *Brain* 124 (7), 1307–1316.
- Giraud, A.L., Truy, E., Frackowiak, R., 2001b. Imaging plasticity in cochlear implant patients. *Audiol. Neurotol.* 6 (6), 381–393.
- Giraud, A.-L., Price, C.J., Graham, J.M., Truy, E., Frackowiak, R.S.J., 2001c. Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron* 30 (3), 657–664.
- Glennon, E., Svirsky, M.A., Froemke, R.C., 2020. Auditory cortical plasticity in cochlear implant users. *Curr. Opin. Neurobiol.* 60, 108–114.
- Gramfort, A., Papadopoulos, T., Olivi, E., Clerc, M., 2010. OpenMEEG: Opensource software for quasistatic bioelectromagnetics. *Biomed. Eng. Online* 9 (1), 1–20.
- Grant, K.W., Walden, B.E., Seitz, P.F., 1998. Auditory-visual speech recognition by hearing-impaired subjects: consonant recognition, sentence recognition, and auditory-visual integration. *J. Acoust. Soc. Am.* 103 (5), 2677–2690.
- Green, K.M.J., Julian, P.J., Hastings, D.L., Ramsden, R.T., 2005. Auditory cortical activation and speech perception in cochlear implant users: effects of implant experience and duration of deafness. *Hear. Res.* 205 (1–2), 184–192.
- Hahlbrock, K.H., 1970. *Sprachaudiometrie: Grundlagen und praktische Anwendung einer Sprachaudiometrie für das deutsche Sprachgebiet*. Georg Thieme Verlag.
- Hansen, P., Kringelbach, M., Salmelin, R., 2010. *MEG: An introduction to methods*. Oxford University Press.
- Hauthal, N., Thorne, J.D., Debener, S., Sandmann, P., 2014. Source localisation of visual evoked potentials in congenitally deaf individuals. *Brain Topogr.* 27 (3), 412–424.
- Heimler, B., Weisz, N., Collignon, O., 2014. Revisiting the adaptive and maladaptive effects of crossmodal plasticity. *Neuroscience* 283, 44–63.
- Henkin, Y., Yaar-Soffer, Y., Steinberg, M., Muchnik, C., 2014. Neural correlates of auditory-cognitive processing in older adult cochlear implant recipients. *Audiol. Neurotol.* 19 (1), 21–26.
- Hinderink, J.B., Krabbe, P.F.M., Van Den Broek, P., 2000. Development and application of a health-related quality-of-life instrument for adults with cochlear implants: the nijmegen cochlear implant questionnaire. *Otolaryngol.-Head Neck Surg.* 123 (6), 756–765.
- Inui, K., Okamoto, H., Miki, K., Gunji, A., Kakigi, R., 2006. Serial and parallel processing in the human auditory cortex: a magnetoencephalographic study. *Cereb. Cortex* 16 (1), 18–30.
- Isaiah, A., Vongpaisal, T., King, A.J., Hartley, D.E.H., 2014. Multisensory training improves auditory spatial processing following bilateral cochlear implantation. *J. Neurosci.* 34 (33), 11119–11130.
- Jachimski, D., Czerwinski, A., Ciszewski, T., 2018. A comparative study of english viseme recognition methods and algorithms. *Multimedia Tools Appl.* 77 (13), 16495–16532.
- Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., McKeown, M.J., Iragui, V., Sejnowski, T.J., 2000. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37 (2), 163–178.
- Kaiser, A.R., Kirk, K.I., Lachs, L., Pisoni, D.B., 2003. Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *J. Speech Language Hear. Res.* 46 (2), 390–404.
- Kalbe, E., Kessler, J., Calabrese, P., Smith, R., Passmore, A.P., Brand, M., Bullock, R., 2004. DemTect: A new, sensitive cognitive screening test to support the diagnosis of mild cognitive impairment and early dementia. *Int. J. Geriatric Psych.* 19 (2), 136–143.
- Kawase, T., Sakamoto, S., Hori, Y., Maki, A., Suzuki, Y., Kobayashi, T., 2009. Bimodal audio–visual training enhances auditory adaptation process. *NeuroReport* 20, 1231–1234.
- Laurienti, P.J., Kraft, R.A., Maldjian, J.A., Burdette, J.H., Wallace, M.T., 2004. Semantic congruence is a critical factor in multisensory behavioral performance. *Exp. Brain Res.* 158 (4), 405–414.
- Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysiol.* 48 (6), 609–621.
- Lim, H.H., Lenarz, T., 2015. Auditory midbrain implant: research and development towards a second clinical trial. *Hear. Res.* 322, 212–223.
- Lin, F.-H., Witzel, T., Ahlfors, S.P., Stufflebeam, S.M., Belliveau, J.W., Hämäläinen, M.S., 2006. Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *Neuroimage* 31 (1), 160–171.
- Lucey, P., Martin, T., Sridharan, S. (2004). Confusability of phonemes grouped according to their viseme classes in noisy environments. *Proc. Of Australian Int. Conf. On Speech Science & Tech*, 265–270.
- Lütkenhöner, B., Steinsträter, O., 1998. High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiol. Neurotol.* 3 (2–3), 191–213.
- Mahavidyalaya, B. V. (2014). Phoneme and viseme based approach for lip synchronization. *Int. J. Signal Process., Image Process. Pattern Recogn.*, 7(3), 385–394.
- Massaro, D.W., Light, J., 2004. Using visible speech to train perception and production of speech for individuals with hearing loss. *J. Speech Lang Hear. Res.* 47 (2), 304–320.
- McKay, C.M., Lim, H.H., Lenarz, T., 2013. Temporal processing in the auditory system: insights from cochlear and auditory midbrain implantation. *J. Assoc. Res. Otolaryngol.* 14 (1), 103–124.
- Meister, H., Fuersen, K., Schreitmüller, S., Walger, M., 2016. Effect of acoustic fine structure cues on the recognition of auditory-only and audiovisual speech. *J. Acoust. Soc. Am.* 139 (6), 3116–3120.
- Merzenich, M.M., Van Vleet, T.M., Nahum, M., 2014. Brain plasticity-based therapeutics. *Front. Hum. Neurosci.* 8, 385.
- Micco, A.G., Kraus, N., Koch, D.B., McGee, T.J., Carrell, T.D., Sharma, A., Nicol, T., Wiet, R.J., 1995. Speech-evoked cognitive P300 potentials in cochlear implant recipients. *Am. J. Otol.* 16 (4), 514–520.
- Michel, C.M., Koenig, T., 2018. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *Neuroimage* 180, 577–593.
- Michel, C.M., Koenig, T., Brandeis, D., Gianotti, L.R., Wackermann, J., 2009. *Electrical Neuroimaging*. Cambridge University Press.
- Michel, C.M., Murray, M.M., 2012. Towards the utilization of EEG as a brain imaging tool. *Neuroimage* 61 (2), 371–385.
- Miller, J., 1982. Divided attention: Evidence for coactivation with redundant signals. *Cogn. Psychol.* 14 (2), 247–279.
- Mitchell, T.V., Maslin, M.T., 2007. How vision matters for individuals with hearing loss. *Int. J. Audiol.* 46 (9), 500–511.
- Moody-Antonio, S., Takayanagi, S., Masuda, A., Auer Jr, E.T., Fisher, L., Bernstein, L.E., 2005. Improved speech perception in adult congenitally deafened cochlear implant recipients. *Otol. Neurotol.* 26 (4), 649–654.
- Mortensen, M.V., Mirz, F., Gjedde, A., 2006. Restored speech comprehension linked to activity in left inferior prefrontal and right temporal cortices in postlingual deafness. *Neuroimage* 31 (2), 842–852.
- Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 20 (4), 249–264.

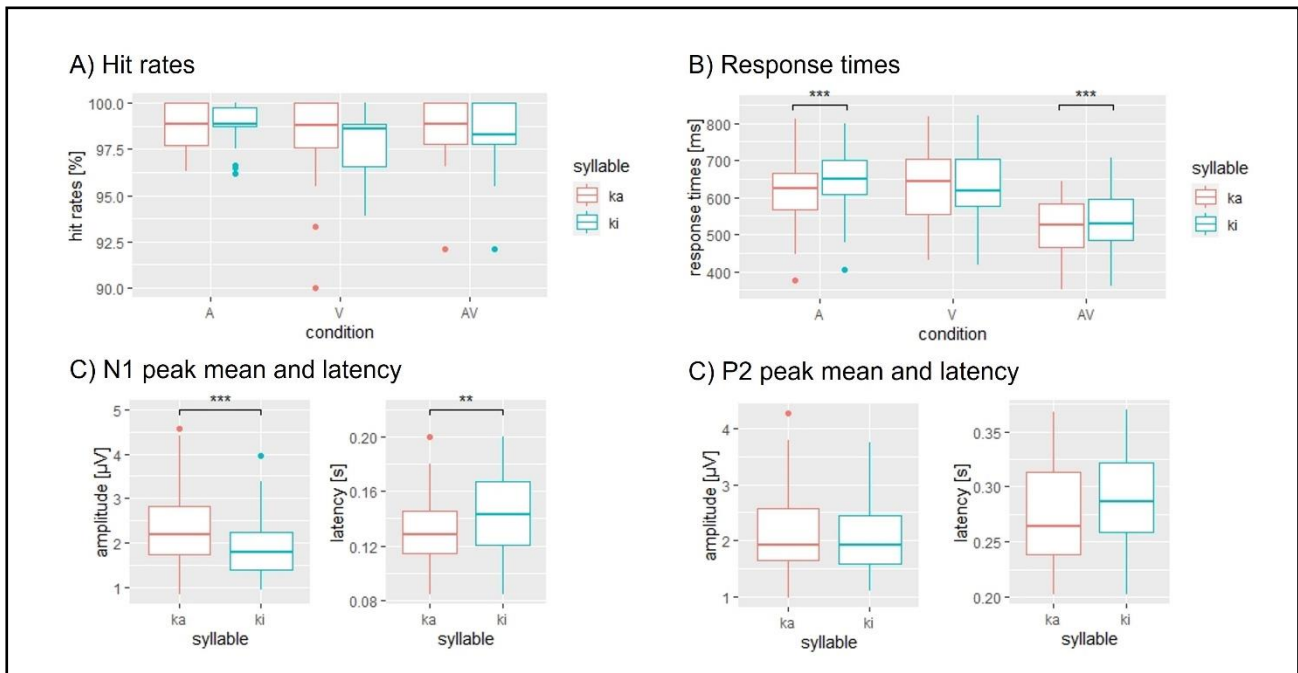
- Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24 (4), 375–425.
- Obleser, J., Elbert, T., Lahiri, A., Eulitz, C., 2003. Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. *Cognitive Brain Res.* 15 (3), 207–213.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia* 9 (1), 97–113.
- Pantev, C., Dinnesen, A., Ross, B., Wollbrink, A., Knief, A., 2006. Dynamics of auditory plasticity after cochlear implantation: a longitudinal study. *Cereb. Cortex* 16 (1), 31–36.
- Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72 (2), 184–187.
- Poeppl, D., 2003. The analysis of speech in different temporal integration windows: cerebral lateralization as “asymmetric sampling in time”. *Speech Commun.* 41 (1), 245–255.
- Ponton, C.W., Vasama, J.-P., Tremblay, K., Khosla, D., Kwong, B., Don, M., 2001. Plasticity in the adult human central auditory system: Evidence from late-onset profound unilateral deafness. *Hear. Res.* 154 (1–2), 32–44.
- Prince, P., Paul, B.T., Chen, J., Le, T., Lin, V., Dimitrijevic, A., 2021. Neural correlates of visual stimulus encoding and verbal working memory differ between cochlear implant users and normal-hearing controls. *Eur. J. Neurosci.*
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Raab, D. H. (1962). Division of psychology: Statistical facilitation of simple reaction times. *Trans. New York Acad. Sci.*, 24(5 Series II), 574–590.
- Ratcliff, R., 1979. Group reaction time distributions and an analysis of distribution statistics. *Psychol. Bull.* 86 (3), 446–461.
- van de Rijt, L.P., Roye, A., Mylanus, E.A., van Opstal, A.J., van Wanrooij, M.M., 2019. The principle of inverse effectiveness in audiovisual speech perception. *Front. Hum. Neurosci.* 13, 335.
- Ross, B., Tremblay, K., 2009. Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hear. Res.* 248 (1–2), 48–59.
- Rouger, J., Fraysse, B., Deguine, O., Barone, P., 2008. McGurk effects in cochlear-implanted deaf subjects. *Brain Res.* 1188, 87–99.
- Rouger, J., Lagleyre, S., Démonet, J.-F., Fraysse, B., Deguine, O., Barone, P., 2012. Evolution of crossmodal reorganization of the voice area in cochlear-implanted deaf patients. *Hum. Brain Mapp.* 33 (8), 1929–1940.
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., Barone, P., 2007. Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proc. Natl. Acad. Sci.* 104 (17), 7295–7300.
- Sandmann, P., Dillier, N., Eichele, T., Meyer, M., Kegel, A., Pascual-Marqui, R.D., Marcar, V.L., Jäncke, L., Debener, S., 2012. Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain* 135 (2), 555–568.
- Sandmann, P., Eichele, T., Buechler, M., Debener, S., Jäncke, L., Dillier, N., Hugdahl, K., Meyer, M., 2009. Evaluation of evoked potentials to dyadic tones after cochlear implantation. *Brain* 132 (7), 1967–1979.
- Sandmann, P., Kegel, A., Eichele, T., Dillier, N., Lai, W., Bendixen, A., Debener, S., Jäncke, L., Meyer, M., 2010. Neurophysiological evidence of impaired musical sound perception in cochlear-implant users. *Clin. Neurophysiol.* 121 (12), 2070–2082.
- Sandmann, P., Plotz, K., Hauthal, N., de Vos, M., Schönfeld, R., Debener, S., 2015. Rapid bilateral improvement in auditory cortex activity in postlingually deafened adults following cochlear implantation. *Clin. Neurophysiol.* 126 (3), 594–607.
- Schiel, F., 1999. Automatic phonetic transcription of non-prompted speech. Open Access LMU.
- Schierholz, I., Finke, M., Kral, A., Büchner, A., Rach, S., Lenarz, T., Dengler, R., Sandmann, P., 2017. Auditory and audio-visual processing in patients with cochlear, auditory brainstem, and auditory midbrain implants: an EEG study. *Hum. Brain Mapp.* 38 (4), 2206–2225.
- Schierholz, I., Finke, M., Schulte, S., Hauthal, N., Kantzke, C., Rach, S., Büchner, A., Dengler, R., Sandmann, P., 2015. Enhanced audio-visual interactions in the auditory cortex of elderly cochlear-implant users. *Hear. Res.* 328, 133–147.
- Schierholz, I., Schönermark, C., Ruigendijk, E., Kral, A., Kopp, B., Büchner, A., 2021. An event-related brain potential study of auditory attention in cochlear implant users. *Clinical* 132 (9), 2290–2305.
- Schreitmüller, S., Frenken, M., Bentz, L., Ortmann, M., Walger, M., Meister, H., 2018. Validating a method to assess lipreading, audiovisual gain, and integration during speech reception with cochlear-implanted and normal-hearing subjects using a talking head. *Ear Hear.* 39 (3), 503–516.
- Senkowski, D., Pomper, U., Fitzner, I., Engel, A.K., Kral, A., 2014. Beta-band activity in auditory pathways reflects speech localization and recognition in bilateral cochlear implant users. *Hum. Brain Mapp.* 35 (7), 3107–3121.
- Sharma, A., Dorman, M.F., Spahr, A.J., 2002. A sensitive period for the development of the central auditory system in children with cochlear implants: Implications for age of implantation. *Ear Hear.* 23 (6), 532–539.
- Stein, B., Meredith, M., 1993. *The merging of the senses*. The MIT Press, Cambridge, MA.
- Stekelenburg, J.J., Vroomen, J., 2007. Neural correlates of multisensory integration of ecologically valid audiovisual events. *J. Cognit. Neurosci.* 19 (12), 1964–1973.
- Stevenson, R.A., Ghose, D., Fister, J.K., Sarko, D.K., Altieri, N.A., Nidiffer, A.R., Kurela, L. R., Siemann, J.K., James, T.W., Wallace, M.T., 2014. Identifying and quantifying multisensory integration: a tutorial review. *Brain Topogr.* 27 (6), 707–730.
- Stevenson, R., Sheffield, S.W., Butera, I.M., Gifford, R.H., Wallace, M., 2017. Multisensory integration in cochlear implant recipients. *Ear Hear.* 38 (5), 521.
- Strelnikov, K., Rouger, J., Barone, P., Deguine, O., 2009. Role of speechreading in audiovisual interactions during the recovery of speech comprehension in deaf adults with cochlear implants. *Scand. J. Psychol.* 50 (5), 437–444.
- Strelnikov, K., Rouger, J., Demonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., Barone, P., 2010. Does brain activity at rest reflect adaptive strategies? Evidence from speech processing after cochlear implantation. *Cereb. Cortex* 20 (5), 1217–1222.
- Strelnikov, K., Rouger, J., Demonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., Barone, P., 2013. Visual activity predicts auditory recovery from deafness after adult cochlear implantation. *Brain* 136 (12), 3682–3695.
- Strelnikov, K., Rouger, J., Lagleyre, S., Fraysse, B., Démonet, J.-F., Déguine, O., Barone, P., 2015. Increased audiovisual integration in cochlear-implanted deaf patients: independent components analysis of longitudinal positron emission tomography data. *Eur. J. Neurosci.* 41 (5), 677–685.
- Stropahl, M., Bauer, A.-K.-R., Debener, S., Bleichner, M.G., 2018. Source-modeling auditory processes of EEG data using EEGLAB and brainstorm. *Front. Neurosci.* 12, 309.
- Stropahl, M., Chen, L.-C., Debener, S., 2017. Cortical reorganization in postlingually deaf cochlear implant users: intra-modal and cross-modal considerations. *Hear. Res.* 343, 128–137.
- Stropahl, M., Debener, S., 2017. Auditory cross-modal reorganization in cochlear implant users indicates audio-visual integration. *NeuroImage: Clinical* 16, 514–523.
- Stropahl, M., Plotz, K., Schönfeld, R., Lenarz, T., Sandmann, P., Yovel, G., De Vos, M., Debener, S., 2015. Cross-modal reorganization in cochlear implant users: auditory cortex contributes to visual face processing. *Neuroimage* 121, 159–170.
- Summy, W.H., Pollack, I., 1954. Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26 (2), 212–215.
- Sun, Z., Seo, J.W., Park, H.J., Lee, J.Y., Kwak, M.Y., Kim, Y., Lee, J.Y., Park, J.W., Kang, W.S., Ahn, J.H., Chung, J.W., Kim, H., 2021. Cortical reorganization following auditory deprivation predicts cochlear implant performance in postlingually deaf adults. *Hum. Brain Mapp.* 42 (1), 233–244.
- Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: A user-friendly application for MEG/EEG analysis. *Comput. Intel. Neurosci.* 2011, 1–13.
- Teder-Sälejärvi, W.A., Di Russo, F., McDonald, J.J., Hillyard, S.A., 2005. Effects of spatial congruity on audio-visual multimodal integration. *J. Cognit. Neurosci.* 17 (9), 1396–1409.
- Timm, L., Agrawal, D., Viola, F.C., Sandmann, P., Debener, S., Büchner, A., Dengler, R., Wittfoth, M., Malmierca, M.S., 2012. Temporal feature perception in cochlear implant users. *PLoS ONE* 7 (9), e45375.
- Tremblay, C., Champoux, F., Lepore, F., Théoret, H., 2010. Audiovisual fusion and cochlear implant proficiency. *Restor. Neurol. Neurosci.* 28 (2), 283–291.
- Tremblay, K., Ross, B., Inoue, K., McClannahan, K., Collet, G., 2014. Is the auditory evoked P2 response a biomarker of learning? *Front. Syst. Neurosci.* 8, 28.
- Ulrich, R., Miller, J., Schröter, H., 2007. Testing the race model inequality: an algorithm and computer programs. *Behav. Res. Methods* 39 (2), 291–302.
- van Wassenhove, V., Grant, K.W., Poeppl, D., 2005. Visual speech speeds up the neural processing of auditory speech. *Proc. Natl. Acad. Sci.* 102 (4), 1181–1186.
- Vasil, K.J., Lewis, J., Tamati, T., Ray, C., Moberly, A.C., 2020. How does quality of life relate to auditory abilities? A subitem analysis of the nijmegen cochlear implant questionnaire. *J. Am. Acad. Audiol.* 31 (04), 292–301.
- Vaughan, H.G., 1982. The neural origins of human event-related potentials. *Ann. N. Y. Acad. Sci.* 388 (1), 125–138.
- Viola, F.C., De Vos, M., Hine, J., Sandmann, P., Bleck, S., Eyles, J., Debener, S., 2012. Semi-automatic attenuation of cochlear implant artifacts for the evaluation of late auditory evoked potentials. *Hear. Res.* 284 (1–2), 6–15.
- Vroomen, J., Stekelenburg, J.J., 2010. Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *J. Cognit. Neurosci.* 22 (7), 1583–1596.
- Wesemann, W., Schiefer, U., Bach, M., 2010. Neue DIN-Normen zur Sehschärfebestimmung New DIN norms for determination of visual acuity. *Der Ophthalmologe* 107 (9), 821–826.
- Wesker, T., Meyer, B., Wagoner, K., Anemüller, J., Mertins, A., Kollmeier, B., 2005. Oldenburg logatome speech corpus (OLLO) for speech recognition experiments with humans and machines. Ninth European Conference on Speech Communication and Technology.
- Widmann, A., Schröger, E., Maess, B., 2015. Digital filter design for electrophysiological data—a practical approach. *J. Neurosci. Methods* 250, 34–46.
- Williams, N., 2017. The borg rating of perceived exertion (RPE) scale. *Occup. Med.* 67 (5), 404–405.
- Zatorre, R.J., Belin, P., 2001. Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11 (10), 946–953.

Supplementary Material

To make sure that there were no group differences and no group-specific effects concerning the two syllable conditions (/ki/ and /ka/), we conducted some supplementary analyses on the behavioural and ERP data. On the behavioural level, we took a closer look at the hit rates and the response times and on the ERP level, we analysed the peak mean and latencies of the global field power (GFP) for the N1 and the P2 time windows.

- 1) **Hit rates:** We computed a three-way mixed-model ANOVA with group (NH, CI) as within-subjects factor and condition (A, V, AV) and syllable (/ki/, /ka/) as within-subjects factors for the hit rates. In the case of violation of the sphericity assumption, a Greenhouse-Geisser correction was applied. Moreover, post-hoc t-tests were performed and corrected for multiple comparisons using a Bonferroni correction, if there were significant main effects or interactions ($p < .05$). The ANOVA showed a main effect of condition ($F_{1,64} = 3.74$; $p \leq .05$, $\eta p^2 = 0.031$). Post-hoc t-tests showed a difference between A and V ($t(67) = 3.11$; $p \leq .01$) with A exhibiting higher hit rates than V.
- 2) **Response times:** We computed a three-way mixed-model ANOVA with group (NH, CI) as within-subjects factor and condition (A, V, AV) and syllable (/ki/, /ka/) as within-subjects factors for the response times. In the case of violation of the sphericity assumption, a Greenhouse-Geisser correction was applied. Moreover, post-hoc t-tests were performed and corrected for multiple comparisons using a Bonferroni correction, if there were significant main effects or interactions ($p < .05$). The ANOVA showed a condition x syllable interaction ($F_{1,64} = 9.53$; $p \leq .01$, $\eta p^2 = 0.002$). Post-hoc t-tests showed a difference between /ki/ and /ka/ for A ($t(33) = -5.56$; $p \leq .001$) and a difference between /ki/ and /ka/ for AV ($t(33) = -4.44$; $p \leq .001$), with /ka/ showing shorter response times than /ki/.
- 3) **GFP peak mean and latency:** We computed a three-way mixed-model ANOVA with group (NH, CI) as within-subjects factor and condition (A, V, AV) and syllable (/ki/, /ka/) as within-subjects factors for the N1 peak mean and latency and the P2 peak mean and latency. In the case of violation of the sphericity assumption, a Greenhouse-Geisser correction was applied. Moreover, post-hoc t-tests were performed and corrected for multiple comparisons using a Bonferroni correction, if there were significant main effects or interactions ($p < .05$).
 - **N1 peak mean:** The ANOVA showed a main effect of syllable ($F_{1,32} = 18.71$; $p \leq .001$, $\eta p^2 = 0.06$). Post-hoc t-tests showed a difference between /ki/ and /ka/ ($t(69) = 5.75$; $p \leq .001$), with /ka/ showing greater amplitudes than /ki/.
 - **N1 peak latency:** The ANOVA showed a main effect of group ($F_{1,32} = 8.21$; $p \leq .01$, $\eta p^2 = 0.068$) and a main effect of syllable ($F_{1,32} = 4.72$; $p \leq .05$, $\eta p^2 = 0.05$). Post-hoc t-tests showed a difference between /ki/ and /ka/ ($t(69) = -2.72$; $p \leq .01$), with /ka/ showing shorter latencies than /ki/. In addition, post-hoc t-tests showed a difference between NH and CI ($t(69) = 3.03$; $p \leq .01$), with CI users exhibiting longer latencies compared to NH.
 - **P2 peak mean:** The ANOVA showed no significant main effects or interactions.
 - **P2 peak latency:** The ANOVA showed a main effect of condition ($F_{1,32} = 4.35$; $p \leq .05$, $\eta p^2 = 0.019$). Post-hoc t-tests showed a difference between AV-V and A ($t(33) = 3.0$; $p \leq .05$), with AV-V showing shorter latencies than A.

The results revealed that there were no group-specific differences or group-specific interactions caused by including /ki/ and /ka/ as separate factors in the ANOVA. The Supplementary Figure 1 illustrates the results from the supplementary analyses for both the behavioural (hit rates, response times) and the ERP data (GFP peak mean and latency for N1 and P2).



Supplementary Figure 1: Analyses including the syllables /ki/ and /ka/ as separate factors. A) Hit rates: no difference between the syllables. B) Response times: difference in response times between the syllables for A and AV, but no group difference. C) N1 peak mean and latency: difference between the syllables (/ka/ shows greater amplitudes and shorter latencies than /ki/, but no group difference. D) P2 peak mean and latency: no significant difference between the syllables and the groups. Significant differences are indicated (*p < 0.05, **p < 0.01, ***p < 0.001).

"SIDE-OF-IMPLANTATION EFFECT ON FUNCTIONAL
ASYMMETRY IN THE AUDITORY CORTEX OF SINGLE-SIDED
DEAF COCHLEAR-IMPLANT USERS"

This chapter is based on the following publication:

Weglage, A., Müller, V., **Layer, N.**, Abdel-Latif, K. H., Lang-Roth, R., Walger, M., & Sandmann, P. (2022). Side-of-implantation effect on functional asymmetry in the auditory cortex of single-sided deaf CI-implant users. *Brain Topography*, 1–22. DOI: 10.1007/s10548-022-00902-3.

Author contributions:

Anna Weglage: Literature review; participant recruitment; data acquisition; data analysis; interpretation and presentation of results to research group; discussion of findings and further procedures; writing initial manuscript draft; incorporating suggestions of co-authors; submission to journal; presentation of results at national conference. **Verena Müller:** Conceptualisation; advisory role (CI technician); participant recruitment; partially writing of part "stimuli"; proofreading of manuscript. **Natalie Layer:** Assistance in data acquisition; assistance in data analysis; group discussions; proofreading of manuscript. **Khaled H. Abdel-Latif:** Programming; partially writing of part "stimuli"; proofreading of manuscript. **Ruth Lang-Roth:** Conceptualisation; proof-reading of manuscript. **Martin Walger:** Conceptualisation; proof-reading of manuscript. **Pascale Sandmann:** Supervisor of Anna Weglage and Natalie Layer; programming of the experiment; pilot measurements; central role in project conceptualisation; funding acquisition; suggestions; discussions; assistance in writing manuscript and proofreading of manuscript.

SYNOPSIS

OBJECTIVES SSD CI users experience a maximally asymmetric hearing experience, as the input highly differs between the CI and NH ears (Gordon et al., 2013; Kral et al., 2013). SSD CI users are a subset of CI users who are a comparably new CI patient group, because it was previously believed that implantation was unnecessary due to the intact NH ear (Arndt et al., 2011; Buechner et al., 2010). Nowadays, it is recognised that an implantation improves sound localisation, speech-in-noise intelligibility, and quality of life due to the restoration of binaural hearing (Kitterick et al., 2015). Nevertheless, it remains unknown whether the side of implantation

affects the benefits of CI implantation in those patients. Furthermore, when tested monaurally, it is unclear whether speech processing via the NH ear of SSD CI users is comparable to that of NH controls. To fill this gap, this EEG study aimed to investigate the effect of implantation side on auditory speech processing by comparing the CI ear with the NH ear of SSD CI users, as well as a control group of NH listeners. Additionally, we intended to demonstrate that the experimental design we chose can be used to objectively differentiate between proficient and non-proficient SSD CI users.

METHODS An EEG with 32 electrodes was recorded from SSD CI users and a group of NH controls. The stimuli were auditory syllables (/ki/, /ka/, /ti/) which were presented in a two-deviant oddball paradigm, whereby /ki/ was the frequent syllable and the other two served as the infrequent deviant syllables. The task of the participants was to press a button whenever a deviant syllable was perceived. The CI users were tested with their CI and NH ears separately, one after the other. The NH controls were also tested monaurally, whereby the other ear was stimulated with a vocoded version of the stimuli. This means that the stimuli were degraded such that they resembled the theoretical sound quality of a CI (Shannon et al., 1995). It was expected that non-proficient SSD CI users are less capable of distinguishing /ki/ from /ti/ because of the subtle difference in the consonant, which only differs in the place of articulation. This is in contrast to the distinction between the syllables /ki/ and /ka/, where the vowels lead to a clear difference.

ANALYSIS We analysed the hit rates and response times of the button presses and the ERPs at N1, P2 and P3b latencies. The P3b is a marker of the central auditory system's capability to discriminate between a standard and a deviant sound (Polich, 2007; Henkin et al., 2009). On the sensor level, the N1 and the P2 were analysed based on the activity at fronto-central electrodes, while the P3b was analysed based on the activity at parietal electrodes. Subsequently, we performed a source analysis with the ipsi- and contralateral activity in the auditory cortex. Both sensor and source ERP data were statistically analysed based on the separate research question, including the factors "group" (proficient/non-proficient SSD CI users or left/right implanted SSD CI users), "stimulation side" (NH/CI ears), "stimulus type" (standard/ deviant 1/ deviant 2), and "hemisphere" (left/ right; for source analysis only).

RESULTS As expected, it was possible to distinguish proficient and non-proficient SSD CI users based on the similarity between the standard and deviant syllables when they were stimulated via the CI ear. In particular, the proficiency level was reflected in the ERP amplitudes at N1 and P3b latency, with reduced and/or prolonged amplitudes for non-proficient compared to proficient CI users. Accordingly, non-proficient CI users showed poorer hit rates than proficient CI users. Comparing the processing of syllables between the NH ear and the CI ear of the SSD CI users revealed generally reduced and/or delayed ERPs for the CI ear, and prolonged response times for the CI ear compared to the NH ear. Comparing the processing of NH controls for original and vocoded stimulation did not reveal any statistical differences, neither at the behavioural nor at the sensory ERP level. Furthermore, the left-ear implanted SSD CI users exhibited an enhanced functional asymmetry in the auditory cortex than

right-ear implanted SSD CI users, showing enhanced auditory-cortex activity in the right compared to the left hemisphere. This was observable independent of whether the syllables were presented via the CI or the NH ears. For right-implanted CI users no such effect was observable. Interestingly, contrasting left with right stimulation with vocoded stimuli in NH controls resulted in the same pattern of functional asymmetry within the auditory cortex as in the left-implanted CI users. Finally, comparing the processing of the NH ears of the CI users with the NH controls revealed differences in functional asymmetry within the auditory cortex. Again, left-ear implanted CI users showed a greater activity in the right hemisphere compared to right-ear implanted CI users. NH controls did not exhibit such effects.

DISCUSSION The results confirm that the proficiency in syllable discrimination can be assessed by N₁ and P_{3b} ERPs. Furthermore, the findings show that speech processing is more difficult via the CI ear compared to the NH ear for SSD CI users. A side-of-implantation effect was observed for left-implanted CI users when they were stimulated both via the CI and the NH ear. When NH controls were stimulated with vocoded (CI input simulated) syllables this resulted in a comparable effect for left-sided stimulation. Whereas, when comparing the NH ears of CI users and NH controls, the functional asymmetry is only observable for the CI users. We hypothesise that the observed side-of-implantation effects for CI users emerge as a result of cortical reorganisation caused by temporal deafness and/or limited sensory input following cochlear implantation. Overall, the findings provide a better understanding of rehabilitation success in SSD CI users by demonstrating that CI-related stimulus degradation and experience-related functional changes in the auditory cortex affect cortical speech processing in SSD CI users.



Side-of-Implantation Effect on Functional Asymmetry in the Auditory Cortex of Single-Sided Deaf Cochlear-Implant Users

Anna Weglage¹ · Verena Müller¹ · Natalie Layer¹ · Khaled H. A. Abdel-Latif² · Ruth Lang-Roth¹ · Martin Walger^{1,2} · Pascale Sandmann¹

Received: 19 November 2021 / Accepted: 10 May 2022 / Published online: 7 June 2022
© The Author(s) 2022

Abstract

Cochlear implants (CIs) allow to restore the hearing function in profoundly deaf individuals. Due to the degradation of the stimulus by CI signal processing, implanted individuals with single-sided deafness (SSD) have the specific challenge that the input highly differs between their ears. The present study compared normal-hearing (NH) listeners ($N = 10$) and left- and right-ear implanted SSD CI users ($N = 10$ left, $N = 9$ right), to evaluate cortical speech processing between CI- and NH-ears and to explore for side-of-implantation effects. The participants performed a two-deviant oddball task, separately with the left and the right ear. Auditory event-related potentials (ERPs) in response to syllables were compared between proficient and non-proficient CI users, as well as between CI and NH ears. The effect of the side of implantation was analysed on the sensor and the source level. CI proficiency could be distinguished based on the ERP amplitudes of the N1 and the P3b. Moreover, syllable processing via the CI ear, when compared to the NH ear, resulted in attenuated and delayed ERPs. In addition, the left-ear implanted SSD CI users revealed an enhanced functional asymmetry in the auditory cortex than right-ear implanted SSD CI users, regardless of whether the syllables were perceived via the CI or the NH ear. Our findings reveal that speech-discrimination proficiency in SSD CI users can be assessed by N1 and P3b ERPs. The results contribute to a better understanding of the rehabilitation success in SSD CI users by showing that cortical speech processing in SSD CI users is affected by CI-related stimulus degradation and experience-related functional changes in the auditory cortex.

Keywords Single-sided deafness · Cochlear implants · Event-related potential · Oddball paradigm · Hemispheric asymmetry · Cortical plasticity

Introduction

Cochlear implants (CIs) can (partially) restore the hearing of individuals with severe to profound sensorineural hearing loss by direct electrical stimulation of the auditory nerve (Zeng et al. 2011). Compared to normal acoustic hearing, the sounds transmitted by a CI are limited in the spectral

and temporal domain and have a smaller dynamic range (Drennan, 2008). Therefore, CI recipients need to adapt to the highly artificial inputs after implantation. However, the speech understanding with a CI remains limited and highly variable across the patients (Lenarz et al. 2012; Roberts et al. 2013). Previous studies have shown that different individual factors contribute to this variability in CI outcome, among them peripheral factors (e.g. positioning of the electrode array), cortical reorganisation as induced by auditory deprivation (before cochlear implantation) and by the (limited) electrical hearing with a CI (after cochlear implantation) (Lazard et al. 2012b; Lazard et al. 2012a, b, c).

The clinical margins for CI indication have been extended over the last years, now including single-sided deaf (SSD) individuals (Arndt et al. 2011a, b; Buechner et al. 2010). SSD CI users are to be distinguished from bilateral (CI on both ears) and bimodal (CI on one ear and hearing aid on the contralateral ear) CI users, since in SSD CI users the

Handling Editor: Micah M. Murray.

✉ Anna Weglage
anna.weglage1@uk-koeln.de

¹ Faculty of Medicine and University Hospital Cologne, Department of Otorhinolaryngology, Head and Neck Surgery, Audiology and Pediatric Audiology, Cochlear Implant Center, University of Cologne, Cologne, Germany

² Jean-Uhrmacher-Institute for Clinical ENT Research, University of Cologne, Cologne, Germany

signal quality of the input highly differs between the two ears, and the normal-hearing (NH) ear typically remains the dominant communication channel. This leads to maximal asymmetric auditory processing in this group of patients (Gordon et al. 2013; Kral et al., 2013). Nevertheless, Arndt et al. (2011a, b) showed improved hearing abilities in SSD patients aided with a CI compared to alternative treatments like contralateral routing of signal (CROS) or bone-anchored hearing aids (BAHA). SSD CI patients particularly benefit in sound localisation, speech understanding in noise and quality of life due to the restored binaural hearing (Kitterick et al. 2015). However, it remains not well understood whether the extent of the benefits in SSD CI users depends on the side of implantation. Similar to the findings from CI users with bilateral hearing loss (Mosnier et al. 2014), first results point to a right-ear advantage for speech recognition ability in SSD CI patients as well (Wettstein and Probst, 2018). In addition to the largely unresolved question regarding the side-of-implantation effects, it remains unclear whether SSD patients—when tested with only the NH ear—show the same speech processing capabilities as NH listeners who use only one of their ears (Arndt et al. 2019; Maslin et al. 2015). To better understand these effects on the CI outcome in SSD CI users, the current study systematically compared speech processing between left- and right-ear implanted SSD users on the one hand, and between these two patient groups and NH listeners on the other hand.

After cochlear implantation, speech intelligibility is typically assessed via behavioural measures, in particular word and sentence tests (Haumann et al. 2010; Hey et al. 2016, 2014; Hahlbrock, 1953; Hochmair-Desoyer et al. 1997). Event-related potentials (ERPs) however allow the *objective* evaluation of speech processing in CI users with a high temporal resolution (Luck, 2014). Previous studies have used ERPs in the electroencephalogram (EEG) to study cortical auditory processing in CI users (Beynon et al. 2005; Finke et al. 2016; Finke et al. 2015; Henkin et al. 2009; Sandmann et al. 2010; Sandmann et al. 2009; Groenen et al. 2001). Most of these studies applied an auditory oddball paradigm, in which a frequent standard sound and an infrequent deviant sound were pseudo-randomly presented, meaning that a deviant sound is followed by at least two standard sounds. Using this type of paradigm allows the study of the N1 ERP (negativity around 100 ms post stimulus) and the P2 ERP (positivity around 200–250 ms post stimulus). These ERPs are elicited in response to both the standard and deviant sounds and originate mainly from the auditory cortex (Crowley and Colrain, 2004; Näätänen and Picton, 1987). An additional deviant-related P3b response (positivity around 300–650 ms) is elicited if the central auditory system can discriminate between the standard and the deviant sound (Henkin et al. 2009). It has been widely assumed that the

P3b reflects the evaluation and classification of incoming auditory events (for a review, see Polich, 2007).

Most of the previous EEG studies on CI users have used an auditory oddball paradigm to study N1, P2 and P3b ERPs in individuals with *bilateral* hearing loss (Beynon et al. 2005; Finke et al. 2016, 2015; Henkin et al. 2009). ERPs of postlingually deafened adult CI users seem to be reduced in amplitude and prolonged in latency when compared to NH listeners, suggesting that CI users have difficulties in the sensory (N1, P2) and higher-level cognitive processing (P3b) of the limited CI input (Beynon et al. 2005; Finke et al. 2016; Henkin et al. 2014; Henkin et al. 2009; Sandmann et al. 2009). Moreover, adult CI users with bilateral hearing impairment show functional changes in the auditory cortex contra- and ipsilateral to the CI ear after implantation (Finke et al. 2015; Green et al. 2005). It is therefore not surprising that this group of patients can show altered functional asymmetry in the auditory cortex when compared to NH listeners, suggesting that auditory deprivation and/or cochlear implantation induce changes in the normal pattern of cortical response asymmetries. In contrast to implanted children with SSD (Lee et al. 2020; Polonenko et al. 2017), not much is known about functional changes in the adult auditory cortex of SSD CI patients. Knowledge about plasticity in the ipsi- and contralateral auditory cortex in SSD patients could help to understand the factors contributing to the CI outcome in these individuals. Thus, one principal aim of the present study was to evaluate the side-of-implantation effect on the functional asymmetry in the auditory cortex of adult SSD CI users.

Most of the previous studies using the oddball paradigm have been restricted to one stimulus pair (Billings et al. 2011; Sasaki et al. 2009). However, it is of clinical interest to develop a time-optimized multi-deviant oddball paradigm, which allows assessing multi-attribute auditory discrimination ‘profiles’. In the present study, we used a two-deviant oddball paradigm with one standard syllable and two types of deviant syllables of different acoustic–phonetic demand. In addition to NH controls, left- and right-ear implanted SSD CI users were tested sequentially on both ears. We systematically compared the behavioural and electrophysiological measures within and between the different groups of adult participants, which allowed us to address the following research questions:

- (1) Can a two-deviant oddball paradigm be used to objectively evaluate the speech discrimination ability in SSD CI users?
- (2) Do SSD CI users show differences in speech processing between the CI ear and the NH ear?
- (3) Is there a side-of-implantation effect on speech processing via the *CI ear* in SSD CI users?

- (4) Is there a side-of-stimulation effect on speech processing via the *NH ear* in SSD CI users and in NH listeners?

Following recent results on adult SSD CI users, we expected differences in behavioural and ERP measures between the CI ear and the NH ear in SSD CI users (Bönitz et al. 2018; Finke et al. 2016). In accordance with previous observations on the CI users with *bilateral* hearing loss, we hypothesised an altered functional asymmetry in the auditory cortex of SSD CI users when compared to NH listeners (Sandmann et al. 2009).

Material and Methods

Participants

Nineteen single-sided post-lingually deafened CI users participated in this study (six male; two left handers). All of them had no history of neurological or psychiatric disorders. All participants used their CI at a daily basis (15.5 ± 0.7 h/day) for at least ten months (mean = 18 months; sd = 8 months). The age ranged from 37 to 66 years (mean = 53.26 years; sd = 8.49 years). The duration of deafness before implantation varied from two months to 36 years (mean = 69 months; sd = 115 months). Since this variable is difficult to determine, anamnestic conversations were used to determine the time point at which a conventional hearing aid was no longer sufficient to understand speech. The duration of deafness was then calculated as the period between this time point and the cochlear implantation. All subjects were unilaterally implanted with a CI, nine of them on the right side and ten on the left side. Apart from two participants, all of the CI users were right-handed, as assessed by the Edinburgh inventory (Oldfield, 1971). According to previous studies with SSD CI users (Bönitz et al. 2018; Finke et al. 2016), the four pure tone average (4PTA—over 0.5, 1, 2 and 4 kHz) of the contralateral NH ear was ≤ 30 dB. Speech comprehension was tested using the Freiburg monosyllabic word test (Hahlbrock, 1970) and the Oldenburg sentence test (Wagener et al. 1999), the latter conducted with and without background noise. Here, all stimuli were presented via a loudspeaker placed at a distance of 1.6 m from the listeners head located at 0° in a soundproofed booth. Additionally, ten age-matched NH controls participated in this study (two male). Their age ranged from 41 to 70 years (mean = 53.2 years; sd = 9.37 years). Detailed information about the implant systems and the demographic variables of the participants can be found in Table 1.

Stimuli

The stimuli consisted of three different syllables which were taken from the Oldenburg logatome (OLLO) corpus, providing natural spoken language stimuli (Welge-Lüssen et al. 1997). The stimuli were generated by cutting the syllables /ki/, /ti/ and /ka/ out of the available logatomes from one speaker (male speaker 2, V6 “normal spelling style” and N3 “dialect”). All syllables had a duration of 300 ms and were normalised using the RMS function of the Adobe Audition software. The syllables differed by the place of articulation in the consonant contrast (/k/ vs. /t/) (Henkin et al. 2009) and by phonetic features in the vowel contrast (/a/ vs. /i/), in particular the vowel height and the vowel place (Micco et al. 1995). The German vowels /i/ and /a/ differ in the central frequencies of the first (F1) and second formant (F2). The formant values of /a/ are 730 Hz for F1 and 1284 Hz for F2. Regarding the vowel /i/, the formant values are 278 Hz (F1) and 2139 Hz (F2). In the German language, these formant values indicate the highest contrast between vowels, which should be perceivable by most of the CI users (Groenen et al. 2001). Unlike the aforementioned vowels, the contrast of the consonants /k/ vs. /t/ is very small, only differing in their place of articulation. The syllables /ki/ and /ti/ differ in rapid spectral changes in the transition of F2, which represents the articulatory movement from the consonant to the vowel (Kent, 1997). Those characteristics are very difficult to distinguish for CI users. In this study, we deliberately used one easier (/ki/ vs. /ka/) and one more difficult stimulus contrast (/ki/ vs. /ti/) to study the effects of auditory discrimination ability on behavioural and ERP measures in SSD CI patients.

In addition to the auditory oddball task with “original”, unprocessed syllables, the NH control group performed three additional blocks with degraded, “vocoded” syllables. This adjusted sound condition allowed to analyse the effects of stimulus degradation comparable to CI processing (Shannon et al. 1995). A noise vocoder was used to degrade the syllables (Gaudrain and Başkent, 2015). The MATLAB code is available online (see Gaudrain, 2016). The vocoder filtered the signal into four bands using 12th order, zero-phase Butterworth bandpass filters. The band boundaries were equally spaced based on a 35-mm basilar membrane distance (Greenwood, 1990) across a frequency range between 0.2 and 20 kHz. To extract the temporal envelope, the output of each band was half-wave rectified and low-pass filtered at 250 Hz (zero-phase fourth order Butterworth filter). The envelope was then multiplied by a wide-band noise carrier, and the resulting signal was summed across bands.

Task and Procedure

The experimental paradigm consisted of an auditory oddball task. The participants were presented with a frequent

Table 1 Demographic variables, audiologic information and implant information of the participants

ID	Age [years]	Handed-ness	CI side	Group	Processor	Etiology (details)	Duration deafness [months]	CI use [months]	PTA CI ear [dB]	PTA NH ear [dB]	Freiburg Monosyllabic test (65 dB) [%]	OLSA quiet [%]	OLSA noise [dB]
vp03	51	Right	Right	Non-proficient	CP910	Morbus meniere	19	24	32.5	19.75	75	96	− 3.2
vp04	38	Right	Right	Proficient	Opus	Sudden deafness	12	36	31.5	3.5	40	97	1.4
vp06	62	Right	Right	Non-proficient	CP910	Cholesteatoma surgery	312	15	34.25	30	30	47	20.1
vp10	43	Right	Right	Non-proficient	CP910	Sudden deafness	12	19	25.75	14.5	65	89	− 0.7
vp11	63	Right	Right	Proficient	CP910	Otitis media	18	12	34.25	14.75	10	23	Not applicable
vp14	61	Right	Right	Non-proficient	CP1000	Sudden deafness	7	11	24	15	65	90	− 0.4
vp15	55	Right	Right	Non-proficient	CP1000	Cogan 1 syndrome	3	10	22.5	10	95	99	− 2.6
vp18	49	Right	Right	Proficient	Sonnet	Otosclerosis	48	17	32.75	23.75	80	77	1.2
vp19	37	Right	Right	Non-proficient	Sonnet	Unknown	26	19	38.5	29	30	91	− 0.7
vp01	58	Right	Left	Non-proficient	CP910	Sudden deafness	84	25	26	12	65	95	− 0.7
vp02	56	Left	Left	Proficient	CP910	Sudden deafness	30	12	36.5	4.25	10	7	Not applicable
vp05	48	Right	Left	Proficient	CP910	Sudden deafness	144	32	23.5	12	100	98	− 1.1
vp07	55	Right	Left	Non-proficient	CP910	Sudden deafness	432	20	43.5	8	55	64	4.3
vp08	53	Left	Left	Proficient	CP910	Sudden deafness	5	32	33.75	15.75	50	79	2.4
vp09	57	Right	Left	Non-proficient	CP910	Sudden deafness	21	10	31.25	11.5	55	92	− 1.2
vp12	66	Right	Left	Non-proficient	CP910	Unknown	120	15	29.5	30	70	94	0.5
vp13	63	Right	Left	Proficient	CP1000	Stapes surgery	19	10	37	12.5	55	75	0.7
vp16	54	Right	Left	Proficient	Sonnet	Acute hearing loss	10	14	28.75	16	80	100	− 2.6
vp17	43	Right	Left	Proficient	CP1000	Petrous bone fracture	2	11	22.5	6.25	55	78	6.1

standard syllable (/ki/, 80% probability) and two infrequent deviant syllables (/ka/ and /ti/, 10% probability each). The participants were instructed to respond to deviant syllables via a button press of a computer mouse. The total of 800 trials were separated into three blocks with reasonable breaks in between. If a participant showed short response times, the inter-stimulus interval of 1400 ms was shortened accordingly. Hence, the measurement time added up to a maximum of 19 min (800 trials $\times$ 1400 ms) per ear. The stimuli were presented in a pseudo-randomised order with the constraint that a deviant syllable was preceded by at least three standard syllables. This was, however, not known by the participants. Prior to the experiment, a short training was performed at each ear. The participants sat in a comfortable chair in a sound attenuated booth. To avoid eye movements, the participants were instructed to look at a fixation cross on a computer monitor throughout the task.

In the *CI-only listening condition*, the processor of the CI users was put inside an aqua case from the manufacturer Advanced Bionics (<https://www.advancedbionics.com>) to avoid an additional stimulation of the NH ear. An earphone was inserted through a hole of the aqua case, where it was directly positioned over the microphone input of the CI. In general, all processors only fit into the aqua case with the compact batteries, which were provided by the clinic for each measurement. A long coil cable was used to connect the processor to the implant.

In SSD CI users, the use of an aqua case (in combination with an insert earphone) is advantageous compared to the stimulation via *direct connect* or loudspeakers for the following two reasons. First, the patients used CI processors from different manufacturers. A presentation of the stimuli via *direct connect* was not used to avoid an additional potential bias through the different ways of stimulus transfer into the different types of sound processors. Second, stimulus delivery via a *loudspeaker* would have been inappropriate, as this condition prevents a sufficient (passive) stimulus masking of the NH (i.e., non-tested) ear (Park et al. 2021).

Regarding the *NH-only condition* in SSD CI users and NH listeners, the stimuli were presented via inserted earphones positioned in the external auditory canal. The contralateral ear, that is, non-tested ear, was masked with an earplug in all conditions and groups. In general, the audio input level was calibrated to an acoustical input at 65 dB SPL. In addition, the participants performed a subjective rating before the start of the experiment. The loudness was readjusted to ensure that it was set to a moderate level in each individual, which is equivalent to a level of 60–70 dB SPL (e.g. Sandmann et al. 2009). This adjustment is important, since recent auditory brain imaging studies showed that the loudness can affect cortical activation (Zhou et al. 2022).

To have a measure of the subjective listening effort, participants were asked after every block of the experiment

to rate the effort of understanding the syllables on a scale between zero (no effort at all) and five (very demanding). In addition, we asked the participants to rate the difficulty to perform the task on a similar scale (between zero (no effort at all) and five (very demanding)).

Data Recording and Analysis

Behavioural Data: Auditory Oddball Task

The percentage of hit rates (hits) and individual mean response times (RT) of correct trials were analysed. Correct responses were defined as the occurrence of a button press in response to deviant syllables from 200 to 1200 ms following stimulus onset.

Electrophysiological Data: Recording and Data Processing

The EEG was continuously recorded with a BrainAmp DC amplifier from 30 active electrode sites, placed according to the extended 10/20 system (Brainproducts, <http://www.brainproducts.com>). An additional electrode was placed under the left eye for recording electrooculography (EOG), and the reference electrode was placed on the nose. The EEG was digitized at 1000 Hz, and the impedance was kept below 5 k Ω throughout the recording.

The data was analysed with EEGLAB (Delorme and Makeig, 2004) running in the MATLAB environment (R2020a; Mathworks). The EEG was downsampled to 500 Hz and offline filtered with a FIR-filter, using a high pass cut-off frequency of 0.1 Hz and a maximum possible transition bandwidth of 0.2 Hz (two times the cut-off frequency) plus a low-pass cut-off frequency of 40 Hz and a transition bandwidth of 2 Hz. For both filters, the Kaiser-window (beta = 5.653, max. stopband attenuation = -60 dB, max. passband deviation = 0.001) approach was used (Widmann et al. 2015). This approach maximises the energy concentration in the main lobe, thus averaging out noise in the spectrum and reducing information loss at the edges of the window (Widmann et al. 2015). Missing channels located over the region of the speech processor and transmitter coil were removed (mean and standard error: 0.6 ± 0.6 electrodes; range: 0–2 electrodes). The EEG data of the CI ear was segmented into epochs from -100 to 400 ms relative to the stimulus onset, and it was baseline corrected (-100 to 0 ms). Similar to previous studies, an independent component analysis (ICA) was then applied to the segmented data to identify the electrical CI artefact which spatially and temporally overlaps with auditory brain activity (Debener et al. 2007; Sandmann et al. 2010, 2009). After applying the ICA weights to the original (down-sampled and filtered (0.1–40 Hz) continuous EEG) data, all components that could be assigned to the electrical CI artefact

were removed. Subsequently, the EEG datasets of both sides (CI ear: after first ICA-based artefact reduction; NH ear: original, i.e. down-sampled and filtered (0.1–40 Hz)) were merged and segmented into 2 s dummy segments. Segments exceeding an amplitude threshold criterion of four standard deviations were removed, and a second ICA was applied. All components assigned to ocular artefacts and other sources of non-cerebral activity were removed (Jung et al. 2000). Afterwards, the removed channels over the CI were interpolated using a spherical spline interpolation, a procedure still allowing good dipole source localisation of auditory event-related potentials (ERPs) in CI users (Debener et al. 2007; Sandmann et al. 2009). Only correct responses (hits for deviant syllables and correct rejections for standard syllables) were included for ERP analysis. Subsequently, a peak analysis of ERPs was performed on single-subject averages measured at different regions-of-interest (ROIs). We defined a frontocentral and a parietal ROI based on the grand average computed across all conditions and participants. The *frontocentral ROI* included the channels FCz, FC1, FC2, Fz and Cz and was used to analyse the N1 and P2 ERPs. The *parietal ROI* included the channels Pz, P3, P4, CP1 and CP2 and was used for the peak detection of the P3b component. For ERP quantification, individual peak amplitudes and latencies were measured by detecting the maximum and latency of ERP peaks in commonly used latency bands of the N1, P2 and P3b ERPs (Luck, 2014; Picton, 2010; N1: 80–160 ms; P2: 180–300 ms; P3b: 300–900 ms).

Source Analysis

Cortical source activities were computed using the Brainstorm software (Tadel et al. 2011) and following the tutorial of Stropahl et al. 2018. Brainstorm applies the method of dynamic statistical parametric mapping of the data (dSPM; Dale et al. 2000). This method uses the minimum-norm inverse maps with constrained dipole orientations to estimate the locations of the scalp-recorded electrical activity of the neurons. It seems to localise deeper sources more accurately than standard minimum norm procedures, but the spatial resolution remains blurred (Lin et al. 2006). Prior to source estimation, the EEG data was re-referenced to the common average. Single-trial pre-stimulus baseline intervals (− 200 to 0 ms) were used to calculate individual noise covariance matrices and thereby estimate individual noise standard deviations at each location (Hansen et al. 2010). As a head model, the boundary element method (BEM) as implemented in OpenMEEG was used, providing three realistic layers and representative anatomical information (Gramfort et al. 2010; Stenroos et al. 2014). Source activities were evaluated in an a-priori defined auditory region-of-interest (ROI). The definition of the ROI was based on the Destrieux-Atlas implemented in Brainstorm (Destrieux

et al. 2010). The used auditory ROI comprised four smaller regions of the original atlas (G_temp_sup-G_T_transv, G_temp_sup-Plan_tempo, Lat_Fis-post, S_temporal_transverse). These regions were combined using the “merge scouts” feature in Brainstorm and approximated Brodmann areas 41 and 42. Peak activation magnitudes and latencies within this ROI were extracted for each individual participant. The activation data is given as absolute values with arbitrary units based on the normalisation within the dSPM algorithm.

Statistical Analyses

The subsequent statistical analysis was performed in R (Version 3.6.3, R Core Team 2020, Vienna, Austria). To address the four different research questions, the amplitudes and latencies of auditory ERPs were analysed separately on the sensor level (frontocentral ROI on head surface: N1, P2; parietal ROI on head surface: P3b) and on the source level (ERP source analysis: activation in ipsi- and contralateral auditory cortex at N1 latency). This was done by computing mixed ANOVAs with the between-subject factor “group” (proficient/non proficient CI users or left/right implanted patients) and the within-subject factors “stimulation side” (CI/NH), “stimulus type” (standard/deviant 1/deviant 2) and “hemisphere” (left/right). Significant interactions and main effects ($p < 0.05$) were followed-up by paired t-tests, and the Holm-Bonferroni approach was used for the correction of multiple comparisons (Holm, 1979). In the case of violation of sphericity, a Greenhouse–Geisser correction was applied.

Results

Question 1: Can a Two-Deviant Oddball Paradigm Be Used to Objectively Evaluate the Speech Discrimination Ability in SSD CI Users?

Behavioural Results

The mean RTs of the nineteen SSD CI users were analysed with a two-way ANOVA, including the within-subject factors “stimulated ear” (CI, NH) and “stimulus type” (deviant 1, deviant 2). A significant main effect of “stimulated ear” ($F_{1,13} = 23.82$, $p_{\text{adj}} \leq 0.001$, $\eta^2 = 0.22$) was followed up by pairwise comparisons, revealing faster RTs when syllables were presented via the NH ear compared to the CI ear ($p \leq 0.001$). A significant main effect of “stimulus type” ($F_{1,13} = 8.56$, $p_{\text{adj}} = 0.012$, $\eta^2 = 0.08$) was followed up by pairwise comparisons, revealing faster response times for deviant 1 compared to deviant 2 ($p = 0.04$).

In contrast to the RTs, the hit rates showed a more complex pattern of results. While all participants could

reliably discriminate deviant 1 from the standard sound (CI: $90.86 \pm 0.07\%$, NH: $92.17 \pm 0.05\%$), only some of the participants were able to reliably differentiate deviant 2 from the standard sound. Hence, the group of participants (including both the left- and right-ear implanted SSD CI users) was divided into two subgroups (proficient, non-proficient CI users) based on the median split in the behavioural performance (median of hit rate = 37.5%). In the following, these subgroups are referred to as proficient (performance > 37.5%) and non-proficient CI users ($\leq 37.5\%$).

The hit rates of the participants were analysed using a three-way mixed ANOVA, with the between-subject factor “group” (proficient, non-proficient) and the within-subject factors “stimulated ear” (CI, NH) and “stimulus type” (deviant 1, deviant 2). A significant three-way interaction ($F_{1,17} = 147.84$, $p_{\text{adj}} \leq 0.001$, $\eta^2 = 0.62$) was followed up by simple two-way interactions and pairwise comparisons. As expected, there was a significant difference between the two groups for the *stimulation of the CI ear* (Fig. 1b): The proficient users had higher hit rates for deviant 2 when compared to the non-proficient users ($p \leq 0.001$). By contrast, there was no group difference for the *stimulation of the NH ear*.

Comparing the response times between the proficient and non-proficient CI users revealed no significant difference between the two groups, neither for the CI ear nor for the NH ear.

ERPs: Proficient vs. Non-Proficient CI Users

The grand average ERPs from both the NH ear and the CI ear revealed an N1 and P2 response (Fig. 1c, Supplementary Fig. 1). In addition, the grand average ERPs showed a P3b ERP around 400–600 ms, which was observed in response to both deviant types in the proficient CI users, but which was restricted to deviant 1 in the non-proficient CI users (Fig. 1c).

In a first step of the ERP analysis, we focused on the N1 and P2 ERPs. We computed a two-way mixed ANOVA for the *N1 and P2 ERPs*, with the between-subject factor “group” (proficient, non-proficient) and the within-subject factor “stimulus type” (deviant 1, deviant 2). This was done separately for the stimulation over the CI ear and the NH ear. Regarding the *stimulation over the CI ear*, we found a main effect of “group” for the N1 amplitude ($F_{1,17} = 4.16$, $p_{\text{adj}} = 0.057$, $\eta^2 = 0.16$) and the N1 latency ($F_{1,17} = 5.68$, $p_{\text{adj}} = 0.029$, $\eta^2 = 0.14$), respectively. The pairwise comparisons revealed a significantly enhanced and prolonged N1 ERP for the non-proficient CI users when compared to the proficient CI users (averaged over all three stimulus types; amplitude: $p = 0.02$; latency: $p = 0.03$). No group differences were found in the P2 component (Supplementary Fig. 1a).

For the *stimulation over the NH ear*, we found no group effect for the N1 ERP, but a main effect of “stimulus type” ($F_{1,17} = 42.99$, $p_{\text{adj}} = 0.001$, $\eta^2 = 0.16$), which was followed up by pairwise comparisons. This analysis revealed an enlarged N1 amplitude for deviant 2 compared to deviant 1 (averaged over both groups; $p = 0.01$). Regarding the P2 component,

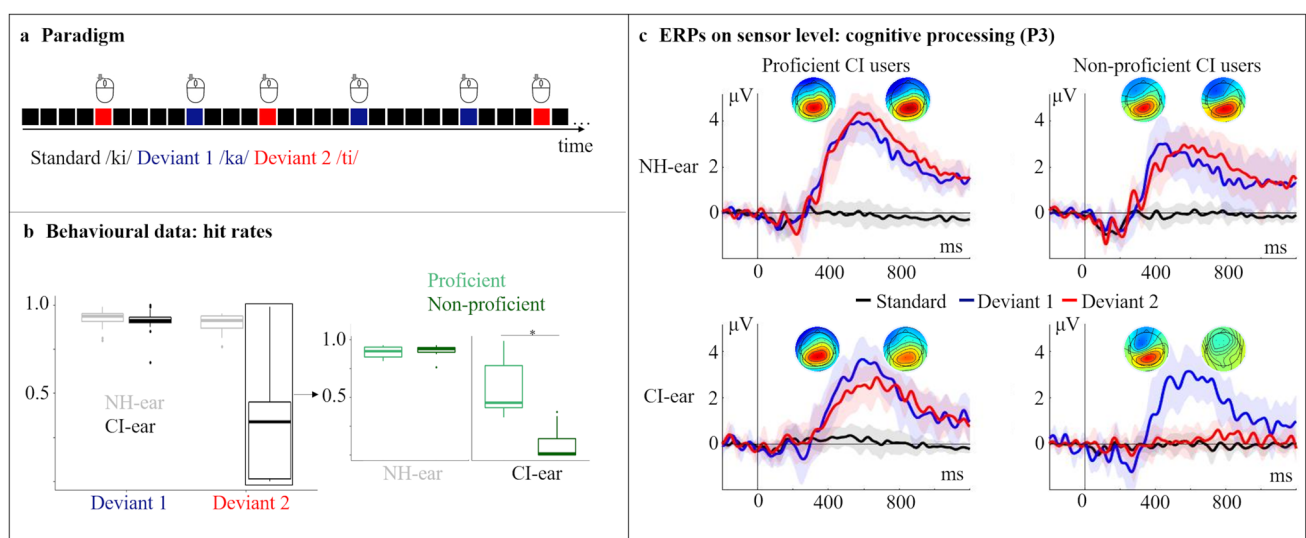


Fig. 1 **a** Depiction of the oddball paradigm with frequent standard and two deviant syllables. **b** Hit rates for the NH- and CI-ear separately for both deviant types plus the separation into two groups based on the hit rates for deviant 2. **c** ERPs showing the objectification of the groups by the P3 component with topographic plots. The

shaded areas indicate the 95% confidence intervals. For illustration purposes, the ERPs are low-pass filtered (10 Hz). The left topography belongs to deviant 1 (blue line) and the right topography belongs to deviant 2 (red line), respectively. Time range for topographic plots: 555–575 ms

we found a significant two-way interaction between the factors “group” and “stimulus type” ($F_{1,17}=4.56$, $p_{\text{adj}}=0.048$, $\eta^2=0.03$). The subsequent pairwise comparisons showed a greater P2 amplitude for deviant 2 for non-proficient CI users compared to the proficient CI users ($p=0.047$).

In a second step, we focused on the P3b ERP. We computed a two-way mixed ANOVA with the between-subject factor “group” (proficient, non-proficient) and the within-subject factor “stimulus type” (deviant 1, deviant 2) separately for the CI ear and the NH. For *stimulation via the CI ear*, the analyses revealed a two-way interaction between the factors “group” and “stimulus type” ($F_{1,17}=6.13$, $p_{\text{adj}}=0.002$, $\eta^2=0.13$), showing a significantly enhanced P3b amplitude for the proficient CI users specifically for deviant 2 when compared to the non-proficient CI users ($p\leq 0.001$). No group difference was observed for the *stimulation via the NH ear* ($p=0.21$). However, the stimulation via the NH ear showed a main effect for “stimulus type” ($F_{1,17}=8.6$, $p_{\text{adj}}=0.009$, $\eta^2=0.13$), resulting in a significantly prolonged P3b latency for deviant 2 when compared to deviant 1 ($p=0.03$). Furthermore, we found a significant positive correlation between the P3b amplitude and the hit rate of deviant 2 in both groups (proficient CI-users: $R=0.8$, $p=0.009$; non-proficient CI-users: $R=0.85$, $p=0.002$).

In sum, the results concerning question 1 revealed that ERPs, which are recorded in the context of a two-deviant oddball paradigm, show differences in initial sensory and later cognitive speech processing between different subgroups of SSD CI users. Specifically, non-proficient and proficient SSD CI users can be distinguished on the basis of the N1 and P3b amplitudes (for stimulation via the CI ear) as well as on the basis of the P2 amplitude (for stimulation via the NH ear). These findings suggest that the two-deviant oddball paradigm can be used to assess speech discrimination proficiency in SSD CI users.

Question 2: Do SSD CI Users Show Differences in Speech Processing Between the CI Ear and the NH Ear?

Since not all participants were able to reliably identify deviant 2, this condition was removed for further statistical analyses.

Behavioural Results: CI Ear vs. NH Ear (SSD CI Users)

The comparison of the behavioural results for deviant 1 between the CI ear and the NH ear in SSD CI users (regardless of the side of implantation) showed no differences in hit rates but significantly faster response times (Fig. 2a) for the stimulation of the NH ear compared to the CI ear ($t(18)=-5.12$, $p\leq 0.001$, $d=0.92$). Furthermore, the listening effort for syllable processing via the

CI ear was significantly enhanced compared to the NH ear ($t(18)=-2.14$, $p=0.047$, $d=0.50$).

ERPs on Sensor Level: CI Ear vs. NH Ear (SSD CI Users)

Figure 2 shows the ERPs in response to the standard and deviant 1, separately for the CI ear and the NH ear. The “stimulus type” (standard, deviant 1) x “stimulated ear” (CI, NH) ANOVA revealed a significant main effect of “stimulated ear” for the N1 amplitude ($F_{1,18}=16.54$, $p_{\text{adj}}\leq 0.001$, $\eta^2=0.13$) and latency ($F_{1,18}=8.72$, $p_{\text{adj}}=0.009$, $\eta^2=0.13$), respectively. The follow up pairwise comparisons revealed a smaller N1 amplitude ($p=0.001$) and a prolonged N1 latency ($p=0.002$) for the CI ear compared to the NH ear (averaged over both stimulus types). Similarly, the P2 amplitude revealed a significant main effect of “stimulated ear” ($F_{1,18}=12.10$, $p_{\text{adj}}=0.003$, $\eta^2=0.07$), resulting in a significantly smaller amplitude for the CI ear than for the NH ear ($p=0.02$). Paired t-tests between the P3b amplitudes and latencies of deviant 1 showed a prolonged latency for the CI ear compared to the NH ear ($t(18)=-27.27$, $p=0.014$, $d=0.62$) but no ear difference in the P3b amplitude (Fig. 2c).

ERPs on Source Level: CI Ear vs. NH Ear (SSD CI Users)

Figure 2d shows the source activity separately for the two stimulation conditions (CI ear, NH ear) in the bilateral auditory cortex. Given that the ERP analysis on the sensor level did not reveal a significant effect of “stimulus type” (standard, deviant 1), the ERP analysis on the source level was restricted to the averages computed across the two stimulus types. The paired t-tests revealed a significantly delayed cortical response at N1 latency for the stimulation via the CI ear when compared to the NH ear ($t(18)=29.64$, $p=0.008$, $d=0.72$). No significant difference was found for the amplitude of the source activity at N1 latency range.

Behavioural and ERP Results: Vocoded vs. Original Sounds (NH Listeners)

To evaluate whether the observed differences between the CI ear and the NH ear originate from the CI-related degradation of the stimuli (hypothesis 1) or from cortical plasticity (hypothesis 2), we compared the behavioural and ERP results between the two stimulus conditions “vocoded” and “original” syllables (separately for the two stimulus types: standard/deviant 1) within the group of NH listeners.

Regarding the *behavioural results*, the NH control group did not show any significant difference between the vocoded and the original syllables (Supplementary Fig. 2a). But the subjective listening effort for syllable processing with

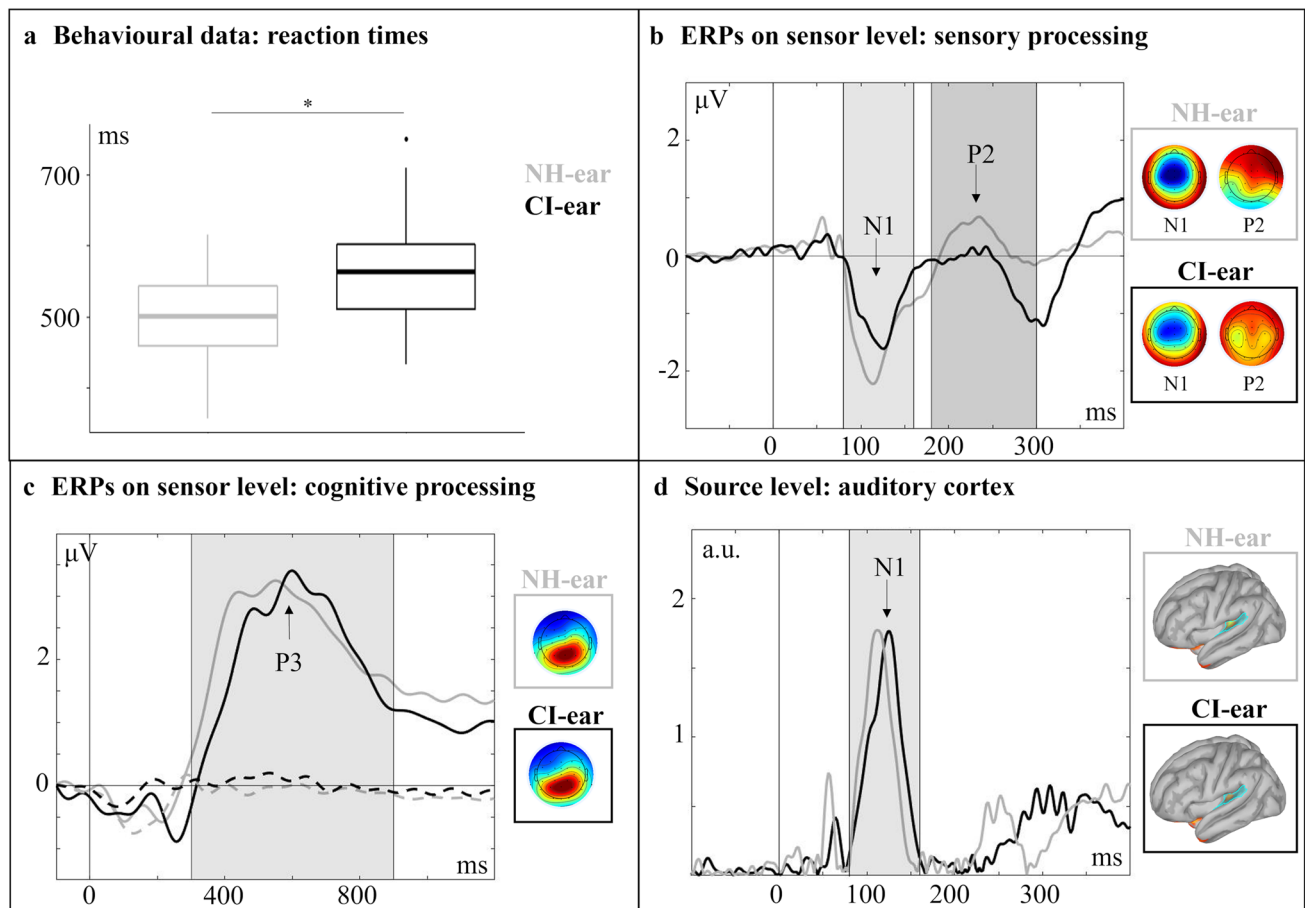


Fig. 2 Comparisons between the NH- and the CI-ear in SSD patients on different levels. **a** Faster reaction times for stimulation over the NH-ear for deviant 1. **b** ERP-averages across standard and deviant 1 show a smaller and prolonged N1 component and a smaller P2 component for stimulation over the CI-ear. Time range for topographic plots: N1=100–130 ms (NH-ear)/110–140 ms (CI-ear), P2=220–250 ms. **c** ERPs to deviant 1 show a prolonged P3 component for

stimulation over the CI-ear. Dotted lines represent responses to the standard syllable. Time range for topographic plots: 530–570 ms (NH-ear)/580–620 ms (CI-ear). **d** Source analysis shows a response in the auditory cortex at N1 latency which is delayed for stimulation via the CI. The blue area represents the used region of interest, red areas show activation. Time points for brain plots: 118 ms (NH-ear)/134 ms (CI-ear)

vocoded stimuli was significantly enhanced compared to the original stimulation ($t(9) = -2.42$, $p = 0.039$, $d = 0.78$).

Concerning the *ERPs on the sensor level*, the supplementary Fig. 2b shows the waveforms of the NH control group separately for the “vocoded” and “original” stimulus conditions. The two-way ANOVA with the within-subject factors “condition” (vocoded/original) and “stimulus type” (standard/deviant 1) revealed no main effects and no significant interaction for the N1 ERP. However, the same ANOVA computed separately for the P2 ERP revealed a significant main effect of “condition” ($F_{1,9} = 17.69$, $p_{\text{adj}} = 0.002$, $\eta^2 = 0.15$), resulting in a significantly larger P2 amplitude for the stimulation with vocoded syllables compared to the stimulation with the original syllables ($p = 0.02$). Finally, paired t-tests comparing the P3b ERP between the two stimulus conditions (vocoded/original) revealed no statistical difference in P3b amplitudes and latencies.

Concerning the *ERPs on the source level*, the supplementary Fig. 2d shows the source waveforms in the bilateral auditory cortex separately for the “vocoded” and “original” stimulus conditions. Paired t-tests comparing the two simulation conditions (vocoded/original) showed no statistical difference in the source activity at the N1 latency range, neither for the amplitude nor for the latency.

In sum, the findings on question 2 revealed that syllable processing via the CI ear—when compared to the NH ear—results in prolonged response times, enhanced subjective listening effort, and ERPs with reduced amplitudes (N1, P2) and prolonged latencies (N1, P2, P3b). These results suggest that the CI-related stimulus degradation leads to difficulties in speech processing in SSD CI users, not only at initial sensory but also at later cognitive processing stages.

Question 3: Is there a Side-of-Implantation Effect on Speech Processing via the CI Ear in SSD CI Users?

Behavioural Results and ERPs on sensor level: Left CI vs. right CI (SSD CI users)

The group of nineteen SSD CI users was divided into two subgroups according to the implantation side. Ten participants were implanted on the left ear and nine on the right ear (Table 1). For the behavioural results (hit rates, response times) and the ERPs (amplitude and latency of N1, P2, P3b ERPs), we computed unpaired t-tests between the groups (left-implanted, right-implanted) separately for the CI ear and the NH ear. However, the results did not show any statistical differences between the left- and right-ear implanted SSD CI users.

ERPs on Source Level: Left CI vs. Right CI (SSD CI Users)

Figure 3a shows the activity in the left and right auditory cortex separately for left- and right-ear implanted SSD CI users when stimulated over the CI. A two-way mixed

ANOVA with the between-subject factor “group” (left/right implanted) and the within-subject factor “hemisphere” (left, right) revealed a significant two-way interaction ($F_{1,17}=9.043$, $p_{\text{adj}}=0.008$, $\eta^2=0.17$). Post-hoc t-test revealed a hemispheric difference for the left-implanted group ($p=0.031$), with enhanced activity in the right than left auditory cortex. By contrast, the right-implanted group did not show a hemispheric difference in auditory-cortex activation.

Does the Side-of-Implantation Effect Arise from the Stimulus Degradation Through the CI or from Intra-Modal Plasticity in the Auditory Cortex?

To verify whether the observed hemispheric differences between the two SSD groups arise from the stimulus degradation through the CI (hypothesis 1) or from intra-modal plasticity in the auditory cortex (hypothesis 2), we compared the behavioural and ERP results between the two stimulus conditions “vocoded” and “original” syllables (separately for the two stimulation sides: left/right) within the group of NH listeners. We computed two-way ANOVAs including the

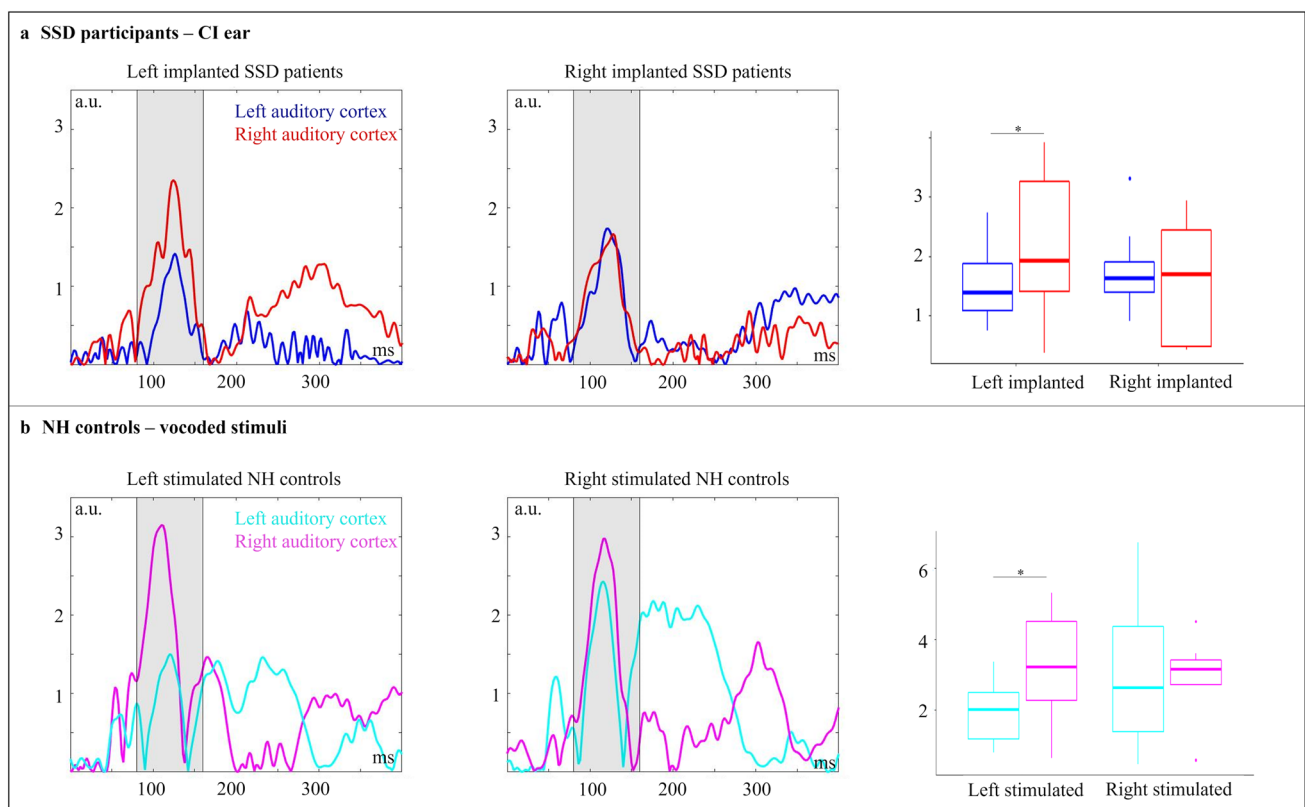


Fig. 3 Activity in the left and right auditory cortex shows differences in the latency range of the N1 component for the groups. **a** Mean activities plus boxplots for all SSD patients, separated by implantation side, stimulated via the CI are shown. **b** Mean activities plus boxplots for all NH control subjects, separated by the stimulation side,

for the stimulation with vocoded syllables are shown. Hemispheric asymmetries specifically for left implanted SSD patients and left stimulated NH controls, as indicated by enhanced N1 amplitude in the right compared to left auditory cortex

within-subject factors “side of stimulation” (left/right) and “condition” (original/vocoded stimuli).

Regarding the behavioural results, we observed a significant main effect of “side of stimulation” for the hit rates ($F_{1,19} = 14.26$, $p_{\text{adj}} = 0.001$, $\eta^2 = 0.10$), which was due to a higher hit rate for stimulation via the right ear when compared to the left ear (averaged over both conditions; $p = 0.006$). All other analyses revealed no statistical differences between the stimulation sides, neither in response times, nor in ERP measures (amplitudes and latencies of N1, P2, P3b ERPs).

Figure 3b illustrates the activity in the left and right auditory cortex for the NH control group when stimulated with vocoded syllables via the left and right ear. A two-way ANOVA with the within-subject factors “side of stimulation” (left/right) and “hemisphere” (left/right) revealed a significant two-way interaction ($F_{1,9} = 9.85$, $p_{\text{adj}} = 0.012$, $\eta^2 = 0.07$). Post-hoc t-tests showed a hemispheric difference in auditory-cortex activation for the stimulation of the left ear ($p = 0.028$), with an enhanced activation in the right than left auditory cortex. By contrast, there was no hemispheric difference in auditory-cortex activation for the stimulation of the right ear.

In sum, the results concerning question 3 revealed no side-of-implantation effect on *behavioural* speech discrimination abilities in SSD CI users, although the NH listeners showed in general enhanced hit rates for the right-ear than the left-ear stimulation condition (i.e., regardless of whether the syllables were “original” or “vocoded”). In contrast to the behavioural findings, however, the ERP analyses revealed a side-of-implantation effect on auditory cortex functions for the SSD CI users, with enhanced hemispheric difference in auditory-cortex activation for the left-ear than the right-ear implanted individuals. A consistent pattern of hemispheric asymmetry was observed in the NH listeners, in particular when these individuals were tested with vocoded stimuli, that is, in approximated sound conditions. This suggests that the side-of-implantation effect on auditory-cortex asymmetry mainly originates from the CI-related degradation of the stimuli (i.e., confirmation of hypothesis 1).

Question 4: Is there a Side-of-Stimulation Effect on Speech Processing via the NH Ear in SSD CI Users and in NH Listeners?

Behavioural Results and ERP Results on Sensor Level: Left Side vs. Right Side in SSD CI Users and NH Controls

We statistically compared the behavioural and ERP measures (for “original” syllables) for the NH ear of SSD CI users between the left and right implanted patients, hence the side-of-implantation effect on the NH ear. No

differences were found for any behavioural measures, nor for the sensory ERP components (N1, P2). The t-test between the two groups for the higher-cognitive P3b component revealed a statistically significant difference in latency ($t(16,98) = -3.18$, $p = 0.005$, $d = 1.45$), with a prolonged latency for the right SSD CI users (NH ear on the left side). The same analysis for the NH control group did not show any differences between the sides of stimulation.

ERPs on Source Level: Left Side vs. Right Side in SSD CI Users and NH Controls

Figure 4a shows the activation of the left and right implanted SSD patients in the left and right auditory cortex when stimulated over their NH ear. A two-way mixed ANOVA with the between-subject factor “group” (left/right implanted) and the within-subject factor “hemisphere” (left/right) showed a significant two-way interaction ($F_{1,17} = 5.91$, $p_{\text{adj}} = 0.026$, $\eta^2 = 0.08$). Post-hoc tests revealed a hemispheric difference in auditory-cortex activation for the stimulation via the right NH ear (left implanted group; $p = 0.009$), with greater activation in the left than right auditory cortex. By contrast, there was no hemispheric difference in the right-implanted SSD group when stimulated via the (left) NH ear. Additionally, a trend for a difference between the two SSD groups (when stimulated via the NH ear) was observed in the left auditory cortex, showing more activation for the stimulation of the right NH ear (left-implanted SSD users) compared to the left NH ear (right-implanted SSD users; $p = 0.076$).

The comparison of the two NH ears of the NH control group revealed no significant hemispheric differences in auditory-cortex activation, neither for the stimulation of the left nor the right ear (Fig. 4b).

In sum, the findings concerning question 4 revealed that stimulation via the NH ear in SSD CI users resulted in comparable *behavioural* speech discrimination abilities between left-ear and right-ear implanted individuals. However, the two groups of SSD CI users differed in the pattern of functional asymmetries in the auditory cortex. When stimulated via the NH ear, the left-ear implanted SSD CI users (NH ear on the right side) revealed a stronger auditory-cortex asymmetry than the right-ear implanted individuals (NH ear on the left side). The NH listeners in general showed a less pronounced side-of-stimulation effect. These results suggest a side-of-implantation effect on speech processing via the NH ear for SSD CI users, possibly caused by implantation-side specific alterations in the ipsilateral and contralateral auditory cortex as induced by temporary deafness and/or degraded sensory input after implantation (i.e., confirmation of hypothesis 2).

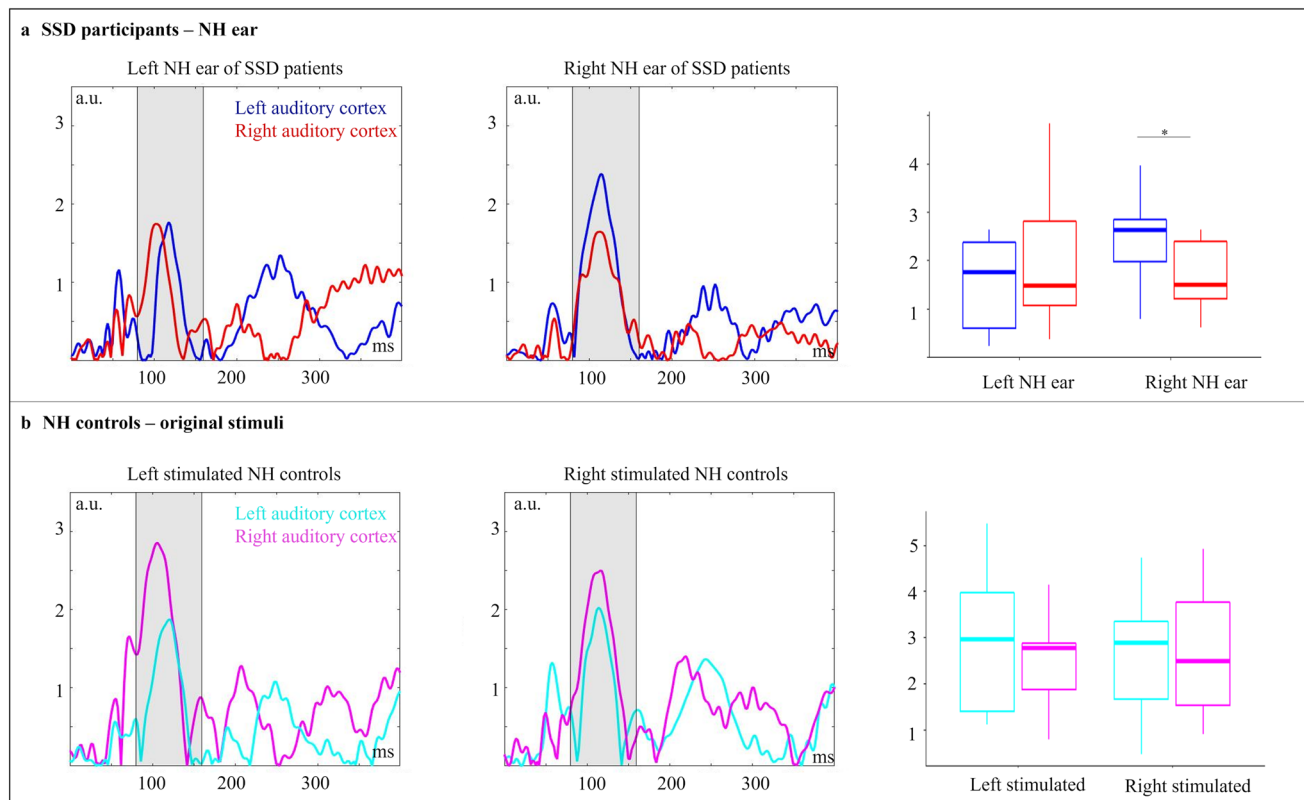


Fig. 4 Activity in the left and right auditory cortex shows differences in the latency range of the N1 component for the groups. **a** Mean activities plus boxplots for all SSD patients, separated by implantation side, stimulated via the NH ear are shown. The left NH ear belongs to the right implanted SSD patients and the right NH ear belongs to the left implanted SSD patients. **b** Mean activities plus

boxplot for all NH control subjects, separated by the stimulation side, for the stimulation with original syllables are shown. Hemispheric asymmetries specifically for right NH ear of the SSD patients, as indicated by enhanced N1 amplitude in the left compared to right auditory cortex

Discussion

The present study compared event-related potentials (ERPs) between left-ear and right-ear implanted single-sided deaf (SSD) cochlear implant (CI) users on the one hand, and between these two patient groups and normal-hearing (NH) controls on the other hand. We used a two-deviant-oddball paradigm and ERP source analysis to evaluate differences in cortical speech processing between acoustic and electrical hearing as well as between proficient and non-proficient SSD CI users. Our results revealed that proficient and non-proficient CI users can be distinguished on the basis of N1 and P3b ERPs to speech sounds. Further, the results suggest that processing via the CI is more difficult than via the NH ear, as indicated by longer response times, higher subjective listening effort and ERPs with reduced amplitudes (N1, P2) and prolonged latencies (N1, P2, P3b; Finke et al. 2016, 2015; Henkin et al. 2009; Sandmann et al. 2009). Further, we found a stronger contralateral dominance of activation in the auditory cortex at N1 latency for left-ear than right-ear implanted SSD CI patients, regardless of whether these

individuals were tested with the CI ear or the NH ear. A contralateral dominance for left-ear stimulation was also observed in the NH control group, which however was particularly present in the “vocalized” sound condition. We conclude that SSD CI users show a side-of-implantation effect on speech processing over both the CI and the NH ear. The next paragraphs focus on the four research questions addressed in the present study.

Question 1: Can a Two-Deviant Oddball Paradigm Be Used to Objectively Evaluate the Speech Discrimination Ability in SSD CI Users?

Speech recognition ability after cochlear implantation is heterogeneous, meaning that many patients reach open-set speech recognition ability while others do not (Blamey et al. 2013). Such variability can also be observed in SSD CI users (Speck et al. 2021), as confirmed by the current results. Although all of our CI users were able to discriminate between the syllables /ki/ (standard) and /ka/ (deviant 1), several CI users showed difficulties in the discrimination

of the syllable contrast /ki/ (standard) versus /ti/ (deviant 2). These findings were expected because the vowel contrast /i/ versus /a/ (deviant 1) is the highest possible contrast in the German language. By contrast, the consonant contrast /k/ versus /t/ (deviant 2) mostly refers to rapid spectral changes in the transition of the second formant (F2), making it more difficult to be distinguished. Not surprisingly, half of our CI users showed a hit rate below 37.5% for deviant 2, which confirms previous observations of impaired discrimination ability in CI users (Sandmann et al. 2010). It seems that difficulties with electrical hearing are caused by different factors, among them the limited spectral and temporal information provided by the implant, the spread of neural excitation in the cochlea, as well as physiological deficiencies in the auditory nerve (Drennan, 2008; Friesen et al. 2001; Kral, 2007; Nadol et al. 1989; Wilson and Dorman, 2008). Auditory deprivation not only reduces metabolism in the auditory cortex contralateral to the hearing-impaired ear (Speck et al. 2020) but also induces a reorganisation of the central auditory system, which may impair the cortical adaptation to the new artificial CI signal after implantation (Lee et al. 2001; Sandmann et al. 2012). Taken together, it seems that several individual factors contribute to the variability in speech discrimination ability observed in SSD CI users.

Our results on the stimulation via the CI ear also revealed a relationship between the (behavioural) auditory discrimination ability and objective ERP measures. CI users who were *better* able to discriminate between the syllables /ki/ and /ti/—here referred to as proficient CI users—showed a *reduced* and *delayed* N1 ERP when compared to the non-proficient CI users, who showed impaired syllable discrimination ability (Supplementary Fig. 1a: bottom left and right; Supplementary Fig. 1b). Importantly, the factor “age” cannot explain the reduced N1 ERPs in the proficient CI users, as the two groups had a comparable age, and a supplementary analysis revealed that the correlation between N1 ERPs and age was not significant (Supplementary Fig. 1c). Thus, our results suggest that the N1 ERP can distinguish between proficient and non-proficient CI users. Similarly, previous neuroimaging studies have suggested significantly different cortical activation patterns between lower and higher CI performers (Kessler et al. 2020), and an association between the recruitment of the auditory cortex and improvement in speech recognition ability over the first months after cochlear implantation (Sandmann et al. 2015).

Despite agreements with previous findings, the present ERP results indicate that the *non-proficient* CI performers revealed an *enhanced* N1 response when compared to the proficient CI performers, which however contradicts previous observations from positron-emission-tomography (PET) studies, reporting an *enhanced* cortical activation for individuals with *better* speech recognition ability (Giraud et al. 2001; Green et al. 2005). This discrepancy in results may

be explained by differences in methodology, including the measurement technique (PET versus EEG), the task (passive versus active), the type of auditory stimuli (words versus syllables), and the tested groups of CI users (bilateral hearing loss versus SSD). In general, an enhanced N1 amplitude (observed in non-proficient CI users) indicates a larger population of activated neurons in the auditory cortex, a stronger synchronisation of this neural activity, or a combination thereof. Thus, the current results may point to an enhanced recruitment and/or more synchronised neural activity in the auditory cortex in the non-proficient as compared to the proficient CI users. It is likely that these functional alterations in non-proficient CI users reflect an enhanced allocation of attentional resources to process small acoustic changes in speech sounds in the context of a discrimination task.

It has been previously suggested for CI users with bilateral hearing loss that ERPs can be used to objectively evaluate the auditory discrimination ability (Henkin et al. 2009; Sandmann et al. 2010; Soshi et al. 2014). Most of these previous studies used an auditory oddball paradigm and focused on the P3b response, which is elicited by infrequent, task-relevant changes in stimuli, and which shows maximal amplitudes over parietal scalp locations (Polich and Comerchero, 2003). Different models exist to describe the P3b component (Verleger, 2020). For instance, the P3b has been proposed to be a correlate of decision making (O’Connell et al. 2012), and to reflect voluntary attention to the task-relevant target stimuli (Polich, 2007). However, the elicitation of such a response requires that the individual can distinguish the acoustical differences between the task-relevant and task-irrelevant events. Indeed, previous results with CI users have pointed to a connection between the behavioural discrimination ability and the P3b ERP. They have revealed that the P3b in response to deviant sounds is comparable between CI users (with bilateral hearing loss) and NH listeners if the acoustic cues are well distinguishable by the participants (Henkin et al. 2009). However, in situations with more difficult stimulus contrasts, the CI users’ P3b response was reduced in amplitude and prolonged in latency (Henkin et al. 2009). The current study with SSD CI users confirms these previous observations. Our proficient and non-proficient CI users were well able to discriminate the syllable contrast /ki/ versus /ka/ (deviant 1), both when tested via the NH ear and via the CI ear (Fig. 1b), and the P3b amplitude elicited to this syllable contrast was comparable across groups and stimulated ears (Fig. 1c). However, we found group differences in the P3b for the syllable contrast /ki/ versus /ti/ (deviant 2), specifically when the CI users were tested via the CI ear (Fig. 1c: bottom left and right). Proficient CI users, but not non-proficient CI users, showed a P3b in response to deviant 2. Additionally, we found significant correlations between the hit rates and the P3b amplitude in both groups, pointing to a direct connection between the discrimination ability

and the P3b amplitude. Taken together, our results confirm previous studies by showing that the P3b is an appropriate ERP component to study higher-order cognitive processing in SSD CI users (Bönitz et al. 2018; Finke et al. 2016; Wedekind et al. 2021). Further, our results extend previous findings by demonstrating that the P3b can serve as an objective index for the behavioural speech discrimination ability in SSD CI users. Regarding the clinical application of the P3b, however, future studies should replicate and extend our results with more complex stimuli, for instance words. The use of similar stimuli in ERP recordings and common clinical test procedures (e.g., Freiburg monosyllabic word test) would allow even better comparability between the results of the electrophysiological P3b response and the behavioural word recognition ability obtained by speech audiometry.

Our results provide further evidence that ERP measures, in particular the N1 and the P3b, can differentiate between proficient and non-proficient CI users. This is consistent with other studies, reporting that ERPs measures—in particular the mismatch negativity (MMN; latency around 150–200 ms) and the P3b (latency around 300–650 ms) — can distinguish between CI users who have better versus lower abilities to discriminate speech sounds (Henkin et al. 2009; Turgeon et al. 2014). Thus, there is converging evidence that objective ERP measures can be used to assess behavioural speech recognition ability in CI users. However, the application of an EEG paradigm in the clinical context poses the challenge that the recording time should be as minimal as possible. Pakarinen et al. (2009) proposed a fast multi-feature paradigm for the recording of the mismatch negativity (MMN) to different speech sounds in the same recording session. Although this type of paradigm is very promising, the MMN is recorded in a passive condition and thus has a much smaller signal-to-noise ratio when compared to the P3b response. It is therefore reasonable to design a time-efficient *active* oddball paradigm, which allows to measure the more pronounced P3b in response to several speech stimuli and in a time short enough to avoid problems with vigilance, motivation, or restlessness of the patient. Our results are promising as they demonstrate that a two-deviant oddball paradigm is suitable to assess syllable discrimination proficiency in SSD CI users. Further, our findings extend previous reports by showing that cortical AEPs can be used in SSD CI users to objectify not only the detection (Távora-Vieira et al. 2018) but also the discrimination of speech sounds. Thus, our results suggest that this paradigm could be useful in the clinical context, as it allows the objective monitoring of the auditory rehabilitation in different acoustic dimensions after cochlear implantation. To extend our findings, which are limited to syllables, the paradigm should be extended to more complex speech stimuli such as words. Importantly, the objective ERP measures could indicate whether the custom setting of the CI

is sufficient for detailed speech discrimination ability, and whether renewed adjustments in certain frequency ranges could be useful. This is particularly important for patients with an ambiguous constellation of behavioural results.

Question 2: Do SSD CI Users Show Differences in Speech Processing Between the CI Ear and the NH Ear?

Our participants showed slower response times for the processing of syllables via the CI ear compared to the NH ear. This is consistent with recent work on SSD CI users, showing for the CI ear prolonged behavioural responses not only to sinusoidal tones (Bönitz et al. 2018) but also to words (Finke et al. 2016). Given the temporally and spectrally limited signal quality of the CI input, it is highly likely that these slower response times reflect enhanced difficulties to process the speech sounds via the CI ear when compared to the NH ear (Beynon et al. 2005; Groenen et al. 2001; Kelly et al. 2005). This interpretation is supported by our observation that SSD CI users report an enhanced listening effort for syllable processing via the CI ear as compared to the NH ear.

Similar to the behavioural results, we found an effect of stimulation type (acoustic versus electric) on ERPs, not only on the sensor level but also in the auditory cortex. For the CI ear, the ERPs (on the sensor level) were smaller in amplitude (N1, P2) and prolonged in latency (N1, P3b). These results are consistent with previous EEG studies on SSD CI users (Bönitz et al. 2018; Finke et al. 2016; Legris et al. 2018). CI-related effects on ERPs are also suggested by the current N1 source analysis, showing a prolonged cortical response to syllables when processed via the CI ear as compared to the NH ear (Fig. 2d). In sum, our ERP results are consistent with our behavioural observations since they suggest difficulties in speech processing via the CI ear, both at initial sensory and later cognitive processing stages. It is likely that the ERP differences between the CI ear and the NH ear are caused by CI-related stimulus degradation and/or cortical reorganisation as induced by temporary deafness and/or cochlear implantation (e.g., Sandmann et al. 2015).

To analyse the specific effect of CI-related stimulus degradation on speech processing, we compared the behavioural and ERP results of NH listeners between the two stimulus conditions “original” and “vocoded” speech sounds. The behavioural results did *not* reveal significant differences in hit rates and response times between the two sound conditions (Supplementary Fig. 2a). Further, N1 ERPs were comparable for “original” and “vocoded” syllables, both on the sensor level (Supplementary Fig. 2d) and in the auditory cortex (Supplementary Fig. 2d). These results indicate that the NH listeners could well distinguish the different syllable pairs, regardless of the CI-related stimulus degradation. Further, our results suggest that the cortical speech processing

at N1 latency was not significantly affected by stimulus degradation in NH listeners. Given that noise-band vocoder simulations used in NH listeners allow a good approximation to the performance levels of CI users (Friesen et al. 2001), our results indicate that the attenuated and prolonged N1 ERP for the CI ear in our SSD patients cannot be explained by the degraded CI input alone. It seems more likely that this latency effect for the CI ear is at least partially caused by intra-modal plasticity in the auditory cortex of SSD CI users. Indeed, a previous prospective longitudinal study on CI users (with bilateral hearing loss) has shown that the N1 latency reduces over the first year of implant use and approaches the levels of NH listeners, but remains delayed, even after one year of CI experience (Sandmann et al. 2015). This observation indicates limitations in the capacity of the auditory cortex to adapt to the CI signal after implantation. Taken together, our results suggest that differences in speech processing between the CI ear and the NH ear in SSD CI users are at least partially related to limitations in cortical adaptation to the implant signal, causing difficulties in the sensory and cognitive processing when speech is perceived via the CI ear.

Unlike our findings about the N1 ERP, we observed an effect of stimulus degradation on the P2 response in the NH control group, with *enhanced* P2 amplitude for the “vocoded” sounds when compared to the “original” sounds (Supplementary Fig. 2b). With regards to the SSD group, however, speech processing via the CI ear resulted in a *smaller* P2 amplitude when compared to the NH ear. Two different mechanisms may account for this group specific differences in P2 amplitude modulation, in particular 1) training-related alterations of sound representation, and 2) allocation of attentional resources. Regarding the *first mechanism*, previous EEG studies have proposed that the training-related enhancement of the auditory P2 response represents an electrophysiological correlate of perceptual learning, memory, and training (Ross and Tremblay, 2009; Tremblay et al., 2001). It seems that P2 amplitude modulations are associated with cortical changes induced by repeated stimulus exposure rather than the learning outcome itself (Tremblay et al. 2014). Thus, our result of a smaller P2 response for the CI ear in SSD CI users can be explained by the fact that the NH ear—when compared to the CI ear—is more experienced and is more exposed to auditory stimuli as it is the dominant communication channel in these individuals. Regarding the *second mechanism*, previous studies with NH listeners have suggested that both the N1 and the P2 ERPs are susceptible to attention (Crowley and Colrain, 2004). An enhanced ERP amplitude at P2 latency can be explained by the attentional shift towards auditory stimuli (Picton and Hillyard, 1974), and seems to be associated with stimulus categorisation (García-Larrea et al. 1992). Following these previous studies, we interpret the larger P2 amplitude for

“vocoded” stimuli in the NH control group as reflecting an enhanced allocation of attentional resources to process the degraded and unfamiliar stimuli. It seems that in this difficult listening condition, the NH listeners’ speech processing is not automatic but explicit and therefore needs the additional recruitment of cognitive resources to reconstruct the limited speech signal (Rönnberg et al. 2013; Zekveld et al. 2010). This interpretation is supported by the finding that the NH control group reported an enhanced subjective listening effort in the “vocoded” than the “original” sound condition.

Question 3: Is there a Side-of-Implantation Effect on Speech Processing via the CI Ear in SSD CI Users?

It is currently not well understood, how the side of implantation affects the rehabilitation success in adult postlingually deafened SSD CI users. Recent results have pointed to a right-ear advantage for speech recognition ability in SSD CI users, independent from their pure tone thresholds (Wettstein and Probst, 2018). The authors have argued that this right-ear advantage in SSD CI users is mostly driven by the left-hemisphere dominance for speech processing. However, the current study could not replicate these previous findings, given that our behavioural results revealed no side-of-implantation effect on syllable processing. One may speculate that this lack of replication can be explained by the fact that the current study focused on syllables, whereas Wettstein and Probst (2018) presented four-syllabic numbers and monosyllabic words. The use of different stimulus types in the two studies obviously limits the comparability between the results. However, in addition to the EEG paradigm, our SSD CI users were also examined with standard clinical speech tests (Table 1). The results revealed no significant differences between left-ear and right-ear implanted SSD CI users regarding the pure-tone thresholds, the word recognition ability (assessed by the Freiburg monosyllabic word test) and the speech intelligibility (assessed by the Oldenburg sentence test). Thus, we speculate that the lack of a replication of a behavioural side-of-implantation effect may be related to the small sample size (used in the present study) in combination with the high variability in behavioural results observed in SSD CI users.

Our data revealed a significant hemispheric difference for the left-implanted participants, but not for the right-implanted participants, both for stimulation via the CI ear and the NH ear (Fig. 3a and Fig. 4a). For the stimulation via the *CI ear* (Fig. 3a), the left-implanted CI users showed a significantly enhanced activation in the right as compared to the left auditory cortex—referred to as contralateral dominance effect. For the stimulation via the *NH ear*, the left-implanted CI users (with a NH ear on the right side) showed a contralateral dominance effect as well, as indicated by a significantly enhanced activation in the left as compared to

the right auditory cortex (Fig. 4a; see section “Question 4: Is there a side-of-stimulation effect on speech processing via the NH ear in SSD CI users and in NH listeners?” for a discussion of the side-of-implantation effect on the NH ear).

In contrast to children with SSD, who develop a normal lateralization to the contralateral auditory cortex when implanted at young age (Lee et al. 2020; Polonenko et al. 2017), not much is known about functional changes in the auditory cortex of postlingually deafened adult SSD CI patients. To discuss the observed differences in auditory-cortex asymmetry between our left- and right-ear implanted SSD CI users, it is important to keep two aspects in mind. First, the contralateral dominance of the auditory cortex has been described for monaural stimulation (Hine and Debener, 2007). In the human auditory system, the pathway from each ear to the contralateral cortical hemisphere consists of more nerve fibres than the pathway to the ipsilateral hemisphere (Rosenzweig, 1951). Therefore, monaural stimulation evokes stronger responses in the contralateral than in the ipsilateral hemisphere (Jäncke et al. 2002). Second, there is a relative hemispheric specialisation of the left and right auditory cortex for the processing of basic acoustic properties (Lazard et al. 2012a). Prior studies have suggested that the left auditory cortex is more proficient in the processing of fast temporal acoustic cues, which are largely contained in speech stimuli (Boemio et al. 2005; Poeppel, 2003). Conversely, the right auditory cortex seems to preferentially process slowly modulated signals and spectral aspects of sounds, which are largely contained in musical stimuli (Liegeois-Chauvel, 1999; Poeppel, 2003; Zatorre et al. 2002). Thus, the hemispheric differences during speech and music processing can be attributed to the *relative* specialisation of the two hemispheres for basic acoustic stimulus properties, in particular fast temporal versus slow spectrotemporal acoustic cues. It has been assumed that the auditory cortex’ preference for basic stimulus properties drives higher-order organisation for speech and music perception (Lazard et al. 2012a).

Our observation of a contralateral dominance effect for the left-ear implanted SSD CI users during speech processing seems to contradict previous results of left-hemisphere dominance for speech processing in NH listeners. However, a strong activation in the right auditory cortex in these patients can be explained by the combination of two factors. First, the left ear shows stronger projections to the contralateral than ipsilateral auditory cortex. Therefore, monaural stimulation of the left ear resulted in an enhanced activation in the right than left auditory cortex. Second, the CI processing remarkably reduces the quality of the speech sounds and affects the spectrotemporal properties of the presented syllables. Given the relative specialisation of the two hemispheres for basic acoustic stimulus properties, the CI-related stimulus degradation may have resulted in a relatively stronger right-than-left-auditory cortex activation. Indeed, some of our SSD CI

patients reported that the speech stimuli were perceived as more noise-like and less speech-like when presented via the CI than the NH ear.

Our results showed a contralateral dominance effect specifically for the left-ear but not for the right-ear implanted SSD CI users. This contrasts with the results of Sandmann et al. (2009), who found a stronger contralateral dominance effect for right- than left-ear stimulated CI users. This discrepancy in results can be due to different reasons. First, in the current study we used syllables, while Sandmann et al. (2009) used dyadic tones, i.e., musical sounds. Speech and musical stimuli are characterised by different acoustic properties. Given the relative specialisation of the left and right auditory cortex for basic stimulus properties (Poeppel, 2003; Zatorre and Belin, 2001), the discrepancies between previous and current results with regards to the pattern of cortical asymmetry could be explained by distinct stimulus properties, resulting in a different recruitment of the left and right auditory cortex. Another reason for discrepant findings between previous and current results is that the current study examined SSD CI users, whereas Sandmann et al. (2009) tested CI users with bilateral hearing loss. In contrast to the current study, the hearing ability of the second ear was reduced, and it was *not* matched between the left- and right-ear stimulated CI users. Given that auditory deprivation reduces the metabolism in the contralateral auditory cortex (Speck et al. 2020) and can induce cortical reorganization in the auditory cortex (Stropahl et al. 2017), the different pattern of hemispheric asymmetry in the auditory cortex might have arisen due to the confounding effect of the hearing loss in the second ear.

It may be argued that handedness is a factor confounding our results regarding functional hemispheric asymmetry, since the probability of a reversed lateralisation for language processing seems to be enhanced in left handers when compared to right handers (Hund-Georgiadis et al. 2002). However, previous studies using different methods have observed that the majority of right- and left-handed individuals show left-sided cerebral dominance for language processing, and only less than 10% of the left-handers show right-sided cerebral dominance for language processing (Khedr et al. 2002; Szaflarski et al. 2002). In addition, a supplementary analysis of our behavioural and ERP data revealed that the two left-handed CI-users lay within the normal range (as defined by mean ± 2 standard deviations) and therefore we conclude that the activity in the left and right auditory cortex was not confounded by the factor handedness.

Our results revealed a similar pattern of auditory-cortex asymmetry between the SSD CI users and the NH control group, when the latter group was presented with “vocoded” sounds (Fig. 3). Specifically, SSD CI users and NH listeners showed a contralateral dominance effect for the left-ear stimulation, with enhanced activation in the right as compared

to the left auditory cortex. By contrast, no hemispheric difference was found for SSD CI users and NH listeners when they were stimulated via the right ear. Given these similarities between CI users and NH listeners (the latter tested with “vocoded” stimuli) and the fact that noise-band vocoder simulations allow a good approximation to sound conditions in CI users (Friesen et al. 2001), we conclude that the different pattern of contralateral dominance between left- and right-ear implanted CI SSD users is mainly driven by the CI-related stimulus degradation. Nevertheless, it is important to note that our sample size is limited, and the current study only allows conclusions to be drawn regarding the processing of syllables. Therefore, future studies should examine the functional asymmetry in the auditory cortex in further SSD CI users and for different auditory stimuli, in particular speech and musical sounds.

Question 4: Is there a Side-of-Stimulation Effect on Speech Processing via the NH Ear in SSD CI Users and in NH Listeners?

A recent multicentre study has reported a significant difference in the hearing threshold between the NH ear of SSD patients and the NH ears of age-matched NH listeners (Arndt et al. 2019). This observation points to a poorer peripheral hearing capacity for the intact ear of SSD patients when compared to NH listeners. Importantly, these previous results suggest that SSD CI users show behavioural alterations not only in the CI ear but also in the NH ear. It seems that these alterations in the NH ear are *not* induced by cochlear implantation, given that the hearing threshold of the NH ear appears to be comparable at the times before and after implantation (Speck et al. 2021). Further, alterations in the NH ear of SSD CI users are not limited to the hearing threshold but can also show up in other auditory tests. For instance, Maslin et al. (2015) have reported that in SSD CI users the intact ear is better able to discriminate intensity differences, suggesting perceptual improvements as induced by cortical plasticity following unilateral deafness.

The present study did also reveal a significant difference between the NH ear of SSD CI users and the matched NH ear of NH listeners with regards to the hearing threshold, but not with regards to the behavioural performance and hit rates in the auditory oddball paradigm. Thus, our results can confirm alterations in the intact ear of SSD CI users, at least on the peripheral hearing capacity. But we could not confirm these alterations on the behavioural level in the oddball paradigm. Several reasons may account for the lack of comparability between the two results in the present study. In addition to the small sample size and the high variability in behavioural measures across participants, there are methodological discrepancies between the two measurements (Arndt et al. 2019; Maslin et al. 2015; Speck et al. 2021),

in particular in terms of the task (pure-tone audiometry/intensity difference limens vs. auditory oddball task) and the stimulus material (pure tones vs. syllables). It may be speculated, that alterations in the NH ear of SSD CI users are stimulus- and task-selective and may be revealed only under specific conditions.

As far as we are aware, the present study is the first to compare the NH ears of SSD CI users and NH listeners in the context of an auditory oddball paradigm. The ERP source analysis revealed that *stimulation of the right NH ear* of (left-implanted) CI users induced an enhanced activation in the left as compared to the right auditory cortex—referred to as contralateral dominance effect (Fig. 4a top right). By contrast, the NH listeners—when stimulated on the right ear—showed no hemispheric difference in auditory-cortex activation, although on the descriptive level, the activation in the left hemisphere was enhanced (Fig. 4a bottom right). Regarding the *stimulation of the left NH ear*, both the (right-implanted) SSD CI users and the NH listeners showed no activation differences between the right and left auditory cortex (Fig. 4a top left and bottom left). These results suggest that specifically the group of left-ear implanted SSD CI users shows cortical alterations for speech-sound processing when stimulated via the (right) NH ear (Fig. 4a top right). Interestingly, the same group also revealed an enhanced contralateral dominance effect when stimulated via the CI ear (Fig. 3a top left; see also section “Question 3: Is there a side-of-implantation effect on speech processing via the CI ear in SSD CI users?” for a discussion of this effect).

The enhanced hemispheric asymmetry for the stimulation of the *right NH ear* in (left-implanted) SSD CI users may be explained by alterations in the left auditory cortex for the processing of rapidly changing stimulus properties contained in speech stimuli (Boemio et al. 2005; Poeppel, 2003). It can be speculated that these improvements are induced by temporary unilateral deafness and/or electrical afferentation with a CI. These improvements may reflect an optimised left-hemispheric speech processing, which is particularly important in difficult listening conditions with reduced or degraded auditory input (processing via the CI ear), but which also affects the processing of the normal acoustic input (processing via the NH ear). Alternatively, but not mutually exclusive, previous animal studies have shown that unilateral deafness results in an enhanced ipsilateral activation, which is due to an increased number and/or enhanced excitability of neurons that are responsive to the intact ear (McAlpine et al. 1997; Mossop et al. 2000). Regarding the current results, the *reduced* hemispheric asymmetry for the stimulation of the *left NH ear* in (right-implanted) SSD CI users may be explained by the fact that SSD patients show enhanced afferent input from the (left) intact ear to the (left) ipsilateral auditory cortex (Maslin et al. 2015). Thus, the speech processing via the *left NH*

ear in right-implanted SSD CI users might evoke a strong activation in the *left* auditory cortex, which counteracts the contralateral dominance effect for left-ear stimulation. Taken together, our results provide first evidence of a side-of-implantation effect on functional auditory-cortex asymmetry in adult postlingually deafened SSD CI users, which is not limited to the CI ear, but which is also shown for the NH ear. However, in order to gain a better understanding of the cortical changes in the intact ear of SSD CI users, future studies are required to examine whether a similar pattern of results on hemispheric asymmetries can be observed with other types of stimuli, for instance musical sounds and more complex speech stimuli.

Limitations

One limitation of this study is the relatively small group size, which results in a low statistical power and therefore makes it difficult to draw conclusions that are transferable to the generality. However, we believe that our results point out important issues in single-sided deaf CI-users that are worth further research to support our findings of asymmetry in the auditory cortex depending on the side of implantation. Furthermore, we did not find correlations between our results of the EEG paradigm and the clinical speech tests reported in this manuscript. One important reason for this lack seems to be the fact that the syllable-discrimination task in the EEG paradigm (discrimination of phonetic contrasts) and the clinical speech tests (monosyllabic word test, Oldenburg Sentence Test (OLSA)) examine speech competencies on different linguistic levels. In particular, the comparison between the standard stimulus (/ki/) and the deviant 2 stimulus (/ti/) only relies on the consonant-contrast, which is very hard to discriminate for some CI-user, particularly in situations without any given context. A second reason could be our exploratory median split procedure. We used this procedure to divide our sample in equal group sizes, but it was not possible to get a division in *clearly poor* and *clearly high* CI performers. Further research should use more diverse and complex stimuli to differentiate the groups (proficient vs. non-proficient) on a more solid basis.

Conclusions

The present study used an auditory oddball task to examine cortical speech processing of the CI ear and the NH ear of SSD CI users. Given that non-proficient and proficient SSD CI users could be distinguished based on the N1 and P3b amplitude, we conclude that the time-efficient two-deviant oddball paradigm can be used to assess speech discrimination proficiency in SSD CI users. Further, our results suggest

that the observed differences in cortical speech processing between the CI ear and the NH ear in SSD CI users are (at least partially) caused by limitations in the cortical adaptation to the implant signal, which leads to difficulties in the sensory and cognitive speech processing for the CI ear. Finally, we found a side-of-implantation effect on auditory-cortex asymmetry for both the CI ear and the NH ear. We suppose that these side-of-implantation effects originate from CI-related degradation of the stimuli and cortical reorganisation as induced by temporary unilateral deafness and/or degraded sensory input after implantation.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10548-022-00902-3>.

Funding Open Access funding enabled and organized by Projekt DEAL. This work was supported by the German Research Foundation (DFG) [project number 415896102 and 416867313] and by the Koeln Fortune Program/Faculty of Medicine, University of Cologne [378/2020] to PS.

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose. The authors have no competing interests to declare that are relevant to the content of this article.

Ethical Approval The study was approved by the Ethics Committee of Cologne University's Faculty of Medicine (application number 13–057) and was conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki, 2013).

Consent to Participate Informed consent was obtained from all individual participants included in the study.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arndt S, Aschendorff A, Laszig R, Beck R, Schild C, Kroeger S et al (2011a) Comparison of pseudobinaural hearing to real binaural hearing rehabilitation after cochlear implantation in patients with unilateral deafness and tinnitus. *Otol Neurotol* 32(1):39–47

- Arndt S, Laszig R, Aschendorff A, Beck R, Schild C, Hassepass F et al (2011b) Einseitige taubheit und cochlear-implant-versorgung. *HNO* 59(5):437–446
- Arndt S, Wesarg T, Stelzig Y, Jacob R, Illg A, Lesinski-Schiedat A, Ketterer MC, Aschendorff A, Speck I (2019) Einfluss einseitiger taubheit auf das hörmögen des besseren ohrs. *HNO* 67:739–749. <https://doi.org/10.1007/s00106-019-00730-1>
- Beynon AJ, Snik AFM, Stegeman DF, van den Broek P (2005) Discrimination of speech sound contrasts determined with behavioral tests and event-related potentials in cochlear implant recipients. *J Am Acad Audiol* 16:042–053. <https://doi.org/10.3766/jaaa.16.1.5>
- Billings CJ, Bennett KO, Molis MR, Leek MR (2011) Cortical encoding of signals in noise: effects of stimulus type and recording paradigm. *Ear Hear* 32:53–60. <https://doi.org/10.1097/AUD.0b013e3181ec5c46>
- Blamey P, Artieres F, Başkent D, Bergeron F, Beynon A, Burke E et al (2013) Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. *Audiol Neurotol* 18(1):36–47
- Boemio A, Fromm S, Braun A, Poeppel D (2005) Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci* 8:389–395. <https://doi.org/10.1038/nn1409>
- Buechner A, Brendel M, Lesinski-Schiedat A, Wenzel G, Frohne-Buechner C, Jaeger B, Lenarz T (2010) Cochlear implantation in unilateral deaf subjects associated with ipsilateral tinnitus. *Otol Neurotol* 31(9):1381–1385
- Bönitz H, Kopp B, Büchner A, Lunner T, Lyxell B, Finke M (2018) Event-related neuronal responses to acoustic novelty in single-sided deaf cochlear implant users: initial findings. *Clin Neurophysiol* 129:133–142. <https://doi.org/10.1016/j.clinph.2017.10.025>
- Crowley KE, Colrain IM (2004) A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clin Neurophysiol* 115:732–744. <https://doi.org/10.1016/j.clinph.2003.11.021>
- Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, Halgren E (2000) Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26(1):55–67
- Debener S, Hine J, Bleeck S, Eyles J (2007) Source localization of auditory evoked potentials after cochlear implantation. *Psychophysiology*. <https://doi.org/10.1111/j.1469-8986.2007.00610.x>
- Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134:9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>
- Destrieux C, Fischl B, Dale A, Halgren E (2010) Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* 53:1–15. <https://doi.org/10.1016/j.neuroimage.2010.06.010>
- Drennan WR (2008) Music perception in cochlear implant users and its relationship with psychophysical capabilities. *JRRD* 45:779–790. <https://doi.org/10.1682/JRRD.2007.08.0118>
- Finke M, Sandmann P, Bönitz H, Kral A, Büchner A (2016) Consequences of stimulus type on higher-order processing in single-sided deaf cochlear implant users. *Audiol Neurotol* 21:305–315. <https://doi.org/10.1159/000452123>
- Finke M, Sandmann P, Kopp B, Lenarz T, Büchner A (2015) Auditory distraction transmitted by a cochlear implant alters allocation of attentional resources. *Front Neurosci*. <https://doi.org/10.3389/fnins.2015.00068>
- Friesen LM, Shannon RV, Baskent D, Wang X (2001) Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am* 110:1150–1163. <https://doi.org/10.1121/1.1381538>
- García-Larrea L, Lukaszewicz A-C, Mauguière F (1992) Revisiting the oddball paradigm. non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia* 30:723–741. [https://doi.org/10.1016/0028-3932\(92\)90042-K](https://doi.org/10.1016/0028-3932(92)90042-K)
- Gaudrain E (2016) Vocoder, v1.0. Online code at <https://github.com/egaudrain/vocoder>. <https://doi.org/10.5281/zenodo.48120>
- Gaudrain E, Başkent D (2015) Factors limiting vocal-tract length discrimination in cochlear implant simulations. *J Acoust Soc Am* 137:1298–1308. <https://doi.org/10.1121/1.4908235>
- Giraud A-L, Price CJ, Graham JM, Truy E, Frackowiak RSJ (2001) Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron* 30:657–664. [https://doi.org/10.1016/S0896-6273\(01\)00318-X](https://doi.org/10.1016/S0896-6273(01)00318-X)
- Gordon KA, Jiwani S, Papsin BC (2013) Benefits and detriments of unilateral cochlear implant use on bilateral auditory development in children who are deaf. *Front Psychol*. <https://doi.org/10.3389/fpsyg.2013.00719>
- Gramfort A, Papadopoulos T, Olivi E, Clerc M (2010) Open-MEEG: opensource software for quasistatic bioelectromagnetics. *BioMed Eng OnLine* 9:45. <https://doi.org/10.1186/1475-925X-9-45>
- Green KJM, Julyan PJ, Hastings DL, Ramsden RT (2005) Auditory cortical activation and speech perception in cochlear implant users: effects of implant experience and duration of deafness. *Hear Res* 205:184–192. <https://doi.org/10.1016/j.heares.2005.03.016>
- Greenwood DD (1990) A cochlear frequency-position function for several species—29 years later. *J Acoust Soc Am* 87(6):2592–2605
- Groenen PAP, Beynon AJ, Snik AFM, van den Broek P (2001) Speech-evoked cortical potentials recognition in cochlear implant users and speech. *Scand Audiol* 30:31–40. <https://doi.org/10.1080/010503901750069554>
- Hahlbrock KH (1953) Über sprachaudiometrie und neue wörterteste. *Archiv Für Ohren-, Nasen-Und Kehlkopfheilkunde* 162(5):394–431
- Hahlbrock KH (1970) Sprachaudiometrie: Grundlagen und praktische Anwendung einer Sprachaudiometrie für das deutsche Sprachgebiet; 157 Abbildungen in 305 Einzeldarstellungen 9 Tabellen. Thieme.
- Hansen P, Kringelbach M, Salmelin R (2010) MEG: an introduction to methods. Oxford University Press, Oxford
- Haumann S, Lenarz T, Büchner A (2010) Speech perception with cochlear implants as measured using a roving-level adaptive test method. *ORL* 72:312–318. <https://doi.org/10.1159/000318872>
- Henkin Y, Swead RT, Roth DA-E, Kishon-Rabin L, Shapira Y, Migirov L, Hildesheimer M, Kaplan-Neeman R (2014) Evidence for a right cochlear implant advantage in simultaneous bilateral cochlear implantation: right ear advantage and bilateral CIs. *Laryngoscope* 124:1937–1941. <https://doi.org/10.1002/lary.24635>
- Henkin Y, Tetin-Schneider S, Hildesheimer M, Kishon-Rabin L (2009) Cortical neural activity underlying speech perception in postlingual adult cochlear implant recipients. *Audiol Neurotol* 14:39–53. <https://doi.org/10.1159/000153434>
- Hey M, Hocke T, Hedderich J, Müller-Deile J (2014) Investigation of a matrix sentence test in noise: reproducibility and discrimination function in cochlear implant patients. *Int J Audiol* 53:895–902. <https://doi.org/10.3109/14992027.2014.938368>
- Hey M, Hocke T, Mauger S, Müller-Deile J (2016) A clinical assessment of cochlear implant recipient performance: implications for individualized map settings in specific environments. *Eur Arch Otorhinolaryngol* 273:4011–4020. <https://doi.org/10.1007/s00405-016-4130-2>
- Hine J, Debener S (2007) Late auditory evoked potentials asymmetry revisited. *Clin Neurophysiol* 118:1274–1285. <https://doi.org/10.1016/j.clinph.2007.03.012>

- Hochmair-Desoyer I, Schulz E, Moser L, Schmidt M (1997) The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users. *Am J Otol* 18(6 Suppl):S83–S83
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scand J Statistics* 6:6570
- Hund-Georgiadis M, Lex U, Friederici AD, von Cramon DY (2002) Non-invasive regime for language lateralization in right and left-handers by means of functional MRI and dichotic listening. *Exp Brain Res* 145:166–176. <https://doi.org/10.1007/s00221-002-1090-0>
- Jung T-P, Makeig S, Humphries C, Lee T-W, McKeown MJ, Iragui V, Sejnowski TJ (2000) Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37:163–178. <https://doi.org/10.1111/1469-8986.3720163>
- Jäncke L, Wüstenberg T, Schulze K, Heinze HJ (2002) Asymmetric hemodynamic responses of the human auditory cortex to monaural and binaural stimulation. *Hear Res* 170:166–178. [https://doi.org/10.1016/S0378-5955\(02\)00488-4](https://doi.org/10.1016/S0378-5955(02)00488-4)
- Kelly AS, Purdy SC, Thorne PR (2005) Electrophysiological and speech perception measures of auditory processing in experienced adult cochlear implant users. *Clin Neurophysiol* 116:1235–1246. <https://doi.org/10.1016/j.clinph.2005.02.011>
- Kent RD (1997) The speech sciences. Singular Publ. Group, San Diego
- Kessler M, Schierholz I, Mamach M, Wilke F, Hahne A, Büchner A, Geworski L, Bengel FM, Sandmann P, Berding G (2020) Combined brain-perfusion SPECT and EEG measurements suggest distinct strategies for speech comprehension in CI users with higher and lower performance. *Front Neurosci* 14:787. <https://doi.org/10.3389/fnins.2020.00787>
- Khedr E, Hamed E, Said A, Basahi J (2002) Handedness and language cerebral lateralization. *Eur J Appl Physiol* 87:469–473. <https://doi.org/10.1007/s00421-002-0652-y>
- Kitterick PT, Lucas L, Smith SN (2015) Improving health-related quality of life in single-sided deafness: a systematic review and meta-analysis. *Audiol Neurotol* 20:79–86. <https://doi.org/10.1159/000380753>
- Kral A (2007) Unimodal and cross-modal plasticity in the ‘deaf’ auditory cortex. *Int J Audiol* 46:479–493. <https://doi.org/10.1080/14992020701383027>
- Kral A, Hubka P, Heid S, Tillein J (2013) Single-sided deafness leads to unilateral aural preference within an early sensitive period. *Brain* 136(1):180–193
- Lazard DS, Collette J-L, Perrot X (2012a) Speech processing: from peripheral to hemispheric asymmetry of the auditory system: auditory asymmetry-didactic review. *Laryngoscope* 122:167–173. <https://doi.org/10.1002/lary.22370>
- Lazard DS, Giraud A-L, Gnansia D, Meyer B, Sterkers O (2012b) Understanding the deafened brain: implications for cochlear implant rehabilitation. *Eur Ann Otorhinolaryngol Head Neck Dis* 129:98–103. <https://doi.org/10.1016/j.anorl.2011.06.001>
- Lazard DS, Vincent C, Venail F, Van de Heyning P, Truy E, Sterkers O, Skarzynski PH, Skarzynski H, Schauwers K, O’Leary S, Mawman D, Maat B, Kleine-Punte A, Huber AM, Green K, Govaerts PJ, Frayssse B, Dowell R, Dillier N, Burke E, Beynon A, Bergeron F, Başkent D, Artières F, Blamey PJ (2012c) Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model over time. *PLoS ONE* 7:e48739. <https://doi.org/10.1371/journal.pone.0048739>
- Lee DS, Lee JS, Oh SH, Kim SK, Kim JW, Chung JK et al (2001) Cross-modal plasticity and cochlear implants. *Nature* 409(6817):149–150
- Lee H-J, Smieja D, Polonenko MJ, Cushing SL, Papsin BC, Gordon KA (2020) Consistent and chronic cochlear implant use partially reverses cortical effects of single sided deafness in children. *Sci Rep* 10:21526. <https://doi.org/10.1038/s41598-020-78371-6>
- Legrís E, Galvin J, Roux S, Gomot M, Aoustin J-M, Marx M, He S, Bakhos D (2018) Cortical reorganization after cochlear implantation for adults with single-sided deafness. *PLoS ONE* 13:e0204402. <https://doi.org/10.1371/journal.pone.0204402>
- Lenarz M, Sönmez H, Joseph G, Büchner A, Lenarz T (2012) Cochlear implant performance in geriatric patients. *Laryngoscope* 122:1361–1365. <https://doi.org/10.1002/lary.23232>
- Liegeois-Chauvel C (1999) Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb Cortex* 9:484–496. <https://doi.org/10.1093/cercor/9.5.484>
- Lin F-H, Witzel T, Ahlfors SP, Stufflebeam SM, Belliveau JW, Hämäläinen MS (2006) Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *Neuroimage* 31:160–171. <https://doi.org/10.1016/j.neuroimage.2005.11.054>
- Luck SJ (2014) An introduction to the event-related potential technique. MIT press, Cambridge
- Maslin MRD, Taylor M, Plack CJ, Munro KJ (2015) Enhanced intensity discrimination in the intact ear of adults with unilateral deafness. *J Acoust Soc Am* 137:EL408–EL414. <https://doi.org/10.1121/1.4914945>
- McAlpine D, Martin RL, Mossop JE, Moore DR (1997) Response properties of neurons in the inferior colliculus of the monaurally deafened ferret to acoustic stimulation of the intact ear. *J Neurophysiol* 78:767–779. <https://doi.org/10.1152/jn.1997.78.2.767>
- Micco AG, Kraus N, Koch DB, McGee TJ, Carrell TD, Sharma A et al (1995) Speech-evoked cognitive P300 potentials in cochlear implant recipients. *Am J Otol* 16(4):514–520
- Mosnier I, Bebear JP, Marx M, Frayssse B, Truy E, Lina-Granade G et al (2014) Predictive factors of cochlear implant outcomes in the elderly. *Audiol Neurotol* 19(Suppl. 1):15–20
- Mossop JE, Wilson MJ, Caspary DM, Moore DR (2000) Down-regulation of inhibition following unilateral deafening. *Hear Res* 147:183–187. [https://doi.org/10.1016/S0378-5955\(00\)00054-X](https://doi.org/10.1016/S0378-5955(00)00054-X)
- Nadol JB Jr, Young YS, Glynn RJ (1989) Survival of spiral ganglion cells in profound sensorineural hearing loss: implications for cochlear implantation. *Ann Otol, Rhinol Laryngol* 98(6):411–416
- Näätänen R, Picton T (1987) The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24(4):375–425
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9(1):97–113
- O’Connell RG, Dockree PM, Kelly SP (2012) A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nat Neurosci* 15:1729–1735. <https://doi.org/10.1038/nn.3248>
- Pakarinen S, Lovio R, Huottilainen M, Alku P, Näätänen R, Kujala T (2009) Fast multi-feature paradigm for recording several mismatch negativities (MMNs) to phonetic and acoustic changes in speech sounds. *Biol Psychol* 82(3):219–226
- Park LR, Preston E, Noxon AS, Dillon MT (2021) Comparison of test methods to assess the implanted ear alone for pediatric cochlear implant recipients with single-sided deafness. *Cochlear Implants Int* 22:283–290. <https://doi.org/10.1080/14670100.2021.1903715>
- Picton TW (2010) Human auditory evoked potentials. Plural Publishing.
- Picton TW, Hillyard SA (1974) Human auditory evoked potentials. II: effects of attention. *Electroencephalogr Clin Neurophysiol* 36:191–200. [https://doi.org/10.1016/0013-4694\(74\)90156-4](https://doi.org/10.1016/0013-4694(74)90156-4)
- Poeppel D (2003) The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time.’ *Speech Commun* 41:245–255. [https://doi.org/10.1016/S0167-6369\(02\)00107-3](https://doi.org/10.1016/S0167-6369(02)00107-3)

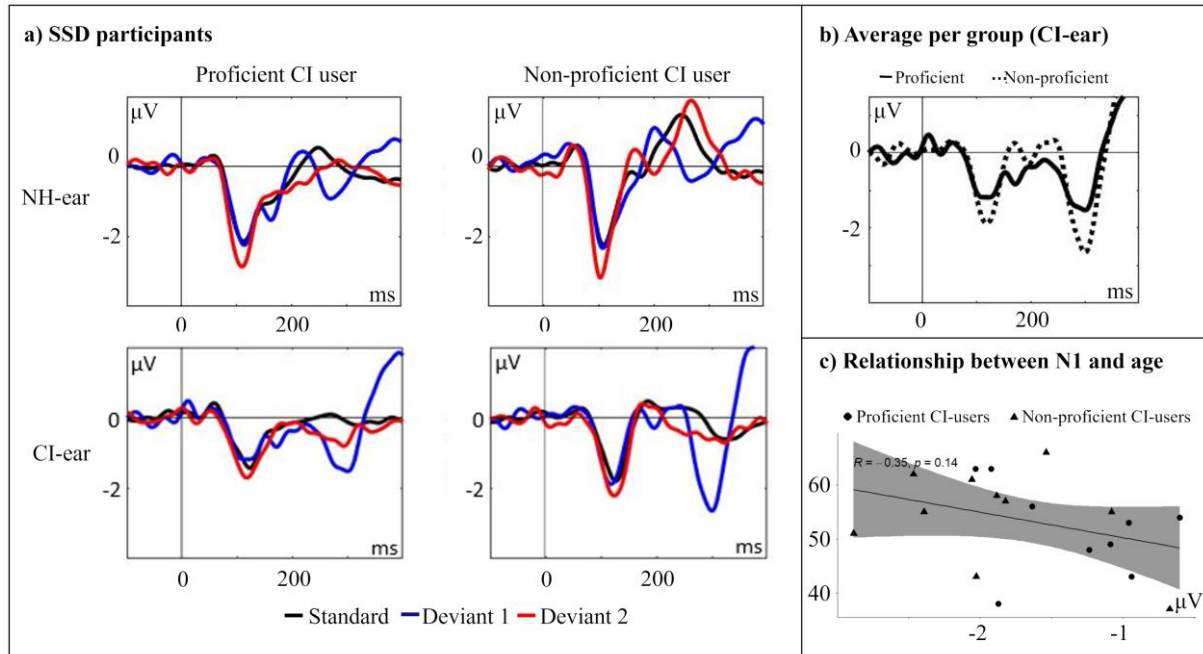
- Polich J (2007) Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 118:2128–2148. <https://doi.org/10.1016/j.clinph.2007.04.019>
- Polich J, Comerchero MD (2003) P3a from visual stimuli: typicality, task, and topography. *Brain Topogr* 15(3):141–152
- Polonenko MJ, Gordon KA, Cushing SL, Papsin BC (2017) Cortical organization restored by cochlear implantation in young children with single sided deafness. *Sci Rep* 7:16900. <https://doi.org/10.1038/s41598-017-17129-z>
- Roberts DS, Lin HW, Herrmann BS, Lee DJ (2013) Differential cochlear implant outcomes in older adults: cochlear implant outcomes in older adults. *Laryngoscope* 123:1952–1956. <https://doi.org/10.1002/lary.23676>
- Rosenzweig MR (1951) Representations of the two ears at the auditory cortex. *Am J Physiology-Legacy Content* 167:147–158. <https://doi.org/10.1152/ajplegacy.1951.167.1.147>
- Ross B, Tremblay K (2009) Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hear Res* 248:48–59. <https://doi.org/10.1016/j.heares.2008.11.012>
- Rönnberg J, Lunner T, Zekveld A, Sörqvist P, Danielsson H, Lyxell B, Dahlström Ö, Signoret C, Stenfelt S, Pichora-Fuller MK, Rudner M (2013) The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front Syst Neurosci*. <https://doi.org/10.3389/fnsys.2013.00031>
- Sandmann P, Dillier N, Eichele T, Meyer M, Kegel A, Pascual-Marqui RD, Marcar VL, Jäncke L, Debener S (2012) Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain* 135:555–568. <https://doi.org/10.1093/brain/awr329>
- Sandmann P, Eichele T, Buechler M, Debener S, Jäncke L, Dillier N, Hugdahl K, Meyer M (2009) Evaluation of evoked potentials to dyadic tones after cochlear implantation. *Brain* 132:1967–1979. <https://doi.org/10.1093/brain/awp034>
- Sandmann P, Kegel A, Eichele T, Dillier N, Lai W, Bendixen A, Debener S, Jancke L, Meyer M (2010) Neurophysiological evidence of impaired musical sound perception in cochlear-implant users. *Clin Neurophysiol* 121:2070–2082. <https://doi.org/10.1016/j.clinph.2010.04.032>
- Sandmann P, Plotz K, Hauthal N, de Vos M, Schönfeld R, Debener S (2015) Rapid bilateral improvement in auditory cortex activity in postlingually deafened adults following cochlear implantation. *Clin Neurophysiol* 126:594–607. <https://doi.org/10.1016/j.clinph.2014.06.029>
- Sasaki T, Yamamoto K, Iwaki T, Kubo T (2009) Assessing binaural/bimodal advantages using auditory event-related potentials in subjects with cochlear implants. *Auris Nasus Larynx* 36:541–546. <https://doi.org/10.1016/j.anl.2008.12.001>
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. *Science* 270(5234):303–304
- Soshi T, Hisanaga S, Kodama N, Kanekama Y, Samejima Y, Yumoto E, Sekiyama K (2014) Event-related potentials for better speech perception in noise by cochlear implant users. *Hear Res* 316:110–121. <https://doi.org/10.1016/j.heares.2014.08.001>
- Speck I, Arndt S, Thurow J, Blazhenets G, Aschendorff A, Meyer PT, Frings L (2020) ¹⁸F-FDG PET imaging of the inferior colliculus in asymmetric hearing loss. *J Nucl Med* 61:418–422. <https://doi.org/10.2967/jnumed.119.231407>
- Speck I, Challier P, Wesarg T, Jakob TF, Aschendorff A, Hassepass F, Arndt S (2021) Is the cochlear implant a successful long-term solution for single-sided deaf and asymmetric hearing-impaired patients? *Eur Arch Otorhinolaryngol* 278:3257–3265. <https://doi.org/10.1007/s00405-020-06411-y>
- Stenroos M, Hunold A, Hauelsen J (2014) Comparison of three-shell and simplified volume conductor models in magnetoencephalography. *Neuroimage* 94:337–348. <https://doi.org/10.1016/j.neuroimage.2014.01.006>
- Stropahl M, Bauer A-KR, Debener S, Bleichner MG (2018) Source-modeling auditory processes of EEG data using EEGLAB and brainstorm. *Front Neurosci* 12:309. <https://doi.org/10.3389/fnins.2018.00309>
- Stropahl M, Chen L-C, Debener S (2017) Cortical reorganization in postlingually deaf cochlear implant users: intra-modal and cross-modal considerations. *Hear Res* 343:128–137. <https://doi.org/10.1016/j.heares.2016.07.005>
- Szaflarski JP, Binder JR, Possing ET, McKiernan KA, Ward BD, Hammeke TA (2002) Language lateralization in left-handed and ambidextrous people: fMRI data. *Neurology* 59(2):238–244
- Tadel F, Baillet S, Mosher JC, Pantazis D, Leahy RM (2011) Brainstorm: A user-friendly application for MEG/EEG analysis. *Comput Intell Neurosci* 2011:1–13. <https://doi.org/10.1155/2011/879716>
- Tremblay K, Kraus N, McGee T, Ponton C, Otis B (2001) Central auditory plasticity: changes in the N1–P2 complex after speech-sound training. *Ear Hear* 22:79–90. <https://doi.org/10.1097/00003446-200104000-00001>
- Tremblay KL, Ross B, Inoue K, McClannahan K, Collet G (2014) Is the auditory evoked P2 response a biomarker of learning? *Front Syst Neurosci*. <https://doi.org/10.3389/fnsys.2014.00028>
- Turgeon C, Lazzouni L, Lepore F, Ellemberg D (2014) An objective auditory measure to assess speech recognition in adult cochlear implant users. *Clin Neurophysiol* 125:827–835. <https://doi.org/10.1016/j.clinph.2013.09.035>
- Távora-Vieira D, Wedekind A, Marino R, Purdy SC, Rajan GP (2018) Using aided cortical assessment as an objective tool to evaluate cochlear implant fitting in users with single-sided deafness. *PLoS ONE* 13:e0193081. <https://doi.org/10.1371/journal.pone.0193081>
- Verleger R (2020) Effects of relevance and response frequency on P3b amplitudes: review of findings and comparison of hypotheses about the process reflected by P3b. *Psychophysiology* 57(7):e13542
- Wagener K, Brand T, Kollmeier B (1999) Entwicklung und evaluation eines satztests für die deutsche sprache I-III: design, optimierung und evaluation des oldenburger satztests. *Zeitschrift Für Audiologie* 38(1–3):4–15
- Wedekind A, Távora-Vieira D, Nguyen AT, Marinovic W, Rajan GP (2021) Cochlear implants in single-sided deaf recipients: near normal higher-order processing. *Clin Neurophysiol* 132(2):449–456
- Welge-Lüssen A, Hauser R, Erdmann J, Schwob C, Probst R (1997) Sprachaudiometrie mit logatomen. *Laryngorhinootologie* 76(02):57–64
- Wettstein VG, Probst R (2018) Right ear advantage of speech audiometry in single-sided deafness. *Otol Neurotol* 39:417–421. <https://doi.org/10.1097/MAO.0000000000001756>
- Widmann A, Schröger E, Maess B (2015) Digital filter design for electrophysiological data—a practical approach. *J Neurosci Methods* 250:34–46. <https://doi.org/10.1016/j.jneumeth.2014.08.002>
- Wilson BS, Dorman MF (2008) Cochlear implants: a remarkable past and a brilliant future. *Hear Res* 242:3–21. <https://doi.org/10.1016/j.heares.2008.06.005>
- Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. *Cereb Cortex* 11(10):946–953
- Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. *Trends Cogn Sci* 6:37–46. [https://doi.org/10.1016/S1364-6613\(00\)01816-7](https://doi.org/10.1016/S1364-6613(00)01816-7)
- Zekveld AA, Kramer SE, Festen JM (2010) Pupil Response as an indication of effortful listening: the influence of sentence intelligibility. *Ear Hear* 31:480–490. <https://doi.org/10.1097/AUD.0b013e3181d4f251>
- Zeng FG, Popper AN, Fay RR (2011) Auditory prostheses: new horizons, vol 39. Springer, Cham

Zhou X, Sobczak GS, McKay CM, Litovsky RY (2022) Effects of degraded speech processing and binaural unmasking investigated using functional near-infrared spectroscopy (fNIRS). *PLoS ONE* 17(4):e0267588

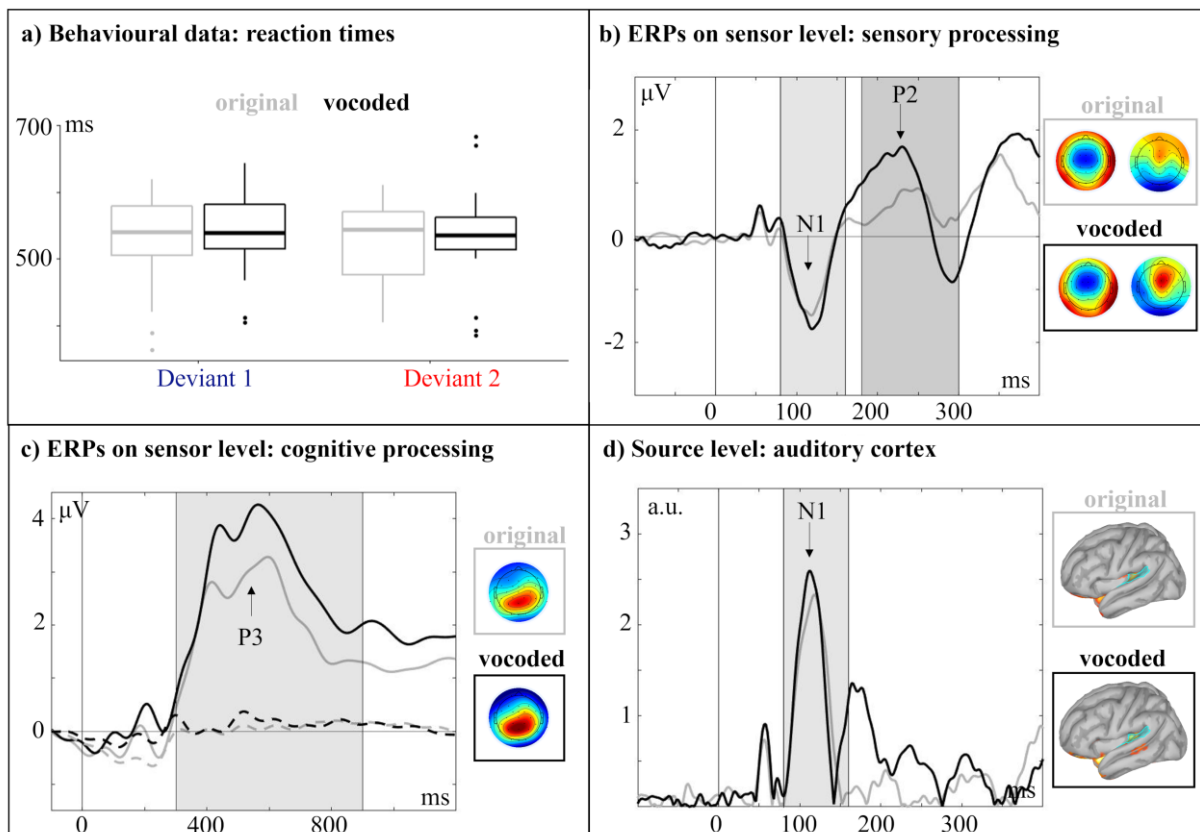
Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary Material

Supplementary Figure 1



Supplementary Figure 2



"ELECTROPHYSIOLOGICAL DIFFERENCES AND SIMILARITIES
IN AUDIOVISUAL SPEECH PROCESSING IN COCHLEAR
IMPLANT USERS WITH UNILATERAL AND BILATERAL
HEARING LOSS"

This chapter is based on the following publication:

Layer, N., Weglage, A., Müller, V., Meister, H., Lang-Roth, R., Walger, M., Murray, M. M. & Sandmann, P. (2022). Electrophysiological differences and similarities in audiovisual speech processing in CI users with unilateral and bilateral hearing loss. *Current Research in Neurobiology*, 100059. DOI: 10.1016/j.crneur.2022.100059.

Author contributions:

Natalie Layer: Conceptualisation; literature review; creation, editing and calibration of stimuli; programming of the experiment; pilot measurements; participant recruitment; data acquisition; data analysis; interpretation and presentation of results to research group; discussion of findings and further procedures; writing initial manuscript draft; incorporating suggestions of co-authors; submission to journal; presentation of results at national and international conferences. **Anna Weglage:** Assistance in data acquisition; assistance in data analysis; group discussions; proof-reading of manuscript. **Verena Müller:** Advisory role (CI technician); initial assistance in participant enrolment; proofreading of manuscript. **Hartmut Meister:** Funding acquisition; conceptualisation; provision of software for stimuli generation; discussion; suggestions; proofreading of manuscript. **Ruth Lang-Roth:** Conceptualisation; proof-reading of manuscript. **Martin Walger:** Conceptualisation; proof-reading of manuscript. **Micah M. Murray:** Presentation of analytical method (topographic analysis); validation of analysis; discussions; suggestions; proofreading of manuscript. **Pascale Sandmann:** Supervisor of Natalie Layer; central role in project conceptualisation; funding acquisition; suggestions; discussions; assistance in writing manuscript and proofreading of manuscript.

SYNOPSIS

OBJECTIVES Given that a CI provides only restricted temporal and spectral information (Drennan and Rubinstein, 2008), CI users are thought to develop compensatory strategies to overcome these limitations. CI users with bilateral hearing loss (provided with either two CIs or one CI plus hearing aid) show enhanced audiovisual interac-

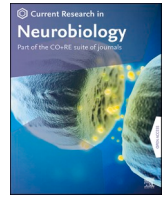
tions which are reflected in cross-modal recruitment of the visual cortex for purely auditory stimuli (Chen et al., 2016; Giraud et al., 2001c) and in better lip-reading abilities compared to NH controls (Rouger et al., 2007). Current ERP studies confirm enhanced multisensory interactions in these CI users for simple non-linguistic stimuli (Schierholz et al., 2015) and linguistic stimuli (Layer et al., 2022; Radecke et al., 2022). However, it is unknown whether these observations are valid for all CI user groups. Given that previous reports were based on CI users with bilateral hearing loss (here referred to as contralateral hearing device (CHD)-CI users), there is no knowledge about audiovisual interactions and lip-reading skills in CI users with unilateral hearing loss (SSD-CI users). To fill this gap, this ERP study extends the results from Layer et al. (2022) by comparing the timecourse of auditory and audiovisual speech processing, as well as the lip-reading abilities in CI-CHD users, CI-SSD users and NH controls.

METHODS In this study, we included postlingual CI-CHD and CI-SSD users and NH controls. CI-CHD users were fitted with either two CIs or a CI plus a hearing aid on the contralateral ear. The CI-SSD users were fitted with one CI and were NH on the contralateral ear. We recorded an EEG using 64 electrodes while the participants performed a speeded response task (the same as in Layer et al. (2022)). The syllables /ki/ and /ka/ were presented in auditory (A), visual (V) and audiovisual (AV) conditions. Each button on a computer mouse was assigned to one of the two syllables. The participants were asked to press a corresponding button as fast as possible when identifying a syllable. Furthermore, after each experimental block, we asked participants to rate how difficult it was to perform the task. Following the EEG experiment, the participants were required to perform a lip-reading test with monosyllabic words, similar to Layer et al. (2022).

ANALYSIS We evaluated both behavioural and ERP measures during the EEG experiment, comparable to Layer et al. (2022). The behavioural measures comprised hit rates and response times of the button presses. The ERP measures were analysed by using an additive model (Barth et al., 1995), which is frequently used to investigate audiovisual interactions (Stevenson et al., 2014). In particular, we contrasted auditory responses (A) with visually modulated auditory responses (AV-V). If these responses were different from one another, non-linear multisensory interactions were assumed to take place (e.g. Vroomen and Stekelenburg, 2010). First, we analysed the ERP data at the sensor level by calculating the global field power during the N1 and P2 peaks' time windows. We compared the amplitudes and latencies of those peaks across groups (NH, CI-CHD, CI-SSD) and conditions (A, AV-V). In addition, we used hierarchical clustering to identify dominant topographic maps in the data, which we then used for a single-subject fitting analysis. In particular, we focused on the *map presence* and the *first onset* of the maps within time windows of the N1 and P2, and statistically examined whether there exist group and condition differences. Finally, we performed a source analysis, concentrating on the visual and auditory cortices. Again, we focused on the N1 and P2 time windows, examined the activity peaks and latencies, and statistically compared groups and conditions.

RESULTS At the behavioural level, we confirmed multisensory integration for each experimental group, as we observed the shortest response times for audiovisual syllables compared to auditory-only and visual-only syllables. In line with this observation, multisensory interactions were reflected in the ERP data, exhibiting generally reduced activity for AV-V compared to A for both the N1 and P2, independent of the group. We also discovered differences between the groups, as shown in the topographic analysis which revealed a significant change of N1 voltage topographies for AV-V compared to A, particularly for CI-CHD users. Moreover, both CI user groups exhibited delayed auditory-cortex and enhanced visual-cortex responses, as well as superior lip-reading results compared to NH controls.

DISCUSSION Taken together, both the behavioural and ERP results show an audiovisual benefit for both CI user groups and NH listeners, which is reflected in the response times and in the visual modulation of the auditory response. Furthermore, group differences are discovered in the topographic analysis, which revealed a particularly strong visual influence on auditory processing in CI-CHD users compared to CI-SSD and NH controls at N1 latency. Nevertheless, both CI user groups appear to have difficulties to process the limited CI input which is reflected in delayed auditory-cortex responses compared to NH listeners. To compensate for this, both CI user groups seem to recruit the visual cortex more strongly, and develop better lip-reading skills compared to NH controls. CI-SSD users seem to compensate the limited CI signal by developing enhanced lip-reading abilities and by cross-modally recruiting visual cortex, similarly to CI-CHD users. This is unexpected, given that CI-SSD users have an intact NH ear. We therefore conclude that these compensatory strategies are independent of the hearing threshold and supply of the contralateral ear. Overall, these findings confirm both differences and similarities among different CI user groups, and they support the benefits of audiovisual information for all CI users, which should be taken into account in clinical practise.



Electrophysiological differences and similarities in audiovisual speech processing in CI users with unilateral and bilateral hearing loss

Natalie Layer^{a,*}, Anna Weglage^a, Verena Müller^a, Hartmut Meister^b, Ruth Lang-Roth^a, Martin Walger^{a,b}, Micah M. Murray^{c,d,e,f}, Pascale Sandmann^a

^a University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Otorhinolaryngology, Head and Neck Surgery, Audiology and Pediatric Audiology, Cochlear Implant Center, Germany

^b Jean-Uhrmacher-Institute for Clinical ENT Research, University of Cologne, Germany

^c The Sense Innovation and Research Center, Lausanne and Sion, Switzerland

^d The LINE (The Laboratory for Investigative Neurophysiology), Department of Radiology, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

^e CIBM Center for Biomedical Imaging of Lausanne and Geneva, Lausanne, Switzerland

^f Department of Hearing and Speech Sciences, Vanderbilt University, Nashville, TN, USA

ARTICLE INFO

Keywords:

Cochlear implant
Single-sided-deafness
Bilateral hearing loss
Event-related potential
Cortical plasticity
Multisensory integration
Audiovisual speech perception

ABSTRACT

Hearing with a cochlear implant (CI) is limited compared to natural hearing. Although CI users may develop compensatory strategies, it is currently unknown whether these extend from auditory to visual functions, and whether compensatory strategies vary between different CI user groups. To better understand the experience-dependent contributions to multisensory plasticity in audiovisual speech perception, the current event-related potential (ERP) study presented syllables in auditory, visual, and audiovisual conditions to CI users with unilateral or bilateral hearing loss, as well as to normal-hearing (NH) controls. Behavioural results revealed shorter audiovisual response times compared to unisensory conditions for all groups. Multisensory integration was confirmed by electrical neuroimaging, including topographic and ERP source analysis, showing a visual modulation of the auditory-cortex response at N1 and P2 latency. However, CI users with bilateral hearing loss showed a distinct pattern of N1 topography, indicating a stronger visual impact on auditory speech processing compared to CI users with unilateral hearing loss and NH listeners. Furthermore, both CI user groups showed a delayed auditory-cortex activation and an additional recruitment of the visual cortex, and a better lip-reading ability compared to NH listeners. In sum, these results extend previous findings by showing distinct multisensory processes not only between NH listeners and CI users in general, but even between CI users with unilateral and bilateral hearing loss. However, the comparably enhanced lip-reading ability and visual-cortex activation in both CI user groups suggest that these visual improvements are evident regardless of the hearing status of the contralateral ear.

1. Introduction

A cochlear implant (CI) can help restore the communication abilities in patients with severe to profound sensorineural hearing loss by electrically stimulating the auditory nerve (Zeng, 2011). However, listening with a CI is completely different from conventional hearing, as the electrical signal provided by the CI transmits only a limited amount of spectral and temporal information (Drennan and Rubinstein, 2008). Consequently, the central auditory system must learn to interpret the

artificially sounding CI input as meaningful information (Giraud et al., 2001c; Sandmann et al., 2015). The ability of the nervous system to adapt to a new type of stimulus is an example of neural plasticity (Glennon et al., 2020; Merzenich et al., 2014). This phenomenon has been investigated in various studies with CI users, manifesting as an increase in activation in the auditory cortex to auditory stimuli during the first months after CI implantation (Giraud et al., 2001c; Green et al., 2005; Sandmann et al., 2015). Additional evidence for neural plasticity in CI users comes from the observation that these individuals recruit

* Corresponding author. at: University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Otorhinolaryngology, Head and Neck Surgery, Audiology and Paediatric Audiology, Cochlear Implant Center, Kerpener Strasse 62, 50937, Cologne, Germany.

E-mail address: natalie.layer@uk-koeln.de (N. Layer).

<https://doi.org/10.1016/j.crneur.2022.100059>

Received 25 April 2022; Received in revised form 24 August 2022; Accepted 7 October 2022

Available online 8 November 2022

2665-945X/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

visual cortices for purely auditory speech tasks (Chen et al., 2016; Giraud et al., 2001c); a phenomenon referred to as cross-modal plasticity (e.g. Glennon et al., 2020).

Previous research has shown that event-related potentials (ERPs) derived from continuous electroencephalography (EEG) are an adequate method for studying cortical plasticity in CI users (Beynon et al., 2005; Finke et al., 2016a; Sandmann et al., 2009, 2015; Schierholz et al., 2015, 2017; Sharma et al., 2002; Viola et al., 2012; Layer et al., 2022). The primary benefit of analysing ERPs is the high temporal resolution, which allows for the tracking of individual cortical processing steps (Biasucci et al., 2019; Michel and Murray, 2012). For instance, the auditory N1 (negative potential around 100 ms after stimulus onset) and the auditory P2 ERPs (positive potential around 200 ms after stimulus onset) are at least partly generated in the primary and secondary auditory cortices (Ahveninen et al., 2006; Bosnyak et al., 2004; Näätänen and Picton, 1987). Current models of auditory signal propagation recognise that there is an underlying anatomy exhibiting a semi-hierarchical and highly parallel organisation (e.g. Kaas and Hackett, 2000). In terms of auditory ERPs this would suggest that prominent components, such as the N1–P2 complex, include generators not only within primary auditory cortices, but within a distributed network along the superior temporal cortices as well as fronto-parietal structures and even visual cortices. Moreover, top-down effects, such as attention or expectation of incoming auditory events mediated by the frontal cortex, can influence these auditory processes (Dürschmid et al., 2019). The majority of previous ERP studies with CI users have used auditory stimuli to show that the N1 and P2 ERPs have a reduced amplitude and a prolonged latency in comparison to normal-hearing (NH) individuals. This observation suggests that CI users have difficulties in processing auditory stimuli (Beynon et al., 2005; Finke et al., 2016a; Henkin et al., 2014; Sandmann et al., 2009) and is consistent with previous behavioural results of impaired auditory discrimination ability in CI users (Sandmann et al., 2010, 2015; Finke et al., 2016a,b).

Although multisensory conditions more likely represent everyday situations, only a few ERP studies so far have been conducted with audiovisual stimuli. These studies primarily concentrated on rudimentary, non-linguistic audiovisual stimuli (sinusoidal tones and white discs) and showed a prolonged N1 response, and a greater visual modulation of the auditory N1 ERPs in CI users compared to NH listeners (Schierholz et al., 2015, 2017). Our previous study extended these results to more complex audiovisual syllables (Layer et al., 2022), using electrical neuroimaging (Michel et al., 2004, 2009; Michel and Murray, 2012) to perform topographic and ERP source analyses. Unlike traditional ERP data analysis, which is based on waveform morphology at specific electrode positions, electrical neuroimaging is reference-independent and takes into account the spatial characteristics and temporal dynamics of the global electric field to distinguish between the effects of response strength, latency, and distinct topographies (Murray et al., 2008; Michel et al., 2009). By using this topographic analysis approach, we previously showed a group-specific topographic pattern at N1 latency and an enhanced activation in the visual cortex at N1 latency for CI users when compared to NH listeners (Layer et al., 2022). These observations confirm a recent report about alterations in audiovisual processing and a multisensory benefit for CI users, if additional (congruent) visual information is provided (Radecke et al., 2022).

Based on previous ERP and behavioural results, one might conclude that multisensory processes, in particular integration of auditory and visual speech cues, remain intact in CI users despite the limited auditory signal provided by the CI. Nevertheless, it remains unclear whether the enhanced visual impact on auditory speech processing applies to all of the CI users, given that large inter-individual differences (e.g. with regards to the hearing threshold in the contralateral ear) have not been taken into account. Most of the aforementioned studies have included CI users with *bilateral hearing loss* (e.g. Finke et al., 2016a; Sandmann et al., 2015; Schierholz et al., 2015; Radecke et al., 2022; Layer et al., 2022), either provided with a CI on both ears (CI + CI on contralateral side) or

on one ear (CI + hearing aid on contralateral side). These CI users will be referred to as CI-CHD users (CHD = ‘contralateral hearing device’) in the following. However, over the last years, the clinical margins for CI indication have been extended to *unilateral hearing loss*, enabling the implantation of single-sided deaf (SSD) patients (CI + NH on contralateral side; Arndt et al., 2011; Arndt et al., 2017; Buechner et al., 2010). This CI user group is particularly interesting, as the signal quality of the input is very different for the two ears, leading to maximally asymmetric auditory processing (Gordon et al., 2013; Kral et al., 2013). The variable hearing ability in the contralateral ear across different CI groups may at least partly account for the large variability in speech recognition ability observed in CI users (Lazard et al., 2012). To better understand the factors contributing to this variability and to extend previous findings on CI-CHD users and NH listeners (Layer et al., 2022), the current study systematically compared the timecourse of auditory and audiovisual speech processing as well as the lip-reading abilities between different groups of CI users, in particular CI-CHD users and CI-SSD users, and to NH listeners. The inclusion of the additional group of patients crucially extends our previous study because it not only evaluates the transferability of our previous results to different patient groups but also provides deeper insights into the influence of individual factors – specifically the hearing ability of the second ear – on audiovisual speech processing in CI users. This is noteworthy because literature comparing CI-SSD to bimodal or bilateral CI users is scarce. However, the few existing studies reported differences in speech-in-noise performance (Williges et al., 2019) and in situations with multiple concurrent speakers (Bernstein et al., 2016) between CI-SSD users and bimodal or bilateral CI users, respectively. But, given this first evidence for purely auditory situations, we hypothesised that further differences would emerge for audiovisual stimulation, which has yet to be reported.

Given that CI-SSD users have an intact ear on the contralateral side, it is reasonable that this NH ear serves as the main communication channel despite the advantages given by the CI (Kitterick et al., 2015; Ludwig et al., 2021). Therefore, we hypothesised that CI-SSD users are less influenced by visual information, benefit less from audiovisual input and show poorer lip-reading skills than CI-CHD users. However, we expected a delay in cortical responses in the CI-SSD group, similar to the group of CI-CHD users, when compared to NH individuals, based on previous results from studies with purely auditory stimuli comparing the CI and the NH ear (Finke et al., 2016b; Bönitz et al., 2018; Weglage et al., 2022).

2. Material and methods

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the medical faculty of the University of Cologne (application number: 18 – 257). Prior to data collection, all participants gave written informed consent, and they were reimbursed.

2.1. Participants

In total, twelve post-lingually deafened CI-SSD patients were invited to participate in this study to extend the results from our previous study (Layer et al., 2022) by including an additional subgroup of CI-SSD users. Among these participants, one had to be excluded due to poor EEG data quality (high artefact load), resulting in a total of eleven CI-SSD patients (two right-implanted). Accordingly, we selected post-lingually deafened CI-CHD patients and NH listeners from our previous study (Layer et al., 2022; $n = 11$ each) such that they matched the CI-SSD patients as best as possible by gender, age, handedness, stimulated ear and years of education. The matched subset datasets from our previous study of CI-CHD users and NH listeners were reused for the current study and were extended by newly acquired data from an additional group of CI-SSD users. The CI-CHD users were implanted either unilaterally ($n = 2$; all left-implanted using a hearing aid on the contralateral ear) or bilaterally

($n = 9$; two right-implanted). All CI users had been using their device continuously for at least one year prior to the experiment. For the experiment, only the ear with a CI was stimulated and in the case of bilateral implantation, the ‘better’ ear (the ear showing the higher speech recognition scores in the Freiburg monosyllabic test) was used as stimulation side.

Thus, for final analyses, thirty-three volunteers were included, with eleven CI-SSD patients (7 female, mean age: 56.5 years $\pm$ 9.2 years, range: 39 – 70, 10 right-handed), eleven CI-CHD patients (7 female, mean age: 61.4 years $\pm$ 9.7 years, range: 39 – 75, 11 right-handed) and eleven NH listeners (7 female, mean age: 60.1 years $\pm$ 10.1 years, range: 34 – 70, 11 righthanded). Detailed information on the implant system and the demographic data are provided in Table 1. To check that cognitive abilities were age-appropriate, the DemTect Ear test battery was used (Brünecke et al., 2018). All participants scored within the normal, age-appropriate range (13 – 18 points). In addition, the German Freiburg monosyllabic speech test (Hahlbrock, 1970) with a sound intensity level of 65 dB SPL (see Table 4 for scores) was used to assess speech recognition abilities. To obtain a hearing threshold (HT) of the contralateral ear, we measured the aided HT in CI-CHD users in free-field and the unaided HT of the NH ear of CI-SSD users with headphones (see Table 1). All participants were native German speakers, had normal or corrected-to-normal vision (assessed by the Landolt test according to the DIN-norm; Wesemann et al., 2010) and none of the participants had a history of psychiatric disorder. Their handedness was assessed by the Edinburgh inventory (Oldfield, 1971).

2.2. Stimuli

The stimuli in this study were identical to those used in our previous study (Layer et al., 2022) and they were presented in three different conditions: visual-only (V), auditory-only (A) and audiovisual (AV). Additionally, there were trials with a black screen only (‘nostim’), to which the participants were instructed to not react. The stimuli were delivered using the Presentation software (Neurobehavioral Systems, version 21.1) and a computer in combination with a duplicated monitor (69 inch). The stimuli consisted of the two syllables /ki/ and /ka/ which are included in the Oldenburg logatome speech corpus (OLLO; Wesker et al., 2005). They were cut from the available logatomes from one speaker (female speaker 1, V6 ‘normal spelling style’, no dialect). These two syllables in particular differed in their phonetic distinctive features (vowel place and height of articulation) in the vowel contrast (/a/ vs. /i/; Micco et al., 1995). These German vowels are different in terms of central frequencies of the first (F1) and second formant (F2) representing the highest contrast between German vowels (e.g. Obleser et al., 2003), making them easily distinguishable for CI users. Importantly, as we presented visual-only syllables as well, the chosen syllables not only highly differ in terms of auditory (phoneme) realisation, but also in their visual articulatory (viseme) realisation. A viseme is the visual equivalent of the phoneme: a static image of a person articulating a phoneme (Dong et al., 2003). The editing of the syllables was done with Audacity (version 3.0.2) by cutting and adjusting them to the same duration of 400 ms. The syllables were normalised (adjusted to the maximal amplitude) in Adobe Audition CS6 (version 5.0.2).

To create a visual articulation of the auditory syllables, we used the

Table 1

Demographic information on the CI participants; HT (hearing threshold; average over 500 Hz, 1 kHz, 2 kHz, 4 kHz); Stim. = stimulated; HL = hearing loss; m = male; f = female.

ID	Group	Sex	Age	Handedness	Fitting	HT (dB HL; contralateral ear)	Stim. ear	Etiology	Age at onset of HL (years)	CI use of the stim. ear (months)	CI manufacturer
1	CI-CHD	m	61	right	bilateral	32	left	unknown	41	15	MedEl
2	CI-CHD	f	75	right	bilateral	31	left	hereditary	57	30	Advanced Bionics
3	CI-CHD	f	39	right	bilateral	26	right	otosclerosis	24	17	Advanced Bionics
4	CI-CHD	f	70	right	bilateral	37	left	unknown	37	56	MedEl
5	CI-CHD	f	70	right	bilateral	37	left	meningitis	69	20	MedEl
6	CI-CHD	m	59	right	bimodal	85	left	unknown	49	33	Advanced Bionics
7	CI-CHD	f	63	right	bilateral	36	left	meningitis	20	106	Advanced Bionics
8	CI-CHD	f	64	right	bilateral	29	left	whooping cough	9	78	Cochlear
9	CI-CHD	m	53	right	bilateral	36	left	unknown	30	235	Cochlear
10	CI-CHD	f	58	right	bimodal	41	left	unknown	49	18	Advanced Bionics
11	CI-CHD	m	56	right	bilateral	35	right	hereditary	19	63	MedEl
12	CI-SSD	f	64	right	SSD	10	left	unknown	49	30	Cochlear
13	CI-SSD	f	40	right	SSD	10	right	sudden hearing loss	34	77	MedEl
14	CI-SSD	m	43	right	SSD	12	left	bike accident	42	12	Cochlear
15	CI-SSD	m	54	right	SSD	17	left	unknown	52	28	MedEl
16	CI-SSD	f	49	right	SSD	23	right	otosclerosis	39	19	MedEl
17	CI-SSD	m	59	left	SSD	17	left	sudden hearing loss	49	54	Cochlear
18	CI-SSD	f	57	right	SSD	12	left	sudden hearing loss	20	53	Cochlear
19	CI-SSD	m	62	right	SSD	13	left	sudden hearing loss	47	63	Cochlear
20	CI-SSD	f	62	right	SSD	15	left	otitis media	14	65	MedEl
21	CI-SSD	f	68	right	SSD	25	left	unknown	60	50	Cochlear
22	CI-SSD	f	52	right	SSD	22	left	hereditary	20	12	MedEl

MASSY (Modular Audiovisual Speech SYNthesizer; [Fagel and Clemens, 2004](#)), which is a computer-based video animation of a talking head. This talking head has been previously validated for CI users ([Meister et al., 2016](#); [Schreitmüller et al., 2018](#)) and is an adequate tool to generate audiovisual and visual speech stimuli ([Massaro and Light, 2004](#)). To generate articulatory movements matching the auditory speech sounds, one has to provide files that transform the previously transcribed sounds into a probabilistic pronunciation model providing the segmentation and the timing of every single phoneme. This can be done by means of the web-based tool MAUS (Munich Automatic Segmentation; [Schiel, 1999](#)). To obtain a video file of the MASSY output, the screen recorder Bandicam (version 4.1.6) was used in order to save the finished video files. Finally, the stimuli were edited in Pinnacle Studio 22 (version 22.3.0.377), making video files of each syllable in each condition: 1) Audiovisual (AV): articulatory movements with corresponding auditory syllables, 2) Auditory-only (A): black screen (video track turned off) combined with auditory syllables, 3) Visual-only (V): articulatory movements without auditory syllables (audio track turned off). Each trial started with a static face (500 ms) and was followed by the video, which lasted for 800 ms (20 ms initiation of articulatory movements + 400 ms auditory syllable + 380 ms completion of articulatory movements). For further analyses, we focused on the moving face (starting 500 ms post-stimulus onset/after the static face), as the responses to static faces comparing NH listeners and CI users have been investigated previously ([Stropahl et al., 2015](#)).

Supplementary video related to this article can be found at <https://doi.org/10.1016/j.crneur.2022.100059>

In general, all participants were assessed monaurally, meaning that in CI-SSD users the CI-ear was measured, in bimodal CI users the CI-ear was measured and in bilateral CI users the better CI-ear was measured. Regarding the *CI-SSD patients*, we followed the procedures of a previous study ([Weglage et al., 2022](#)) and positioned the processor inside an aqua case (Advanced Bionics; <https://www.advancedbionics.com>) to specifically stimulate the CI-ear without risking a cross-talking to the NH ear. Note that stimulation via a loudspeaker is not possible in this group, because an ear-plug is not enough to cover and mask the NH-ear, as it only reduces the intensity level by maximally 30 dB ([Park et al., 2021](#)). We refused the option of using noise to mask the NH ear, as this noise would have to be very loud to fully mask the information. This stimulation option would rather represent a speech-in-noise condition and would differ much more from the other two groups. An insert earphone (3M E-A-RTONE 3A) was put inside a hole of the aqua case where it was placed right over the microphone of the CI. A long coil cable was used to connect the processor with the implant. Regarding the *CI-CHD users*, the stimuli were presented via a loudspeaker (Audiometer-Box, type: LAB 501, Westra Electronic GmbH) which was placed in front of the participant. The hearing aid or the CI at the contralateral side was removed during the experiment and the ear was additionally covered with an ear-plug. For *NH participants*, the ear of the matched CI user was stimulated via an insert earphone (3M E-A-RTONE 3A), and the contralateral ear was closed with an ear-plug as well to avoid a cross-talking to the contralateral ear. The stimuli were calibrated to 65 dB SPL to ensure that the intensity level was equal for each stimulation technique. All participants rated the perceived loudness of the syllables with a seven-point loudness rating scale (as used in [Sandmann et al., 2009, 2010](#)), to ensure that the syllable intensity was perceived at a moderate level of 60–70 dB ([Allen et al., 1990](#)). The stimuli (video files) are provided as supplementary material and can be downloaded.

2.3. Procedure

The procedure was identical to our previous study (see [Layer et al., 2022](#)). The additional CI-SSD users were seated comfortably in an electromagnetically shielded and dimly lit booth at a viewing distance of 175 cm to the screen. The participants were instructed to discriminate as fast and as accurately as possible between the syllables /ki/ and /ka/.

The given response was registered using a mouse, with each of the two buttons assigned to one syllable. The sides were counterbalanced across the participants to prevent confounds caused by the used finger.

For each condition (AV, A, V, 'nostim'), 90 trials each were presented per syllable, resulting in a total number of 630 trials (90 repetitions x 3 conditions (AV, A, V) x 2 syllables (/ki/, /ka/) + 90 'nostim'-conditions). Each trial began with a 'nostim'-condition or a static face of the talking head (500 ms) followed by a visual-only, auditory-only or an audiovisual syllable. Afterwards, a fixation cross was shown until the participant pressed a button. In the case of a 'nostim'-trial, the participants were asked not to respond to. The trials were pseudo-randomised such that no trial of the same condition and syllable appeared twice in a row. The experiment lasted for 25 min excluding breaks, composed of five blocks of approximately 5 min each. A short break was given after each block. To ensure that the task was understood by the participants, we presented a short practice block consisting of five trials per condition before starting the recording. An illustration of the experimental paradigm can be found in [Fig. 1A](#).

To obtain further behavioural measures apart from the ones registered in the EGG task (hit rates, response times), we asked the additionally measured CI-SSD users to rate the exertion of performing the task after each experimental block by using the 'Borg Rating of Perceived Exertion'-scale (Borg RPE-scale; [Williams, 2017](#)). Further, we measured the lip-reading abilities by means of a behavioural lip-reading test consisting of monosyllabic words from the German Freiburg test ([Hahlbrock, 1970](#)) which were visualised by various speakers and filmed ([Stropahl et al., 2015](#)). The participants were asked to watch the short videos of the muted monosyllabic word performances and to report which word was understood. This test was used in previous studies with CI patients ([Stropahl et al., 2015](#); [Stropahl and Debener, 2017](#)), as well as in our previous study with CI-CHD users and NH listeners ([Layer et al., 2022](#)), whose scores we compared to CI-SSD users in the current study.

2.4. EEG recording

Similar to our previous study ([Layer et al., 2022](#)), the EEG data of the additionally measured CI-SSD users were continuously recorded by means of 64 AG/AgCl ActiCap slim electrodes using a BrainAmp system (BrainProducts, Gilching, Germany) and a customised electrode cap with an electrode layout (EasyCap, Herrsching, Germany) according to the 10-10 system. To record an electrooculogram (EOG), two electrodes were placed below and beside the left eye (vertical and horizontal eye movements, respectively). The nose-tip was used as reference, and a midline electrode placed slightly anterior to Fz served as ground. Data recording was performed using a sampling rate of 1000 Hz. The online analog filter was set between 0.02 and 250 Hz. Electrode impedances were maintained below 10 k Ω during data acquisition.

2.5. Data analysis

The subset data taken from our previous study including the newly acquired data of the CI-SSD users were analysed in MATLAB 9.8.0.1323502 (R2020a; Mathworks, Natick, MA) and R (version 3.6.3; [R Core Team \(2020\)](#), Vienna, Austria). Topographic analyses were carried out in CARTOOL (version 3.91; [Brunet et al., 2011](#)). Source analyses were performed in Brainstorm ([Tadel et al., 2011](#)). The following R packages have been used: ggplot2 (version 2.3.3) for creating plots; dplyr (version 1.0.4), tidyverse (version 1.3.0) and tidyr (version 1.1.3) for data formatting; ggpubr (version 0.4.0) and rstatix (version 0.7.0) for statistical analyses.

2.5.1. Behavioural data

In a first step, we collapsed the syllables /ki/ and /ka/ for each condition (A, V, AV), as they did not show substantial differences between each other (mean RTs $\pm$ one standard deviation of the mean: CI-CHD: /ki/ = 620 ms $\pm$ 88.3 ms, /ka/ = 611 ms $\pm$ 88.0 ms; CI-SSD: /ki/ =

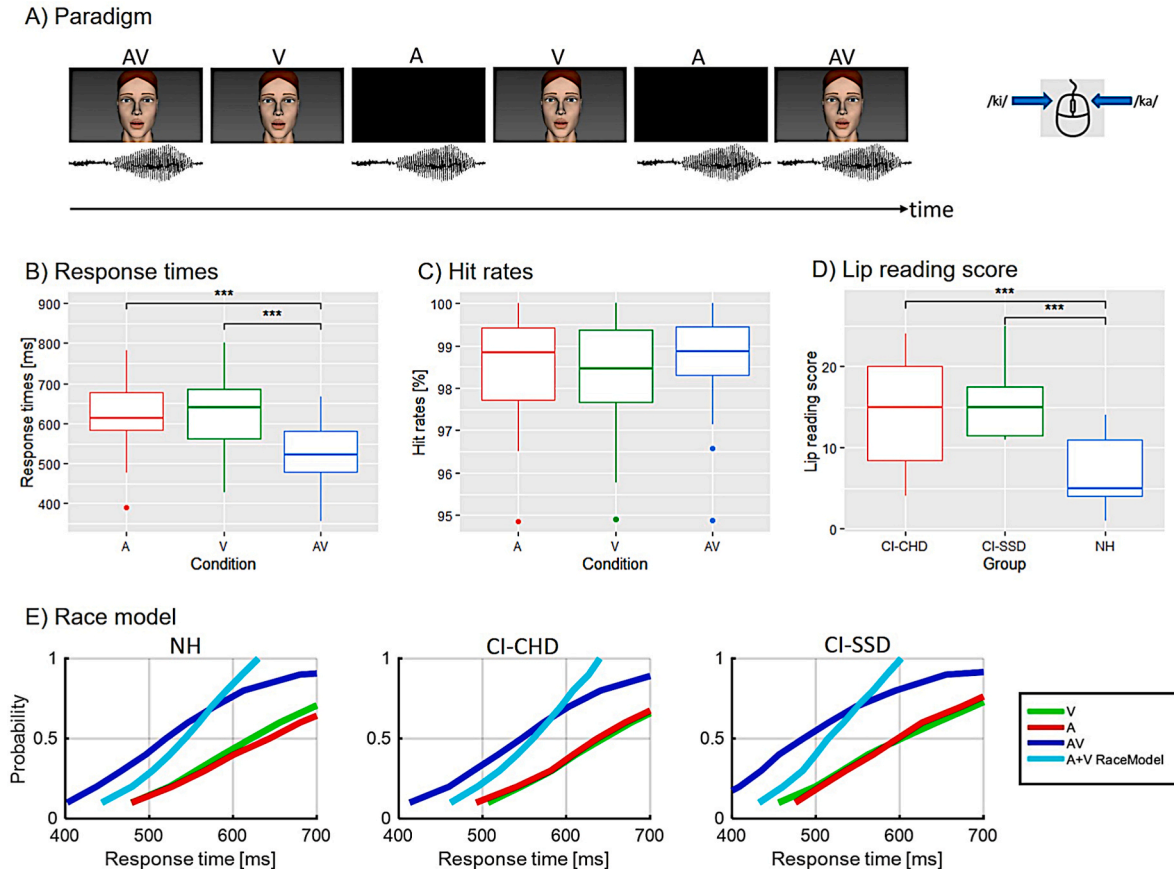


Fig. 1. Behavioural results. A) Overview of the paradigm (adapted from Layer et al., 2022). B) Mean response times for auditory (red), visual (green) and audiovisual (blue) syllables averaged over all groups, demonstrating that audiovisual syllables had shorter response times than auditory-only and visual-only RTs. C) Mean hit rates for auditory (red), visual (green) and audiovisual (blue) syllables averaged across all groups, with no differences between the three conditions. D) Cumulative distribution functions for CI-CHD, CI-SSD and NH. The race model is violated for all three groups because they show that the probability of faster response times is higher for audiovisual stimuli (blue line) than for those estimated by the race model (cyan line). Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

574 ms $\pm$ 100.0 ms, /ka/ = 566 ms $\pm$ 94.0; NH: /ki/ = 607 ms $\pm$ 91.6 ms, /ka/ = 587 ms $\pm$ 102.0 ms; all $p \geq 0.354$). Second, false alarms or missing responses were discarded from the dataset. Trials that exceeded the individual mean by more than three standard deviations for each condition were declared as outliers and were removed from the dataset. Then, RTs and hit rates were computed for each condition (A, V, AV) for each individual. To analyse the performance for each condition and group, a 3 $\times$ 3 mixed ANOVA was used separately for the RTs and the hit rates, with condition (AV, A, V) as the within-subjects factor and group (NH, CI-CHD, CI-SSD) as the between-subjects factor. In the case of violation of the sphericity assumption, a Greenhouse-Geisser correction was applied. Moreover, post-hoc t-tests were carried out and corrected for multiple comparisons using a Bonferroni correction, in the case of significant main effects or interactions ($p \leq .05$). As the hit rates were very high in our previous study (Layer et al., 2022), we did not expect CI-SSD users to deviate from this pattern. Concerning the RTs, we expected similar results for CI-SSD users as for CI-CHD users and NH listeners, with shorter RTs for AV conditions compared to unisensory (A, V) conditions.

In a next step, we analysed the origin of the redundant signals effect, which is the effect of achieving faster RTs for audiovisual stimuli in comparison to unimodal stimuli (A, V) (Miller, 1982). For this purpose, we reused a subset of our previously reported data of the CI-CHD users and NH listeners (Layer et al., 2022) and extended these by the additional group of CI-SSD users. There are two accounts explaining this issue: the race model (Raab, 1962) and the coactivation model (Miller, 1982). Briefly, the race model claims that due to statistical facilitation it

is more probable that either of the stimuli (A and V) will result in shorter response times in comparison to one stimulus alone (A or V). Therefore, one can assume that RTs of redundant signals (AV) are significantly faster, and that no neural integration is required to observe a redundant signals effect (Raab, 1962). In contrast, the coactivation model (Miller, 1982) assumes an interaction between the unimodal stimuli which forms a new product before initiating a motor response, leading to faster RTs. A widely used method in multisensory research is to test for the race model inequality (RMI; Miller, 1982) to explain whether the redundant signals effect was caused by multisensory processes or by statistical facilitation. According to the RMI, the cumulative distribution function (CDF) of the RTs in the multisensory condition (AV) can never exceed the sum of the CDFs of the two unisensory (A, V) conditions:

$$P(RT_{AV} \leq t) \leq P(RT_A \leq t) + P(RT_V \leq t), \text{ for all } t \leq 0,$$

where $P(RT_x \leq t)$ represents the likelihood of a condition $x \in \{AV, A, V\}$ being less than an arbitrary value t . Violation of this model, for any given value of t , is an indication for multisensory processes (see also Ulrich et al. (2007) for details). By applying the RMITest software by Ulrich et al. (2007), the CDFs of the RT distributions for each condition (AV, A, V) and for the sum of the modality-specific conditions (A + V) were estimated. The individual RTs were rank ordered for each condition to obtain percentile values (Ratcliff, 1979). Next, for each group separately (NH, CI-CHD, CI-SSD), the CDFs for the redundant signals conditions (AV) and the modality-specific sum (A + V) were compared for the five fastest deciles (bin width: 10 %). We used one-tailed paired t-tests

followed by a Bonferroni correction to account for multiple comparisons. Significance at any decile bin was treated as violation of the race model, suggesting multisensory interactions in the behavioural responses. Here, we expected a similar redundant signals effect for CI-SSD users, as CI-CHD users and NH listeners both showed a violation of the race model inequality in our previous study (Layer et al., 2022).

To assess differences between the CI user groups and the NH listeners in the lip-reading task and in the subjective rating of exertion, we performed one-way ANOVAs. Whenever a significant main effect of group was present, we performed follow-up tests with a Bonferroni correction to account for multiple comparisons. Concerning the lip-reading ability, we anticipated that CI-SSD users performed worse compared to CI-CHD users due to their intact contralateral ear, which may reduce the need to rely on lip movements in their everyday life. In terms of subjective exertion rating, we expected no difference between experimental groups because our previous study (Layer et al., 2022) found no difference, which was likely due to the easy task.

2.5.2. EEG pre-processing

The pre-processing of the EEG data was done with EEGLAB (version v2019.1; Delorme and Makeig, 2004), a software working within the MATLAB environment (Mathwork, Natick, MA). The raw data were down-sampled (500 Hz) and filtered with a FIR-filter, having a high pass cut-off frequency of 0.5 Hz with a maximum possible transition bandwidth of 1 Hz (cut-off frequency multiplied by two), and a low pass cut-off frequency of 40 Hz with a transition bandwidth of 2 Hz. For both filters, the Kaiser-window (Kaiser- $\beta = 5.653$, max. stopband attenuation = -60 dB, max. passband deviation = 0.001; Widmann et al., 2015) was used to maximise the energy concentration in the main lobe by averaging out noise in the spectrum and minimising information loss at the edges of the window. Electrodes in the proximity of the speech processor and transmitter coil were removed for CI users (mean: 2.8 electrodes; range: 1 – 4). Afterwards, the datasets were epoched into 2 s dummy epoch segments, and pruned of unique, non-stereotype artefacts using an amplitude threshold criterion of four standard deviations. An independent component analysis (ICA) was computed (Bell and Sejnowski, 1995) and the resulting ICA weights were applied to the epoched original data (1 – 40 Hz, -200 to 1220 ms relative to the stimulus onset (including the static and moving face)). Independent components reflecting vertical and horizontal ocular movements, electrical heartbeat activity, as well as other sources of non-cerebral activity were rejected (Jung et al., 2000). Independent components exhibiting artefacts of the CI were identified based on the side of stimulation and the time course of the component activity, showing a pedestal artefact around 700 ms after the auditory stimulus onset (520 ms). The identified components were removed from the EEG data. In a next step, we interpolated the missing channels using a spherical spline interpolation (Perrin et al., 1989) which allows for a solid dipole source localisation of auditory ERPs in CI users (Debener et al., 2008; Sandmann et al., 2009). Only trials yielding correct responses (NH: 91.0 % $\pm$ 3.9 %; CI-CHD: 88.1 % $\pm$ 3.8 %; CI-SSD: 86.6 % $\pm$ 4.5 %) were kept for further ERP analyses.

2.5.3. EEG data analysis

We compared event-related potentials (ERPs) of all conditions (AV, A, V) between the two CI user groups and the NH participants. The additive model which is denoted by the equation $AV = A + V$ (Barth et al., 1995) was used to investigate multisensory interactions. The model is satisfied and suggests independent and linear processing if the multisensory (AV) responses equal the sum of the unisensory (A, V) responses. Whereas, if the model is not satisfied, non-linear interactions between the unisensory modalities are assumed (Barth et al., 1995). Similar to our previous study (Layer et al., 2022), we rearranged the equation to $A = AV - V$ such that we could compare the directly measured auditory ERP response (A) with the term $[AV - V]$, denoting an ERP difference wave representing a visually-modulated auditory ERP response. Hence, $[AV - V]$ is an estimate of an auditory response elicited

in a multisensory context. In the case of a lack of interaction between the two unisensory (A, V) modalities, both A and AV-V should be identical. However, if the auditory (A) and the modulated auditory (AV-V) ERPs are *not* identical, this would point to non-linear multisensory interactions (Besle et al., 2004; Murray et al., 2005; Cappe et al., 2010; Foxe et al., 2000; Molholm et al., 2002). Such non-linear effects can be either super-additive ($A < AV - V$) or sub-additive ($A > AV - V$). But, since interpreting these effects is not straightforward, it is necessary to obtain reference-independent measurements of power or of source estimates (e.g. Cappe et al., 2010). Before creating the difference waves (AV-V), we randomly reduced the number of epochs based on the condition with the lowest number of epochs for each individual to guarantee that there was an equal contribution of each condition to the resulting difference wave. The difference waves were only created for the CI-SSD users in this study, and the difference waves for the NH listeners and CI-CHD users were reused from our previous study (Layer et al., 2022).

As in our previous study (Layer et al., 2022), we analysed our ERP data within an electrical neuroimaging framework (Murray et al., 2008; Michel et al., 2009; Michel and Murray, 2012), comprising topographic and ERP source analysis to compare auditory (A) and modulated (AV-V) ERPs within and between groups (NH, CI-CHD, CI-SSD). We investigated the global field power (GFP) and the global map dissimilarity (GMD) to quantify ERP differences in response strength and response topography, respectively (Murray et al., 2008). First, we looked at the GFP, at the time window of the N1 and the P2 (N1: 80 – 200 ms; P2: 200 – 370 ms), which were chosen based on visual inspection of the GFP computed for the grand average ERPs across conditions and groups. The GFP is the spatial standard deviation of all electrode values at a specific time point (Murray et al., 2008) and was first described by Lehmann and Skrandies (1980). The reason for choosing the GFP instead of selecting specific channels of interest is that this approach allows for a more objective peak detection. The GFP peak mean amplitudes and latencies were detected for each individual, condition (A, AV-V) and time window (N1, P2) and were statistically analysed by using a 3 x 2 mixed ANOVA with group (NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) as the within-subjects factor for each peak separately. Based on previous observations with CI-CHD users (Beynon et al., 2005; Finke et al., 2016a; Henkin et al., 2014; Sandmann et al., 2009; Layer et al., 2022) and CI-SSD users (Finke et al., 2016b; Bönitz et al., 2018; Weglage et al., 2022), we expected delayed N1 and reduced P2 responses for all CI user groups compared to NH controls.

Second, we analysed the GMD (Lehmann and Skrandies, 1980) to quantify topographic dissimilarities (and by extension, dissimilar configurations of neural sources; Vaughan Jr, 1982) between experimental conditions and groups, regardless of the signal strength (Murray et al., 2008). The GMD was analysed in CARTOOL by computing a 'topographic ANOVA' (TANOVA; Murray et al., 2008) to quantify differences in topographies between groups for each condition. Even though the name is misleading, this is no analysis of variance, but a non-parametric randomisation test. This randomisation test was executed with 5.000 permutations and by calculating sample-by-sample p-values. To control for multiple comparisons, an FDR correction was applied (FDR = false discovery rate; Benjamini and Hochberg, 1995). Since ERP topographies remain stable for a certain period of time before changing to another topography (called 'microstates'; Michel and Koenig, 2018) and to account for temporal autocorrelation, the minimal significant duration was adjusted to 15 consecutive time frames, corresponding to 30 ms.

2.5.4. Hierarchical clustering and single-subject fitting analysis

Whenever differences in topographies (GMD) for two groups or conditions are found, this is an indication for distinct neural generators contributing to these topographies (e.g. Vaughan Jr, 1982). However, it is also possible that a GMD is caused by a latency shift of the ERP, meaning that the same topographies are present but just shifted in time (Murray et al., 2008). To disentangle these two possible GMD causes, we performed a hierarchical topographic clustering analysis with

group-averaged data (NH(A), NH(AV-A), CI-CHD(A), CI-CHD(AV-V), CI-SSD(A), CI-SSD(AV-V)) to identify template topographies within the time windows of interest (N1, P2). Again, we used CARTOOL for this analysis and chose the atomize and agglomerate hierarchical clustering (AAHC) which has been especially designed for ERP-data (detailed in Murray et al., 2008). This method includes the global explained variance of a cluster and prevents blind combinations (agglomerations) of short-duration clusters. Thus, this clustering method identifies the minimal number of topographies accounting for the greatest variance within a dataset (here NH(A), NH(AV-A), CI-CHD(A), CI-CHD(AV-V), CI-SSD(A), CI-SSD(AV-V)).

In a next step, the template maps detected by the AAHC were entered into a single-subject fitting (Murray et al., 2008) to identify the distribution of specific templates by calculating sample-wise spatial correlations for each individual and condition between each template topography and the observed voltage topographies. Each sample was matched to the template map with the largest spatial correlation. For statistical analyses, we focused on the *first onset of maps* (latency) and the *map presence* (number of samples in time frames) which are two of many other output options provided by CARTOOL. Particularly, we performed a mixed ANOVA with group (NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as within-subject factors, separately for each time window (N1, P2). In the case of significant three-way interactions, group-wise mixed ANOVAs (condition x template map) were computed. A Greenhouse-Geisser correction was applied whenever there was a violation of the sphericity assumption. Post-hoc t-tests were computed and corrected for multiple comparisons using a Bonferroni correction. We anticipated that the analysis of the first onset of maps would confirm a delayed N1 latency for both CI-CHD and CI-SSD users based on previous results. In terms of map presence at N1 latency range, we speculated that there would be a pattern between CI-CHD users and NH listeners for CI-SSD users, as they have both a CI and a NH ear. However, we are not aware of previous studies reporting similar results for CI-SSD users.

2.5.5. Source analysis

We performed an ERP source analysis for each group and condition by means of Brainstorm (Tadel et al., 2011) to find out whether topographic differences can be explained by fundamentally different configurations of neural generators. The tutorial provided by Stropahl et al. (2018) served as guideline for conducting the source analysis. As in our previous study (Layer et al., 2022) and in various studies with CI patients (Bottari et al., 2020; Stropahl et al., 2015; Stropahl and Debener, 2017), we selected the method of dynamic statistical parametric mapping (dSPM, Dale et al. (2000)). dSPM works more precisely in identifying deeper sources than standard norm methods, even though the spatial resolution stays relatively low (Lin et al., 2006). It takes the minimum-norm inverse maps with constrained dipole orientations to approximate the locations of electrical activity recorded on the scalp. This method can be successfully used to localise small cortical areas such as the auditory cortex (Stropahl et al., 2018). First, individual noise covariances were calculated from single-trial pre-stimulus onset baseline intervals (−50 to 0 ms) to estimate single-subject based noise standard deviations at each location (Hansen et al., 2010). As a head model, the boundary element method (BEM) which is implemented in OpenMEEG was used. The BEM gives three realistic layers and representative anatomical information (Gramfort et al., 2010). The final activity data is then displayed as absolute values with arbitrary units based on the normalisation within the dSPM algorithm. Consistent with the procedures of our previous study (Layer et al., 2022), we defined an *auditory* and a *visual* ROI by combining smaller regions within the Destrieux-atlas (Destrieux et al., 2010; Tadel et al., 2011; auditory: G_temp_sup-G_T_transv, S_temporal_transverse, G_temp_sup-Plan_tempo and Lat_Fis-post; visual: G_cuneus, S_calcarine, S_parieto_occipital). These ROIs were chosen in accordance with several previous studies (Stropahl et al., 2015; Stropahl and Debener, 2017; Giraud et al., 2001b,

; Prince et al., 2021; Layer et al., 2022). In specific, the chosen parts of the auditory ROI have been reported as both N1 (Näätänen and Picton, 1987; Godey et al., 2001; Woods et al., 1993; Bosnyak et al., 2004) and P2 (Crowley and Colrain, 2004 (for review); Hari et al., 1987; Bosnyak et al., 2004; Ross and Tremblay, 2009) generators. The selected ROIs can be viewed in Fig. 4A.

Source activities for each ROI, condition and group were exported from Brainstorm for each participant. Afterwards, the peak means and latencies for each time window of interest (N1: 80 – 200 ms, P2: 200 – 370 ms) were extracted. A mixed-model ANOVA was performed separately for each time window with group (NH, CI-CHD, CI-SSD) as between-subject factor and condition (A, AV-V), ROI (auditory, visual) and hemisphere (left, right) as within-subjects factors. A Greenhouse-Geisser correction was applied in the case of violation of the sphericity assumption. In the case of significant interactions or main effects, post-hoc t-tests were computed and corrected for multiple comparisons using a Bonferroni correction. Similar to our hypotheses for the fitting data, we speculated that a pattern between the one of CI-CHD users and NH listeners would emerge for the recruitment of the visual cortex, which we observed for CI-CHD users in our previous study (Layer et al., 2022). In addition, in accordance with the fitting data and the GFP, we expected a delayed auditory cortex response for CI-SSD users as well. Finally, based on our previous study, we expected to find indications for multisensory processing, with different activity for AV-V compared to A for CI-SSD users, too.

3. Results

3.1. Behavioural results

Overall, all participants showed hit rates of $\geq 98\%$ in all conditions, and the mean RTs were between 504 ms and 638 ms (see Table 2). The 3 x 3 mixed ANOVA with condition (AV, A, V) as within-subject factor and group (CI-CHD, CI-SSD, NH) as between-subject factor revealed for RTs no main effect of group ($F_{2,30} = 1.06$; $p = .36$, $\eta^2 = 0.058$) and no group x condition interaction, but a main effect of condition ($F_{1,31.39.3} = 100.63$; $p \leq .001$, $\eta^2 = 0.31$). Follow-up post-hoc t-tests showed that RTs to redundant signals (AV) were significantly faster when compared to V ($t(32) = 12.4$; $p \leq .001$) or A ($t(32) = 20.3$; $p \leq .001$). There was no difference in RTs between the unisensory stimuli A and V ($t(32) = -0.84$; $p = 0.41$). These results are displayed in Fig. 1B.

For the hit rates, the 3 x 3 mixed ANOVA with condition (AV, A, V) as within-subject factor and group (CI-CHD, CI-SSD, NH) as between-subject factor showed no main effects or interactions (see Fig. 1C).

Concerning the race model, the one sample t-tests were significant in at least one decile for each group (see Table 3). This means that the likelihood of faster response times for redundant signals (AV) is higher than for those estimated by the race model (A + V). Fig. 1D displays the results of the race model. Overall, the violation of the race model in CI-CHD, CI-SSD users and NH listeners confirms the existence of multisensory integration in all tested groups.

For the other behavioural measures, we calculated one-way ANOVAs with subsequent t-tests to assess differences in *auditory word recognition ability* and *(visual) lip-reading abilities* between CI-CHD, CI-SSD users and

Table 2
Mean hit rates (in %) and mean response times (in ms).

Condition	Hit rates			Response times		
	NH	CI-CHD	CI-SSD	NH	CI-CHD	CI-SSD
A	99.0 ± 0.7	98.6 ± 1.1	98.1 ± 1.6	638 ± 84.1	623 ± 92.4	607 ± 79.4
	98.5 ± 1.2	98.0 ± 1.3	98.5 ± 1.3	638 ± 98.4	623 ± 96.1	602 ± 87.7
AV	98.7 ± 1.0	98.3 ± 1.3	99.1 ± 0.8	526 ± 81.9	530 ± 89.4	504 ± 74.0

Table 3

Redundant signals and modality-specific sum in each decile. AV is the redundant signals condition. A + V is the modality-specific sum. Paired-samples one-tailed t-tests were conducted for each group (with Bonferroni correction for multiple comparisons). An asterisk indicates a statistically significant result ($p \leq .05/5$).

Decile	NH			CI-CHD			CI-SSD		
	AV	A + V	p	AV	A + V	p	AV	A + V	p
.10	402	444	.000*	413	462	.000*	373	432	.000*
.20	437	480	.000*	460	494	.008*	409	461	.000*
.30	467	504	.000*	489	520	.014	435	484	.000*
.40	496	525	.003*	519	541	.056	456	450	.000*
.50	520	543	.011	547	560	.169	485	515	.000*

Table 4

Other behavioural measures for CI users and NH listeners. In the Freiburg monosyllabic word test and in the lip-reading test, a score of 100% means that all words have been repeated correctly. A higher value for the exertion rating means it was more effortful to perform the task (range: 6–20; 6 = no effort, 20 = highly effortful).

Group	Freiburg test (%)	Lip-reading test	Exertion rating
CI-CHD	70.9 ± 12.2	14.5 ± 6.7	12.5 ± 2.0
CI-SSD	71.4 ± 13.1	15.5 ± 4.4	12.2 ± 1.8
NH	98.2 ± 4.0	6.9 ± 4.7	11.9 ± 1.6

NH listeners. The ANOVAs showed a main effect of group for both the auditory word recognition ability ($F_{2,30} = 23.99$; $p < .001$, $\eta^2 = 0.615$) and (visual) lip-reading abilities ($F_{2,30} = 8.6$; $p = .001$, $\eta^2 = 0.364$). Follow-up t-tests revealed poorer speech recognition ability in the Freiburg monosyllabic test ($p \leq .001$), but better lip-reading skills for all CI users when compared to NH listeners (CI-CHD vs. NH: $p = .002$, CI-SSD vs. NH: $p = .004$). There was no difference between the two CI-user groups (CI-CHD vs. CI-SSD: $p = 0.6$). Concerning the *subjective exertion* measured during the EEG task, the ANOVA did not show a difference between the CI-CHD, the CI-SSD users and NH listeners

($F_{2,30} = 0.25$; $p = 0.78$, $\eta^2 = 0.016$). These results indicate that none of the tested groups perceived the task as more effortful than another group. The scores of these tests can be found in Table 4.

3.2. ERP results on the sensor level: GFP

In Fig. 2 the GFP of the grand averaged ERPs for the unisensory auditory (A) and the visually modulated auditory (AV-V) responses are shown for each group. Approximately between 120 and 140 ms, the first prominent peak is visible for all three groups. This peak fits into the time window of a N1 ERP. The next peak is around 240 ms and seems to be more prominent in NH listeners when compared to the two CI user groups. In the following, this peak is labelled as the P2 ERP. The GFP of the other conditions (V, AV) are also shown in the supplementary material (including the GMD between groups for each condition) to give an idea of the “raw,” non-difference wave data. First, we calculated a 3 x 2 mixed ANOVA with group (NH, CI-CHD, CI-SSD) as between-subject variable and condition (A, AV-V) as within-subject factor for the N1 GFP peak mean amplitude and the GFP peak latency. For the *N1 peak amplitude*, no statistically significant main effects or interactions were found. However, the ANOVA with *N1 latency* revealed a significant main effect of group ($F_{2,30} = 4.76$; $p \leq .05$, $\eta^2 = 0.18$). Follow-up t-tests

Global field power (GFP) and global map dissimilarity (GMD)

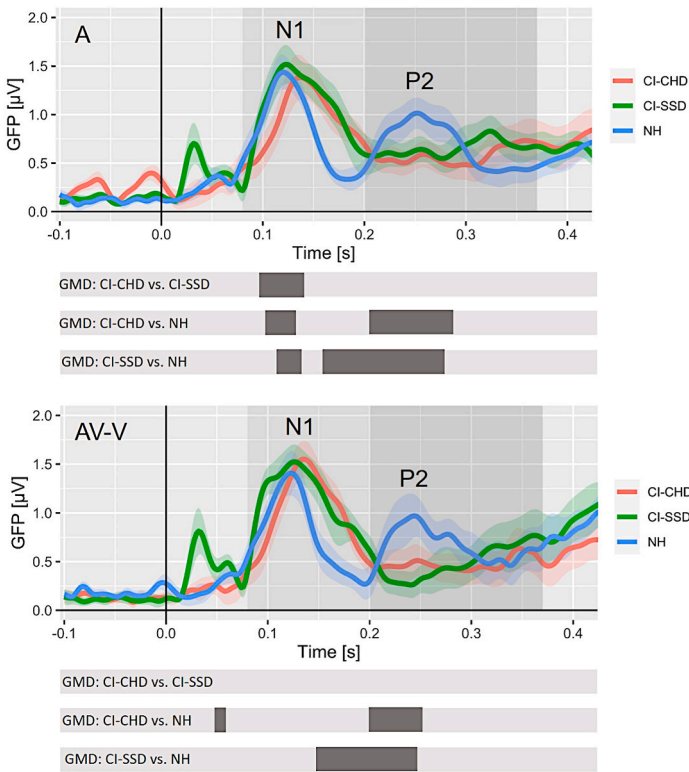


Fig. 2. ERP results on the sensor level. A) GFP of conditions A and AV-V for CI-CHD users (red), CI-SSD users (green) and NH listeners (blue), including standard error. It is important to note that the GFP only provides positive values because it represents the standard deviation across all electrodes separately for each time point. The ERP topographies at the GFP peaks (N1(A) = CI-CHD: 147 ms, CI-SSD: 136 ms, NH: 118 ms; N1(AV-V) = CI-CHD: 137 ms, CI-SSD: 135 ms, NH: 118 ms; P2(A) = CI-CHD: 305 ms, CI-SSD: 288 ms, NH: 256 ms; P2(AV-V) = CI-CHD: 284 ms, CI-SSD: 307 ms, NH: 245 ms) are given separately for each group (displayed on the right). The grey-shaded areas represent the N1 and P2 time windows for detecting peak and latency. The grey bars below representing the time window in which significant GMDs between the three groups were observed. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

showed a prolonged N1 latency for both the CI-CHD ($t(21) = 3.14$; $p \leq .05$) and the CI-SSD users ($t(21) = 3.83$; $p \leq .05$) compared to NH individuals. There was no significant difference between the two CI groups ($t(21) = -0.03$; $p = .97$).

We performed the same 3 x 2 mixed ANOVA for the P2 GFP peak mean amplitude and latency. For both the P2 peak amplitude and latency, there was no significant main effects or interactions.

3.3. ERP results on the sensor level: GMD

The GMD was analysed sample-by-sample to identify if and when ERP topographies significantly differ between conditions and groups. We compared CI-CHD with NH listeners (CI-CHD vs. NH), CI-SSD with NH listeners (CI-SSD vs. NH) and both CI groups (CI-CHD vs. CI-SSD) separately for each condition (A and AV-V). For the auditory condition (A), the results revealed topographic differences for all group comparisons within the time window of the N1 (CI-CHD vs. NH: 90–132 ms, CI-SSD vs. NH: 114–128 ms, CI-CHD vs. CI-SSD: 100–122 ms). Concerning the topographic differences within the P2 time window, there were no differences between the two groups of CI users, but we observed differences between the NH listeners and the two CI groups (CI-CHD vs. NH: 208–280 ms, CI-SSD vs. NH: 152–272 ms).

Regarding the GMD for the modulated condition (AV-V), there was a difference between NH listeners and CI-CHD at the N1 time window (NH vs. CI-CHD: 104–126 ms). Within the P2 time window, again there were no differences between the two CI groups, however there were differences between the NH listeners and each CI group (CI-CHD vs. NH: 204–264 ms, CI-SSD vs. NH: 154–242 ms). In addition, the GMD duration at the P2 time window was shorter for AV-V compared to A. The exact durations displaying differences between the groups are illustrated in Fig. 2 (grey bars beneath the GFP plots).

3.4. ERP results on the sensor level: Hierarchical clustering and single-subject fitting results

To better understand the underlying topographic differences (i.e. GMD) between the three groups, we conducted a hierarchical topographic clustering analysis by using the group-averaged data (CI-CHD (A); CI-CHD(AV-V); CI-SSD(A); CI-SSD(AV-V); NH(A); NH(AV-V)) in order to find template topographies within the N1 and P2 time windows. For that purpose, we chose a segment ranging from –100 ms to 470 ms (50–235 time frames). Specifically, we employed the atomize and agglomerate hierarchical clustering (AAHC) to identify the minimal amount of topographies that can explain the variance in our data set as best as possible. This method detected 17 template maps in 18 clusters that explained 88.08 % of all data. To be precise, we detected two maps within the N1 time window (map A and Map B) and three prominent maps within the P2 time window (Map C, Map D, Map E). With these template maps, we performed a single-subject analysis (Murray et al., 2008) to determine how well each of the template maps spatially correlated with the data from each participant. As the template Map B matches the topography from a conventional N1 peak (Fig. 2; Finke et al., 2016a; Sandmann et al., 2015), this template map will be referred to as N1 topography hereafter. Template Map C looks like a typical P2 topography (Fig. 2; Finke et al., 2016a; Schierholz et al., 2021) and therefore we will refer to this template map as the P2 topography. The template Map E is particularly prominent in the two CI-user groups and will be referred to as P2-like topography due to its similarity to the P2 topography (Fig. 2).

Dissimilarities within the topography across groups and conditions (see section ‘ERP results on the sensor level: GMD’) can be explained by a latency shift of the ERPs and/or by distinct neural generator configurations. To shed light on the origin of these differences, we analysed the first onset of maps and the map presence for the N1 and the P2 time windows. These results are presented in the following two subsections.

3.4.1. N1 time window

On the descriptive level, the CI-CHD users showed a map A and a Map B (= N1 topography) which were both present in the auditory-only condition (A; number of samples map A: 18.3 ± 18.6 ; Map B: 30.5 ± 19.0). Interestingly, specifically in the modulated condition (AV-V), the Map B (= N1 topography) was clearly more frequent compared to map A (number of samples 39.9 ± 18.2 (Map B) vs. 9.27 ± 16.6 (map A)). By contrast, both the NH listeners and the CI-SSD users showed a greater presence of Map B (= N1 topography) in general, irrespective of condition (A: number of samples: NH 40.8 ± 10.8 ; CI-SSD 48.0 ± 6.71 ; AV-V: number of samples: NH 36.9 ± 12.5 ; CI-SSD 47.5 ± 8.26).

To obtain an explanation for a potential ERP latency shift, we statistically analysed the first onset of maps by using a mixed-model ANOVA with group (NH, CI-CHD; CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as the within-subject factors for the N1 time window. The three-way mixed ANOVA revealed a significant group $\times$ map interaction ($F_{2,29} = 13.4$; $p \leq .001$, $\eta^2 = 0.39$). Follow-up t-tests showed that the onset of Map B (=N1 topography) was earlier in the NH listeners when compared with the CI-CHD ($t(20) = 3.82$; $p \leq .01$) and the CI-SSD ($t(21) = 6.75$; $p \leq .001$). There was no group difference in the onset of the N1 topography between CI-CHD and CI-SSD ($t(20) = -0.54$; $p = .6$). The results suggest that the N1 is generated later in CI users compared to NH individuals, regardless of the hearing threshold of the contralateral ear.

Second, we statistically analysed the number of time frames of the maps that showed the highest spatial correlations to the single-subject data, i.e. the map presence. This variable can provide an explanation for potentially distinct underlying neural generators between the three groups (CI-CHD, CI-SSD, NH) and the two conditions (A, AV-V). As above, we calculated a mixed-model ANOVA with group (NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as the within-subjects factors for the N1 time window.

For the N1 template maps, the ANOVA results showed a group $\times$ map $\times$ condition interaction ($F_{2,30} = 7.12$; $p \leq .005$, $\eta^2 = 0.04$). Post-hoc t-tests revealed for the CI-CHD users, but not for the CI-SSD users or NH listeners, that the presence of Map B (=N1 topography) was significantly enhanced for the modulated (AV-V) compared to the auditory-only (A) condition ($t(10) = -2.85$; $p \leq .05$). These results are illustrated in Fig. 3A. Given that template Map B corresponds to a conventional N1 topography, the results suggest that CI-CHD users in specific generate a N1 ERP map for the modulated response (AV-V) more frequently compared to the unisensory (A) condition. This visual modulation effect at the N1 latency was not observable for the NH listeners and the CI-SSD users.

Taken together, our results for the N1 on the first onset of maps and the map presence suggest that the observed topographic group differences at N1 latency can be explained by the following two reasons: 1) there are generally delayed cortical N1 ERPs in CI users, regardless of the condition (auditory-only or modulated response) and regardless of whether these patients have unilateral or bilateral hearing loss, and 2) there is a distinct pattern of ERP topographies specifically for the CI-CHD users compared to NH listeners and CI-SSD users. The visual modulation effect in the N1 topography was only observed for CI-CHD users, which suggests that this CI group in particular has a strong visual impact on auditory speech processing. By contrast, the visual impact in the CI-SSD users seems to be less pronounced and appears to be comparable to the NH listeners.

3.4.2. P2 time window

Similar to the analysis on the N1 time window, we analysed the first onset of maps by using a mixed-model ANOVA with group (NH, CI-CHD; CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as the within-subject factors for the P2 time window. The results did not reveal any significant main effects or interactions.

In a second step, we analysed the map presence for the P2 time window. As above, we calculated a mixed-model ANOVA with group

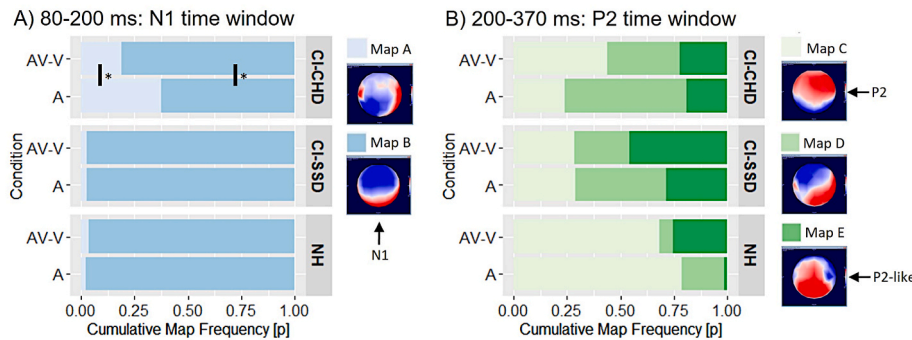


Fig. 3. Results from the hierarchical clustering and the single-subject fitting. A) Cumulative map frequency of the N1 maps: the CI-CHD users, but not the NH listeners or CI-SSD users, show a condition effect, with more frequent N1 map presence for AV-V compared to A. The corresponding map topographies are displayed on the right side, with Map B being referred to as the N1 topography. B) Cumulative map frequency of the P2 maps: there is a group effects (independent of the condition): NH listeners reveal a more frequent presence of a P2 topography (Map C) compared to CI-SSD users, and CI-SSD users show a more frequent presence of a P2-like topography (Map E) compared to NH listeners. CI-CHD users show a more frequent presence of Map D compared to NH listeners. Additionally, there is a condition effect (independent of the group): The

presence of the P2-like topography (Map E) is enhanced for AV-V compared to A. This suggests a visual modulation of auditory speech processing at P2 latency in all groups. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

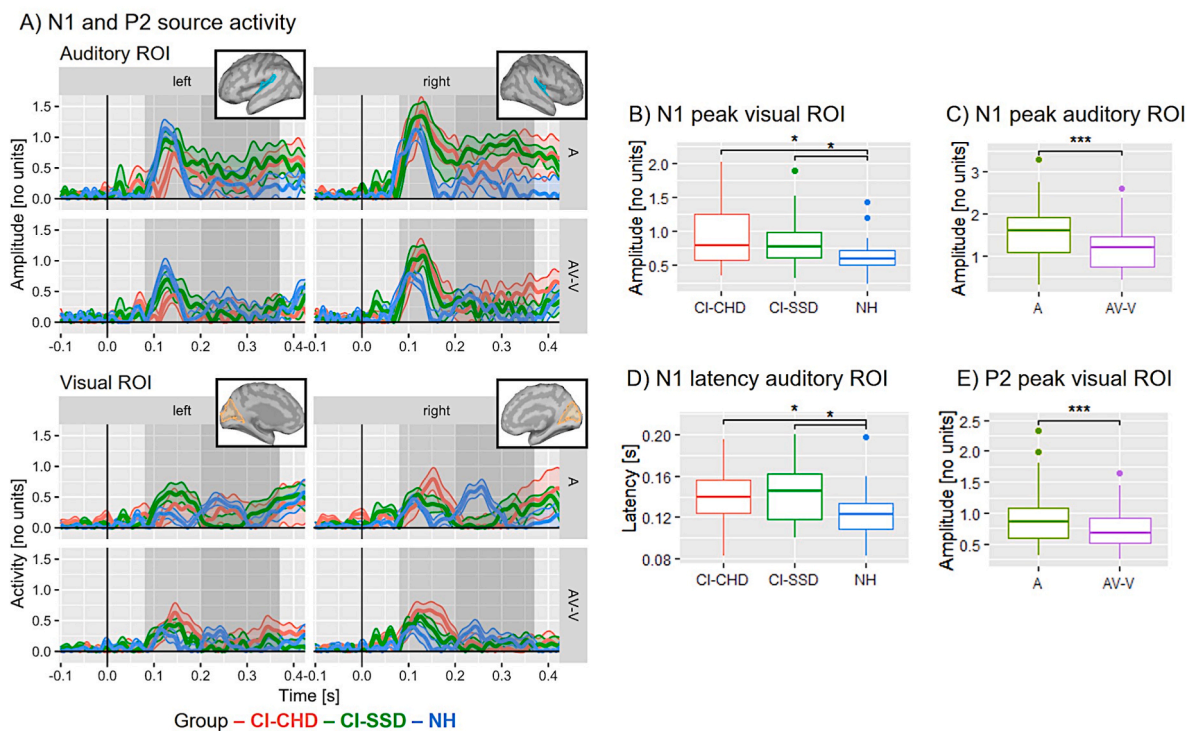


Fig. 4. ERP results on the source level. A) N1 and P2 source activity for CI users (red), CI-SSD users (green) and NH listeners (blue) separately for each ROI and each hemisphere with standard error (standard error shading was capped at zero). The source activity is displayed as absolute values with arbitrary units based on the normalisation within Brainstorm's dSPM algorithm. The grey shaded areas mark the N1 (light grey) and the P2 (dark grey) time windows. The boxes depict the location of the defined ROIs, with auditory ROIs in blue and visual ROIs in yellow. B) Group effect of the N1 peak mean in the visual cortex: both CI-CHD and CI-SSD users show more activity in the visual cortex compared to NH listeners, regardless of condition. C) Condition effect of the N1 peak mean in the auditory cortex: there is a significantly reduced auditory-cortex activation for AV-V compared to A, indicating multisensory interactions in all groups. D) N1 latency effect in the auditory cortex: Both CI and CI-SSD users show a prolonged N1 latency compared to NH listeners in the auditory cortex, regardless of the condition. This suggests a delayed auditory-cortex activation in CI users, independent of the hearing threshold in the contralateral ear. E) P2 condition effect in the visual cortex: there is a significantly reduced visual-cortex activation for AV-V compared to A, pointing towards multisensory interactions in all groups. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

(NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) and template map as the within-subjects factors for the P2 time window. The three-way mixed ANOVA showed a significant group $\times$ map ($F_{4,60} = 3.47$; $p \leq .05$, $\eta^2 = 0.12$) and a condition $\times$ map ($F_{2,60} = 3.41$; $p \leq .05$, $\eta^2 = 0.03$) interaction. For the group $\times$ condition interaction, follow-up t-tests revealed for the NH listeners a significantly enhanced presence of Map C (= P2 topography) compared to CI-SSD users ($t(21) = -3.18$; $p \leq .01$). Vice versa, CI-SSD users showed a significantly enhanced presence of Map E (= P2-like topography)

compared to NH listeners ($t(21) = 3.65$; $p \leq .01$), regardless of the condition. Finally, for Map D, there was a significant difference between CI-CHD users and NH individuals ($t(21) = 3.46$; $p \leq .01$), with CI-CHD users showing a more dominant presence of this map compared to NH controls. These results are shown in Fig. 3B. Following the condition $\times$ map interaction, follow-up t-tests revealed significant differences between A and AV-V only for Map E ($t(32) = -2.5$; $p \leq .01$). This result suggests that a P2-like topography (Map E) is generated more often for modulated responses (AV-V) compared to unmodulated responses (A),

which is shown in Fig. 3B.

In sum, our results about the *first onset of maps* and the *map presence* at *P2 latency* suggest group-specific topographic differences at P2 latency, with a stronger presence of a conventional P2 topography (Map C) in NH listeners compared to CI-SSD users and a stronger presence of the P2-like topography (Map E) in CI-SSD users compared to NH listeners. Together with the observation that Map D is more present in CI-CHD users than in NH listeners, these results confirm our GMD results, showing a significant group difference between NH listeners and the two CI-user groups for both conditions (A, AV-V) at P2 latency (Fig. 2). Finally, all groups show a P2-like topography (Map E) that is more frequent in the modulated than in the auditory-only condition, which points to alterations in the cortical processing at P2 latency due to the additional visual information in the speech signal.

3.5. Results from ERP source analysis

We conducted a source analysis to further analyse the differences between the three groups, focusing on the auditory and visual cortex activity in both hemispheres. Single-subject source activities for each ROI, condition and group were exported from Brainstorm and were statistically analysed. The source waveforms for the N1 and the P2 are illustrated in Fig. 4A, showing the response in the auditory cortex (N1 peak latency mean: CI-CHD = 141 ms $\pm$ 27 ms; CI-SSD = 143 ms $\pm$ 28 ms; NH = 122 ms $\pm$ 22 ms) and in the visual cortex (N1 peak latency mean: CI-CHD = 143 ms $\pm$ 29 ms; CI-SSD = 136 ms $\pm$ 28 ms; NH = 136 ms $\pm$ 35 ms) for all groups. The peak mean amplitudes and latencies were the dependent variables for the following ANOVA. We performed a mixed-model ANOVA with group (NH, CI-CHD, CI-SSD) as the between-subjects factor and condition (A, AV-V) and hemisphere (left, right) as the within-subject factors for each time window of interest (N1, P2) and each ROI (auditory, visual) separately.

Concerning the N1 *peak mean* in the *visual cortex*, the mixed-model ANOVA showed a significant main effect of group ($F_{2,30} = 4.39$; $p \leq .05$, $\eta^2 = 0.15$). Post-hoc t-tests confirmed a significant difference between the NH listeners and both CI groups (NH vs. CI-CHD: $t(21) = 3.01$; $p \leq .05$; NH vs. CI-SSD: $t(21) = 3.18$; $p \leq .05$), but no difference between the two CI groups (CI-CHD vs. CI-SSD: $t(21) = -0.24$; $p = 0.81$). Thus, both CI user groups showed more recruitment of the visual cortex compared to NH listeners, regardless of hemisphere and condition (see Fig. 4B).

For the N1 *peak mean* in the *auditory cortex*, the mixed-model ANOVA revealed a significant main effect of hemisphere ($F_{1,29} = 9.46$; $p \leq .005$, $\eta^2 = 0.11$) and a significant main effect of condition ($F_{1,29} = 34.25$; $p \leq .001$, $\eta^2 = 0.11$). Resolving the main effect of hemisphere, follow-up t-tests showed a greater amplitude for the right hemisphere compared to the left hemisphere ($t(63) = -3.66$; $p \leq .001$), regardless of group and condition. Following the main effect of condition, the subsequent t-tests revealed reduced amplitudes for AV-V compared to A ($t(64) = 5.95$; $p \leq .001$), regardless of hemisphere and group, which points to multisensory interaction processes (see Fig. 4C).

For the N1 *peak latency* in the *auditory cortex*, the mixed-model ANOVA identified a significant main effect of group ($F_{2,29} = 4.31$; $p \leq .05$, $\eta^2 = 0.13$) and a significant main effect of hemisphere ($F_{1,29} = 5.68$; $p \leq .05$, $\eta^2 = 0.02$). Following the main effect of hemisphere, the post-hoc t-test revealed a significant difference between the left and the right auditory cortex ($t(63) = 2.20$; $p \leq .05$) with the right hemisphere showing faster latencies compared to the left hemisphere. Resolving the main effect of group, follow-up t-tests revealed a significantly shorter latency of the auditory-cortex response in the NH listeners compared to both CI groups (NH vs. CI-CHD: $t(21) = 5.64$; $p \leq .001$; NH vs. CI-SSD: $t(21) = 6.01$; $p \leq .001$), but no difference between the two CI groups (CI-CHD vs. CI-SSD: $t(21) = -0.20$; $p = 0.80$). Hence, both CI user groups showed a delayed auditory-cortex response compared to NH listeners, regardless of hemisphere and condition (see Fig. 4C). For the

N1 *peak latency* in the *visual cortex*, the mixed-model ANOVA did not show any significant main effects or interactions.

Concerning the P2 *peak mean* in the *auditory cortex*, the mixed-model ANOVA found a significant main effect of condition ($F_{1,29} = 11.25$; $p \leq .01$, $\eta^2 = 0.04$). Resolving this main effect, the post-hoc t-tests revealed a significant difference between A and AV-V ($t(64) = 2.9$; $p \leq .005$), with A showing greater amplitudes than AV-V, regardless of group and hemisphere. This points to multisensory interaction processes in the auditory cortex at P2 latency.

For the P2 *peak mean* in the *visual cortex*, the mixed-model ANOVA found a significant main effect of condition ($F_{1,30} = 17.06$; $p \leq .001$, $\eta^2 = 0.06$) as well. Follow-up t-tests revealed a significant difference between A and AV-V ($t(65) = 3.92$; $p \leq .001$), with A showing greater amplitudes than AV-V, regardless of group and hemisphere. This points to multisensory interaction processes in the visual cortex at P2 latency as well (see Fig. 4E).

Regarding the P2 *peak latency* in the *auditory and visual cortices*, the mixed-model ANOVA found neither significant main effects nor significant interactions.

3.6. Correlations

We performed correlations for each CI user group (CI-SSD and CI-CHD), using the Pearson's correlation and the Benjamini-Hochberg (BH) procedure to control for multiple comparisons (Benjamini and Hochberg, 1995). First, we wanted to check whether lip-reading abilities are related to the CI experience and the age at onset of hearing loss (Stropahl et al., 2015; Stropahl and Debener, 2017; Layer et al., 2022). The results revealed a trend for a positive relationship between lip-reading abilities and CI experience (CI-CHD: $r = 0.51$; $p = 0.038$; p corrected = 0.076; CI-SSD: $r = 0.47$; $p = 0.12$; p corrected = 0.147) and a negative relationship between lip-reading abilities and the age of onset of hearing loss for both CI user groups (CI-CHD: $r = -0.84$; $p = 0.001$; p corrected = 0.005; CI-SSD: $r = -0.69$; $p = 0.01$; p corrected = 0.027). Thus, for both CI user groups it holds that the earlier the onset of hearing impairments, the more pronounced are the lip-reading abilities. Moreover, we aimed to reproduce the relationship between CI experience and the activation in the visual cortex (Giraud et al., 2001c; Layer et al., 2022). The results did not reach a significance level (CI-CHD: $r = 0.51$; $p = 0.1$; p corrected = 0.16; CI-SSD: $r = 0.34$; $p = 0.29$; p corrected = 0.29).

4. Discussion

In this follow-up study, we used behavioural and EEG measures to investigate audiovisual interactions in CI users with unilateral (CI-SSD) and bilateral (CI-CHD) hearing loss and in a group of NH controls. This study was conducted to extend the results from our previous study comparing CI-CHD users with NH listeners (Layer et al., 2022), by including a third group of participants; namely the CI-SSD users. A subset of our previously reported data was reused and compared to the additional group of CI-SSD users. The inclusion of the additional group of patients significantly extends our previous study because it not only examines the transferability of our previous findings to different CI patient groups, but also provides valuable insights into the influence of individual factors - specifically the hearing ability of the second ear - on audiovisual speech processing in CI users.

At the behavioural level, we confirmed multisensory interactions for all three groups, as evidenced by the shortened response times for the audiovisual condition compared to each of the two unisensory conditions (Fig. 1B) and by the violation of the race model (Fig. 1E). This was in line with the ERP analyses, confirming a multisensory effect for all groups by exhibiting a reduced activation in the auditory and visual cortex for the modulated (AV-V) response compared to the auditory-only (A) response at both the N1 and P2 latencies (Fig. 4C and E,

respectively). In addition to this multisensory effect across all groups, we found group-specific differences. First, specifically the group of CI-CHD users, showed a change of N1 voltage topographies when additional visual information accompanied the auditory information (Fig. 3A), which suggests a particularly strong visual impact on auditory speech processing in CI users with bilateral hearing loss. Second, both groups of CI users revealed a delayed auditory-cortex activation (Fig. 4D), enhanced lip-reading abilities (Fig. 1D) and stronger visual-cortex activation (Fig. 4B) when compared to the NH controls. Thus, the current results extend the results of our previous study (Layer et al., 2022) by showing distinct multisensory processes not only between NH listeners and CI users in general, but even between CI users with unilateral (CI-SSD) and bilateral (CI-CHD) hearing loss.

4.1. Behavioural multisensory integration in all groups

The behavioural results revealed that both the NH listeners and the two CI user groups had faster reaction times for audiovisual syllables than for unisensory (auditory-alone, visual-alone) syllables (Fig. 1; Table 2). No difference was found between the auditory and visual conditions. Hence, all groups exhibited a clear redundant signals effect for audiovisual syllables, implying that the benefit of cross-modal input is comparable between the CI user groups and NH listeners on a behavioural level (Laurienti et al., 2004; Schierholz et al., 2015; Layer et al., 2022), at least when considering syllables that are combined with a talking head. The violation of the race model for each group (CI-CHD, CI-SSD, NH) suggests that multisensory integration was the cause for the observed redundant signals effect in both CI user groups and NH listeners. However, the behavioural responses of the CI users were not slower compared to the NH listeners, even though the signal provided by the CI is known to be limited in comparison to a natural hearing experience (Drennan and Rubinstein, 2008). The observation of comparable response times in CI users can be explained by the fact that there were only two syllables, and that the difficulty of the task was correspondingly low. Compatible with this, all groups were equally able to perform the task, and the subjective rating of the listening effort showed no difference between the groups.

One would assume that the CI users might be better and faster at identifying the purely visual syllables due to results from previous studies with congenitally deaf individuals and CI users, showing visual enhancements, in particular visually induced activation in the auditory cortex (Bottari et al., 2014; Finney et al., 2003; Hauthal et al., 2014; Bavelier and Neville, 2002; Heimler et al., 2014; Sandmann et al., 2012). This cross-modal activation seems to be driven by auditory deprivation and might form the neural basis for specific superior visual abilities (Lomber et al., 2010). Importantly, auditory impairment is not only experienced in CI users before receiving a CI, but also after the implantation when only a limited auditory input is provided by the CI. Thus, it is not surprising that CI users reveal compensatory visual strategies, such as enhanced lip-reading abilities, in order to overcome the limited CI signal (Rouger et al., 2007; Schreitmüller et al., 2018; Stropahl et al., 2015; Stropahl and Debener, 2017). Our results extend previous observations of enhanced visual abilities in CI users by showing that not only CI-CHD users, but surprisingly also CI-SSD users demonstrate a better lip-reading ability when compared to NH listeners. Importantly, the lip-reading ability was comparable between the two patient groups, and both groups showed a positive correlation with the age of the onset of hearing loss, indicating that an earlier onset of hearing loss triggers improved behavioural visual abilities. Our results demonstrate that this visual improvement develops across different groups of CI patients, independent of the hearing abilities of the contralateral ear.

However, behavioural visual improvements in CI users seem to be stimulus- and task-specific, as indicated by our finding that the two CI user groups showed comparable behavioural results to NH listeners in the speeded response task. Our finding is consistent with previous

studies, using a speeded response task with simple tones and white discs as auditory and visual stimuli, respectively (Schierholz et al., 2015, 2017). It seems that in our study the task with the basic stimuli and the two syllables was too easy, leading to ceiling effects in all groups. This estimation is in line with our observation that the perceived exertion effort was comparable between all three groups. Importantly, behavioural group differences have well been reported in a previous study using more complex stimuli presented in the context of a difficult recognition paradigm, showing an enhanced audiovisual gain in CI users when compared to NH listeners (Radecke et al., 2022). This is consistent with the view that behavioural advantages due to additional visual information in CI users are task- and stimulus-selective, and that they become evident under specific circumstances, for instance in conditions with semantic information (Moody-Antonio et al., 2005; Rouger et al., 2008; Tremblay et al., 2010; Radecke et al., 2022). Thus, future studies should use linguistically complex stimuli, such as words or sentences presented in auditory, visual and audiovisual conditions, in order to better understand the behavioural advantages for visual and audiovisual speech conditions in CI users compared to NH individuals.

4.2. Electrophysiological correlates of multisensory speech perception

Similar to the behavioural data, we also discovered commonalities among groups at the ERP level. Nonetheless, group differences were found as well, which will be discussed in the following sections.

4.2.1. Group similarities in multisensory speech processing

Similar to the behavioural results, we found evidence for multisensory effects in the ERP responses. The topographic clustering analysis with subsequent single-subject fitting confirmed multisensory interactions for both CI user groups and NH individuals by revealing an increase in *P2-like topographies* for the modulated ERPs (AV-V) compared to the purely auditory condition (A). This observation points to a visual modulation of the auditory ERPs in the two CI user groups as well as in the NH individuals.

These findings are supported by the source analysis. Specifically, we investigated the visual impact on the auditory cortical response by comparing the ERPs of the auditory condition (A) with the visually modulated ERPs (AV-V). A difference between these two conditions can be seen as evidence for non-linear multisensory interactions (Besle et al., 2004; Murray et al., 2005; Cappe et al., 2010; Foxe et al., 2000; Molholm et al., 2002). As previous studies have shown, a difference between two conditions can either be superadditive ($A < AV - V$; e.g. Schierholz et al., 2015) or subadditive ($A > AV - V$; e.g. Cappe et al., 2010) (Stevenson et al., 2014). Our source analyses confirmed the multisensory interactions in all groups, both within the N1 and the P2 time windows by demonstrating a subadditive effect, indicated by a reduced activation in the auditory cortex for the modulated (AV-V) compared to the auditory-only (A) condition. A similar subadditive effect was even observed for the visual cortex at P2 latency in all groups. In sum, these findings are highly consistent with the results from other ERP studies (e.g. Cappe et al., 2010) as well as fMRI studies (e.g. Martuzzi et al., 2007), showing multisensory interactions in both the auditory and visual cortex and confirming the behavioural results of multisensory integration for all groups.

4.2.2. Group differences in multisensory speech processing

As measured by global field power (GFP), the signal strength of cortical responses did not differ between the two CI user groups and the NH listeners. Nevertheless, for the response topography, as quantified by the global map dissimilarity (GMD), we detected various differences between CI-CHD, CI-SSD and NH listeners for both the auditory-only condition (A) and the modulated response (AV-V). Thus, group differences were not caused by signal strength but rather by differences in their electric field topographies and by extension the configuration of active brain networks. To ascertain whether distinct topographies were

caused by a simple latency shift of the ERP responses and/or by variable configurations of the neural sources, we conducted various analyses, in particular hierarchical clustering, single-subject fitting and source analysis.

Within the N1 time window, hierarchical clustering and single-subject fitting showed specifically for the CI-CHD users – but not for CI-SSD users and not for NH listeners – a greater presence of N1 topographies (Map B) for modulated responses (AV-V) compared to purely auditory responses (A). The difference in N1 topography between CI-CHD users and NH listeners was already reported in our previous study (Layer et al., 2022), which was now expanded by including a group of CI-SSD users. Interestingly, the NH listeners and CI-SSD users showed no differences in map presence between auditory and modulated responses, indicating that these two groups use similar neural processes to evaluate the audiovisual speech stimuli. An increasing presence of a conventional N1 topography (Map B) specifically for the modulated condition indicates alterations in audiovisual processing and a multisensory benefit for CI-CHD users, when additional visual information is present. This modulation in CI-CHD users reflects most likely a strategy that they develop to compensate for the limited CI input. Interestingly, this modulation was not detected in CI-SSD users, which leads to the assumption that CI-CHD have a higher benefit from additional visual input compared to CI-SSD users. This group-specific effect may be due to the fact that CI-SSD users have a NH ear on the contralateral side which might serve as the main communication channel. Therefore, it is not surprising that CI-SSD users and NH listeners show comparable topographic results, at least within the N1 latency range. Previous research comparing different groups of CI users is limited. Nevertheless, first evidence of differences in speech-in-noise performance between CI-SSD users and bimodal CI users (CI on one ear and hearing aid on the contralateral ear) was reported (Williges et al., 2019). On the other hand, differences between CI-SSD users and bilateral CI users (both ears fitted with CIs) were observed in situations with multiple concurrent speakers (Bernstein et al., 2016). As a result, we expected group differences to emerge not only for auditory stimulation but also for audiovisual stimulation. However, as far as we are aware, this has not been investigated yet. Our results therefore indeed confirm first indications of different processing strategies among different CI user groups.

Given that our results are restricted to the processing of syllables, we propose that further studies should use more complex speech stimuli in order to refine and deepen the current findings about different groups of CI users. Similar to the current analyses, future studies should apply electrical neuroimaging (Michel et al., 2009), including topographic and ERP source analysis, given that it is a powerful approach to investigate multisensory interactions (Stevenson et al., 2014). It is likely that the testing of CI-CHD in more complex speech conditions, in particular with semantic information, results in even more enhanced cortical audiovisual interactions (Radecke et al., 2022), as can be assumed based on previous behavioural results (Moody-Antonio et al., 2005; Rouger et al., 2008; Tremblay et al., 2010). We speculate that in these demanding stimulus conditions, an increase in audiovisual interactions may become even detectable in the group CI-SSD users.

Apart from this modulation effect, which was specific for CI-CHD users, we found generally delayed cortical responses for both CI user groups compared to NH listeners, which was consistently reflected in various analyses. On the sensor level, ERP data revealed that both CI-CHD users and CI-SSD users had a prolonged N1 latency compared to NH listeners for both the modulated and the auditory-only responses. By comparing the first onset of Map B (N1 topography) between the three groups of participants, hierarchical clustering and single-subject fitting analyses confirmed the delayed onset of the N1 topography in the two groups of CI users. Similarly, the source analysis revealed a delayed N1 response in the auditory cortex for both CI-CHD and CI-SSD users compared to NH listeners. In fact, since the signal does not have to pass through the middle and inner ear due to direct stimulation of auditory nerve fibres, one might assume that the time until the electrical signal

reaches the auditory nerve with a CI is faster than in NH listeners. Interestingly, Schierholz et al. (2017) discovered delayed cortical responses not only in patients with a CI, but also in patients with an auditory brainstem implant (ABI; direct stimulation of the cochlear nucleus) and an auditory midbrain implant (AMI; direct stimulation of the inferior colliculus). Despite the fact that central auditory implants bypass more structures than CIs, ABI and AMI patients showed even more delayed cortical and poorer behavioural responses when compared to CI patients, appearing to be specifically related to the insufficient input provided by central auditory implants (Schierholz et al., 2017). Analogously, delayed N1 ERP responses in CI users are likely to reflect difficulties in processing speech sounds with the CI compared to natural hearing. In line with this, previous research with NH listeners has shown that difficult acoustic listening conditions, such as speech in background noise, cause a delay in the N1 response (Billings et al., 2011; Finke et al., 2016a). Furthermore, our findings of slowed cortical N1 ERPs in CI users are consistent with the findings of several other studies which used auditory stimuli of varying acoustic complexity, and which showed delayed N1 latency and poorer auditory discrimination ability in CI-CHD users compared to NH listeners (Beynon et al., 2005; Finke et al., 2016a; Sandmann et al., 2009, 2015; Senkowski et al., 2014). This is also an observation that has been reported in studies with CI-SSD users comparing the cortical responses from the CI ear and the NH ear (Finke et al., 2016b; Weglage et al., 2022).

To sum up, the results on the N1 ERPs revealed several group differences. First, the hierarchical clustering and single-subject fitting analyses showed a *condition effect* specifically for the group of CI-CHD users, showing distinct patterns of N1 topography between the modulated auditory response (AV-V) and the auditory-only response (A; Fig. 3A). However, for both the CI-SSD users and the NH listeners, such a modulation was not given, which leads to the conclusion that these two groups do not benefit from additional visual input to the same extent as CI-CHD users. CI-CHD users seem to have a particularly strong visual impact on auditory speech processing, allowing these individuals to compensate the limited CI input. By contrast, CI-SSD users appear to be less dependent on additional visual input, probably due to the fact that they can rely on the contralateral ear, which is normal-hearing.

Apart from the CI-CHD-specific condition effect, however, our results revealed a *map dissimilarity at N1 latency* between the three groups, at least for the auditory condition (Fig. 2A). The results from the topographic and source analyses suggest that these group differences can at least partially be explained by a delayed N1 auditory-cortex response in both CI-CHD and CI-SSD users when compared to NH listeners. However, as discussed in the following section (4.2.3), the results from the source analyses indicate that this map dissimilarity at N1 latency is also caused by a different configuration of neural sources. Both CI user groups showed an additional activation in the visual cortex compared to NH listeners (see section 4.2.3 for more details).

Within the P2 time window, the hierarchical clustering and single-subject fitting analyses showed a condition effect (independent of group; Fig. 3B) and some group effects (independent of condition; Fig. 3B). The *condition effect*, which revealed an enhanced map presence of the P2-like topography (Map E) for the modulated (AV-V) compared to the auditory response (A), indicates a visual modulation of the auditory ERPs in all groups of participants. Regarding the *group effects*, we found a greater map presence of the P2 topography (Map C) for NH individuals compared to CI-SSD users, and a greater map presence of a P2-like topography (Map E) for CI-SSD users compared to NH listeners. This difference is in line with the results from the GMD observations, pointing to a difference in response topography at the P2 time window (Fig. 2). At the same latency window, we also found a difference between CI-CHD users and NH listeners for the Map D, showing a greater presence of this topography in CI-CHD users than NH listeners. This difference was also confirmed by the GMD analysis, showing a difference in response topography at the P2 time window (Fig. 2). Importantly, no significant differences between CI-CHD users and CI-SSD users

were detected in the topographic analysis at P2 latency, which corresponds to the lack of differences in the GMD analysis. These results however are not easy to interpret. One may speculate that the two CI user groups do not differ statistically at P2 latency but exhibit distinct topographic patterns when compared to NH listeners. Therefore, a source analysis is unlikely to yield clear results revealing group-specific cortical activation patterns. Indeed, the source analysis revealed no group differences in auditory or visual cortex activation. Nevertheless, the source analysis provided evidence for audiovisual interactions by identifying reduced amplitudes for the modulated condition (AV-V) compared to the purely auditory condition (A), independent of the groups (see section 4.1). Taken together, we conclude that CI users and NH listeners recruit the auditory and visual cortices similarly during auditory and audiovisual speech processing, at least at the P2 latency range. Future research is required to determine if group-specific differences within the P2 time window emerge with other stimuli/tasks and whether they appear beyond the visual and auditory sensory systems, as in frontal or other temporal areas (Campbell and Sharma, 2013; Giraud et al., 2001a).

4.2.3. Cross-modal plasticity in the visual cortex

The results from the source analysis within the N1 time window revealed a greater recruitment of the visual cortex for both the CI-CHD users and CI-SSD users when compared to NH listeners. This cross-modal activation of the visual cortex is in line with previous research findings from CI-CHD users, reporting that cortical alterations are not limited to the auditory cortex (Campbell and Sharma, 2013, 2016; Chen et al., 2016; Giraud et al., 2001a,b) but seem to extend to the visual and even inferior frontal areas (Rouger et al., 2012). As far as we are aware, this is the first study to show that auditory-induced activation in the visual cortex is not restricted to CI-CHD users, as shown in our previous study (Layer et al., 2022) but is also present in CI-SSD users.

It has been suggested that the additional recruitment of the visual cortex in CI users is a way to compensate the limited auditory input delivered through the CI (Doucet et al., 2006; Giraud et al., 2001c; Strelnikov et al., 2010, 2013). In the PET (positron emission tomography) study of Giraud and colleagues (Giraud et al., 2001c), a greater auditory activation in the visual cortex was reported for unilaterally-implanted CI users (with bilateral hearing loss) in comparison to NH controls for meaningful sounds. The authors found an association between this visual recruitment with longer CI experience and pronounced lip-reading abilities (Giraud et al., 2001c). By now, cross-modal recruitment of the visual cortex for processing auditory stimuli has been observed in CI users for syllables, words, environmental sounds, pure tones, and reversed words (Chen et al., 2016; Giraud et al., 2001b; Giraud et al., 2001c). This visual cortex activation when processing auditory stimuli may represent an enhanced synergy of the auditory and visual modalities in CI users. This might be beneficial for speech comprehension recovery following cochlear implantation, when CI users must learn to match novel auditory speech inputs with corresponding visual speech cues (Strelnikov et al., 2013). Consistent with this view, our previous study showed a positive correlation between CI experience and visual-cortex activation, indicating that CI users increasingly recruit the visual cortex with increasing duration of CI usage (Layer et al., 2022).

The present study extends previous findings by showing visual recruitment for auditory (and audiovisual) syllables not only in CI-CHD users but also in CI-SSD users. This observation in CI-SSD users may appear surprising because these individuals have at least one ear on the contralateral side on which to rely. However, in natural conditions with bilateral stimulation, the CI-SSD users' auditory performance is still impaired compared to normal hearing patients (Dorman et al., 2015; Ludwig et al., 2021), particularly in difficult listening conditions, which may explain why these individuals develop compensatory strategies, as indicated by enhanced lip-reading skills and cross-modal recruitment of the visual cortex. Taken together, our results suggest that

auditory-induced activation of the visual cortex is independent of the hearing abilities on the contralateral ear when being stimulated over the CI ear. However, future studies are required to systematically compare the stimulation of the CI ear alone, the contralateral ear alone and both ears together. The measured cortical activation differences would provide a clearer picture of how much of the large variability in speech recognition ability in different CI users can be attributed to the contralateral ear. Further, we suggest that future studies should further split the CI groups into bimodal (CI – HA), bilateral (CI–CI) and SSD (CI–NH) CI users and use more ecologically valid stimuli such as complex speech stimuli, and different task conditions (e.g. passive vs. active) in order to deliver a portrait of the characteristics for each group in terms of audiovisual speech processing and cortical reorganisation. This could allow for developing methods for accelerating and improving auditory rehabilitation after implantation that are specific for each CI user group.

5. Limitations

An important limitation of this study is the small sample size of $n = 11$ per group. A small sample size is common in the scarce literature on CI-SSD users, because this patient group was not implanted a few years back due to the intact NH ear. Only after proving the benefits of an implantation (improved hearing abilities compared to other techniques such as bone-anchored hearing aids (BAHA) or contralateral routing of signal (CROS) (Arndt et al., 2011, 2017); better speech intelligibility, sound localisation and quality of life (Kitterick et al., 2015)), CI-SSD users were increasingly fitted with a CI. Consequently, results with small sample sizes always must be interpreted with caution. Future studies, however, should aim to increase the number of participants in order to improve statistical power and apply more extensive statistical correction methods, as suggested by Cramer et al. (2016).

One might ask the question whether the presented results are transferrable to other audiovisual stimuli or whether these are restricted to language-specific stimuli. We assume that the significant audiovisual benefit observed in our CI users was strongly driven by the linguistic property of the presented stimuli, and we hypothesise that this audiovisual benefit may even be more pronounced for more difficult linguistic stimuli (words or sentences). Nevertheless, when solely focusing on the additional recruitment of the visual cortex, as observed in the present and our recent results (Layer et al., 2022), previous studies have suggested that this effect is not restricted to purely linguistic tasks. This effect has been shown for syllables, words and environmental sounds (Giraud et al., 2001c). Another study (Chen et al., 2016) extended these results to pure tones and reversed words, showing that intelligibility is not necessary for eliciting auditory-induced recruitment of the visual cortex. However, these previous studies do not provide answers concerning processing differences and similarities between CI user groups, which is the novelty of this study. We hypothesise that further and more pronounced differences between CI-CHD users and CI-SSD users will become evident for more difficult linguistic stimuli (words/sentences vs. syllables) and more difficult task conditions (semantic processing vs. discrimination of syllables). Importantly, future studies should use the same experimental settings (i.e. the same paradigm) for different types of stimuli. They should compare cortical response patterns between non-linguistic stimuli (e.g. basic and environmental sounds) and linguistic stimuli (e.g., syllables and words), to see if our findings are transferable to both non-linguistic and more difficult linguistic stimulus conditions within the same patients.

6. Conclusion

The present study applied electrical neuroimaging, including topographic and source analysis, to investigate whether the cortical processing of audiovisual syllables is different between CI-CHD users and CI-SSD users. These two CI user groups were also compared to NH listeners. This study expands on our previous study that compared CI-CHD

users to NH listeners by reusing these data to include an additional group of CI users. Our findings showed a clear multisensory effect in both CI user groups and NH listeners, as evidenced by faster behavioural response times and reduced auditory-cortex activation at N1 and P2 latencies. Furthermore, we discovered a condition effect for the N1 ERP topography in the group of CI-CHD users only, demonstrating a distinct pattern of N1 topography between the auditory-only (A) and the modulated auditory response (AV-V). This finding indicates a strong visual influence on auditory speech processing especially in CI-CHD users, allowing these patients to compensate for the limited CI input. Finally, we discovered that both CI user groups showed a delay in the auditory-cortex response at N1 latency, indicating difficulties in cortical processing of the CI's limited signal. Nonetheless, we found that both CI user groups showed pronounced lip-reading abilities and an additional recruitment of the visual cortex compared to NH listeners. This finding extends previous results by demonstrating that this cross-modal visual activation is also present in CI-SSD users, who have an intact NH ear on the contralateral side. One may speculate that these cortical alterations allow the CI users to combine multisensory information, to refine expectations and to sharpen perception. We conclude that an auditory-induced activation in the visual cortex is independent of the hearing threshold and supply on the contralateral ear when being stimulated via the CI ear. This auditory-induced activation in the visual cortex is an important and insightful similarity between the two CI user groups. Overall, these results confirm existing differences in multisensory processing both when comparing NH listeners and CI users and when comparing different groups of CI users. This emphasises the importance of developing individual rehabilitation methods tailored to different groups of CI users.

Funding

This work was supported by the German Research Foundation (Deutsche Forschungsgemeinschaft; DFG) [SA 3615/1-1 (project number: 415896102) and SA 3615/2-1 (project number: 416867313)] and by the Köln Fortune Program/Faculty of Medicine, University of Cologne [278/2020] to PS. MMM is supported by the Swiss National Science Foundation (169206).

CRedit authorship contribution statement

Natalie Layer: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization, Project administration. **Anna Weglage:** Software, Validation, Formal analysis, Investigation, Writing – review & editing. **Verena Müller:** Conceptualization, Writing – review & editing. **Hartmut Meister:** Conceptualization, Software, Resources, Writing – review & editing. **Ruth Lang-Roth:** Conceptualization, Writing – review & editing. **Martin Walger:** Conceptualization, Writing – review & editing. **Micah M. Murray:** Software, Formal analysis, Validation, Writing – review & editing. **Pascale Sandmann:** Conceptualization, Methodology, Validation, Resources, Writing – original draft, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the Deutsche Forschungsgemeinschaft (DFG) [SA 3615/1-1 (project number: 415896102) and SA 3615/2-1 (project number: 416867313)] and Köln Fortune [278/2020] for providing funding for this research. Also, we thank Pauline Burkhardt and Lina Wiesel for their

help in data collection and Moritz Layer for his support with the implementation of the analysis. Finally, we thank all participants for taking part in this study.

Appendix A. Peer Review Overview and Supplementary data

A Peer Review Overview and (sometimes) Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.crneur.2022.100059>.

References

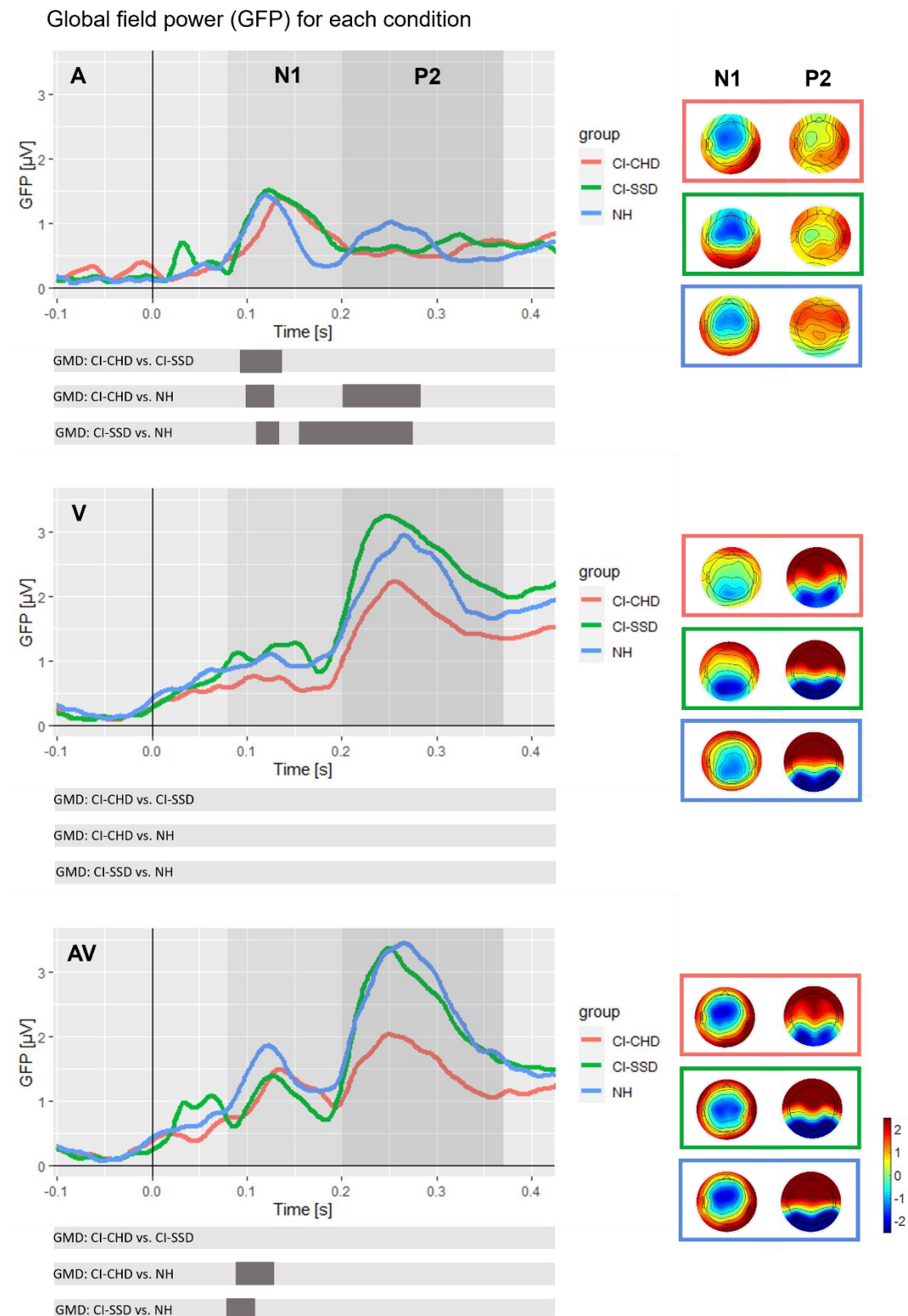
- Ahveninen, J., Jääskeläinen, I.P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.-H., Sams, M., Shinn-Cunningham, B.G., 2006. Task-modulated “what” and “where” pathways in human auditory cortex. *Proc. Natl. Acad. Sci. USA* 103 (39), 14608–14613.
- Allen, J.B., Hall, J., Jeng, P., 1990. Loudness growth in 1/2-octave bands (LGOB)—a procedure for the assessment of loudness. *J. Acoust. Soc. Am.* 88 (2), 745–753.
- Arndt, S., Aschendorff, A., Laszig, R., Beck, R., Schild, C., Kroeger, S., Ihorst, G., Wesarg, T., 2011. Comparison of pseudobinaural hearing to real binaural hearing rehabilitation after cochlear implantation in patients with unilateral deafness and tinnitus. *Otol. Neurotol.* 32 (1), 39–47.
- Arndt, S., Laszig, R., Aschendorff, A., Hassepass, F., Beck, R., Wesarg, T., 2017. Cochleaimplantatversorgung bei einseitiger Taubheit oder asymmetrischem Hörverlust. *HNO* 65 (7), 586–598.
- Barth, D.S., Goldberg, N., Brett, B., Di, S., 1995. The spatiotemporal organization of auditory, visual, and auditory-visual evoked potentials in rat cortex. *Brain Res.* 678 (1–2), 177–190.
- Bavelier, D., Neville, H.J., 2002. Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* 3 (6), 443–452.
- Bell, A.J., Sejnowski, T.J., 1995. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7 (6), 1129–1159.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B* 57 (1), 289–300.
- Bernstein, J.G., Goupell, M.J., Schuchman, G.I., Rivera, A.L., Brungart, D.S., 2016. Having two ears facilitates the perceptual separation of concurrent talkers for bilateral and single-sided deaf cochlear implantees. *Ear Hear.* 37 (3), 289.
- Besle, J., Fort, A., Delpuech, C., Giard, M.-H., 2004. Bimodal speech: early suppressive visual effects in human auditory cortex. *Eur. J. Neurosci.* 20 (8), 2225–2234.
- Beynon, A., Snik, A., Stegeman, D., Van den Broek, P., 2005. Discrimination of speech sound contrasts determined with behavioral tests and event-related potentials in cochlear implant recipients. *Ear Hear.* 26 (4), 442–453.
- Biasiucci, A., Franceschiello, B., Murray, M.M., 2019. Electroencephalography. *Current Biology* 29 (3), R80–R85.
- Billings, C.J., Bennett, K.O., Molis, M.R., Leek, M.R., 2011. Cortical encoding of signals in noise: effects of stimulus type and recording paradigm. *Ear Hear.* 32 (1), 53.
- Bönitz, H., Kopp, B., Büchner, A., Lunner, T., Lyxell, B., Finke, M., 2018. Event-related neuronal responses to acoustic novelty in single-sided deaf cochlear implant users: initial findings. *Clin. Neurophysiol.* 129 (1), 133–142.
- Bosnyak, D.J., Eaton, R.A., Roberts, L.E., 2004. Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebr. Cortex* 14 (10), 1088–1099.
- Bottari, D., Heimler, B., Caclin, A., Dalmolin, A., Giard, M.-H., Pavani, F., 2014. Visual change detection recruits auditory cortices in early deafness. *Neuroimage* 94, 172–184.
- Bottari, D., Bednaya, E., Dormal, G., Villwock, A., Dzhelyova, M., Grin, K., Pietrini, P., Ricciardi, E., Rossion, B., Röder, B., 2020. EEG frequency-tagging demonstrates increased left hemispheric involvement and crossmodal plasticity for face processing in congenitally deaf signers. *Neuroimage* 223, 117315.
- Brünecke, I., Hölsken, S., Kessler, J., 2018. Der DemTect Eye+ Ear—Neues kognitives Screeningverfahren bei schwerhörigen Menschen mit Demenzverdacht. *Z. für Audiol.* 57 (3), 121.
- Brunet, D., Murray, M.M., Michel, C.M., 2011. Spatiotemporal analysis of multichannel EEG: CARTOOL. *Comput. Intell. Neurosci.* 2011.
- Buechner, A., Brendel, M., Lesinski-Schiedat, A., Wenzel, G., Frohne-Buechner, C., Jaeger, B., Lenarz, T., 2010. Cochlear implantation in unilateral deaf subjects associated with ipsilateral tinnitus. *Otol. Neurotol.* 31 (9), 1381–1385.
- Campbell, J., Sharma, A., 2013. Compensatory changes in cortical resource allocation in adults with hearing loss. *Front. Syst. Neurosci.* 7, 71.
- Campbell, J., Sharma, A., 2016. Visual cross-modal re-organization in children with cochlear implants. *PLoS One* 11 (1), e0147793.
- Cappe, C., Thut, G., Romei, V., Murray, M.M., 2010. Auditory-visual multisensory interactions in humans: timing, topography, directionality, and sources. *J. Neurosci.* 30 (38), 12572–12580.
- Chen, L.-C., Sandmann, P., Thorne, J.D., Bleichner, M.G., Debener, S., 2016. Cross-modal functional reorganization of visual and auditory cortex in adult cochlear implant users identified with fNIRS. *Neural Plast.* 4382656–4382656.
- Cramer, A.O., van Ravenzwaaij, D., Matzke, D., Steingrover, H., Wetzels, R., Grasman, R.P., et al., 2016. Hidden multiplicity in exploratory multiway ANOVA: prevalence and remedies. *Psychonomic Bull. Rev.* 23 (2), 640–647.
- Crowley, K.E., Colrain, I.M., 2004. A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clin. Neurophysiol.* 115 (4), 732–744.

- Dale, A.M., Liu, A.K., Fischl, B.R., Buckner, R.L., Belliveau, J.W., Lewine, J.D., Halgren, E., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26 (1), 55–67.
- Debener, S., Hine, J., Bleck, S., Eyles, J., 2008. Source localization of auditory evoked potentials after cochlear implantation. *Psychophysiology* 45 (1), 20–24.
- Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134 (1), 9–21.
- Destrieux, C., Fischl, B., Dale, A., Halgren, E., 2010. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage* 53 (1), 1–15.
- Dong, L., Foo, S.W., Lian, Y., 2003. Modeling continuous visual speech using boosted viseme models. In: Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint, 3, pp. 1394–1398.
- Dorman, M.F., Zeidler, D., Cook, S.J., Loisel, L., Yost, W.A., Wanna, G.B., Gifford, R.H., 2015. Interaural level difference cues determine sound source localization by single-sided deaf patients fit with a cochlear implant. *Audiol. Neurotol.* 20 (3), 183–188.
- Doucet, M., Bergeron, F., Lassonde, M., Ferron, P., Lepore, F., 2006. Cross-modal reorganization and speech perception in cochlear implant users. *Brain* 129 (12), 3376–3383.
- Drennan, W.R., Rubinstein, J.T., 2008. Music perception in cochlear implant users and its relationship with psychophysical capabilities. *J. Rehabil. Res. Dev.* 45 (5), 779.
- Dürschmid, S., Reichert, C., Hinrichs, H., Heinze, H.-J., Kirsch, H.E., Knight, R.T., Deouell, L.Y., 2019. Direct evidence for prediction signals in frontal cortex independent of prediction error. *Cerebr. Cortex* 29 (11), 4530–4538.
- Fagel, S., Clemens, C., 2004. An articulation model for audiovisual speech synthesis—determination, adjustment, evaluation. *Speech Commun.* 44 (1–4), 141–154.
- Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., Sandmann, P., 2016a. On the relationship between auditory cognition and speech intelligibility in cochlear implant users: an ERP study. *Neuropsychologia* 87, 169–181.
- Finke, M., Sandmann, P., Bönitz, H., Kral, A., Büchner, A., 2016b. Consequences of stimulus type on higher-order processing in single-sided deaf cochlear implant users. *Audiol. Neurotol.* 21 (5), 305–315.
- Finney, E.M., Clementz, B.A., Hickok, G., Dobkins, K.R., 2003. Visual stimuli activate auditory cortex in deaf subjects: evidence from MEG. *Neuroreport* 14 (11), 1425–1427.
- Foxe, J.J., Morocz, I.A., Murray, M.M., Higgins, B.A., Javitt, D.C., Schroeder, C.E., 2000. Multisensory auditory–somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Cognit. Brain Res.* 10 (1–2), 77–83.
- Giraud, A.L., Price, C.J., Graham, J.M., Frackowiak, R.S., 2001a. Functional plasticity of language-related brain areas after cochlear implantation. *Brain* 124 (7), 1307–1316.
- Giraud, A.L., Truy, E., Frackowiak, R., 2001b. Imaging plasticity in cochlear implant patients. *Audiol. Neurotol.* 6 (6), 381–393.
- Giraud, A.-L., Price, C.J., Graham, J.M., Truy, E., Frackowiak, R.S., 2001c. Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron* 30 (3), 657–664.
- Glennon, E., Svirsky, M.A., Froemke, R.C., 2020. Auditory cortical plasticity in cochlear implant users. *Curr. Opin. Neurobiol.* 60, 108–114.
- Godey, B., Schwartz, D., De Graaf, J.B., Chauvel, P., Liegeois-Chauvel, C., 2001. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clin. Neurophysiol.* 112 (10), 1850–1859.
- Gordon, K.A., Jiwan, S., Papsin, B., 2013. Benefits and detriments of unilateral cochlear implant use on bilateral auditory development in children who are deaf. *Front. Psychol.* 4, 719.
- Gramfort, A., Papadopoulos, T., Olivi, E., Clerc, M., 2010. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed. Eng. Online* 9 (1), 1–20.
- Green, K.M., Julyan, P.J., Hastings, D.L., Ramsden, R.T., 2005. Auditory cortical activation and speech perception in cochlear implant users: effects of implant experience and duration of deafness. *Hear. Res.* 205 (1–2), 184–192.
- Hahlbrock, K.-H., 1970. Sprachaudiometrie: Grundlagen und praktische Anwendung einer Sprachaudiometrie für das deutsche Sprachgebiet. Georg Thieme Verlag.
- Hansen, P., Kringelbach, M., Salmelin, R., 2010. MEG: an Introduction to Methods. Oxford university press.
- Hari, R., Pelizzzone, M., Mäkelä, J.P., Hällström, J., Leinonen, L., 1987. Neuromagnetic responses of the human auditory cortex to on-and offsets of noise bursts. *Audiology* 26, 31–43.
- Hauthal, N., Thorne, J.D., Debener, S., Sandmann, P., 2014. Source localisation of visual evoked potentials in congenitally deaf individuals. *Brain Topogr.* 27 (3), 412–424.
- Heimler, B., Weisz, N., Collignon, O., 2014. Revisiting the adaptive and maladaptive effects of crossmodal plasticity. *Neuroscience* 283, 44–63.
- Henkin, Y., Yaar-Soffer, Y., Steinberg, M., Muchnik, C., 2014. Neural correlates of auditory-cognitive processing in older adult cochlear implant recipients. *Audiol. Neurotol.* 19 (Suppl. 1), 21–26.
- Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., Mckeown, M.J., Iragui, V., Sejnowski, T.J., 2000. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37 (2), 163–178.
- Kaas, J.H., Hackett, T.A., 2000. Subdivisions of auditory cortex and processing streams in primates. *Proc. Natl. Acad. Sci. USA* 97 (22), 11793–11799.
- Kitterick, P.T., Lucas, L., Smith, S.N., 2015. Improving health-related quality of life in single-sided deafness: a systematic review and meta-analysis. *Audiol. Neurotol.* 20 (Suppl. 1), 79–86.
- Kral, A., Hubka, P., Heid, S., Tillein, J., 2013. Single-sided deafness leads to unilateral aural preference within an early sensitive period. *Brain* 136 (1), 180–193.
- Laurienti, P.J., Kraft, R.A., Maldjian, J.A., Burdette, J.H., Wallace, M.T., 2004. Semantic congruence is a critical factor in multisensory behavioral performance. *Exp. Brain Res.* 158 (4), 405–414.
- Layer, N., Weglage, A., Müller, V., Meister, H., Lang-Roth, R., Walger, M., Murray, M., Sandmann, P., 2022. The Timecourse of Multisensory Speech Processing in Unilaterally Stimulated Cochlear Implant Users Revealed by ERPs. *Neuroimage: Clinical*, 102982.
- Lazard, D.S., Vincent, C., Venail, F., Van de Heyning, P., Truy, E., Sterkers, O., Skarzynski, P.H., Skarzynski, H., Schauwers, K., O’Leary, S., Mawman, D., Maat, B., Kleine-Punte, A., Huber, A.M., Green, K., Govaerts, P.J., Frayssé, B., Dowell, R., Dillier, N., Burke, E., Beynon, A., Bergeron, F., Başkent, D., Artières, F., Blamey, P.J., 2012b. Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model over time. *PLoS One* 7, e48739.
- Lehmann, D., Skrandies, W., 1980. Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalogr. Clin. Neurophysiol.* 48 (6), 609–621.
- Lin, F.-H., Witzel, T., Ahlfors, S.P., Stufflebeam, S.M., Belliveau, J.W., Hämäläinen, M.S., 2006. Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *Neuroimage* 31 (1), 160–171.
- Lomber, S.G., Meredith, M.A., Kral, A., 2010. Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nat. Neurosci.* 13 (11), 1421–1427.
- Ludwig, A.A., Meuret, S., Battmer, R.-D., Schönwiesner, M., Fuchs, M., Ernst, A., 2021. Sound localization in single-sided deaf participants provided with a cochlear implant. *Front. Psychol.* 12.
- Martuzzi, R., Murray, M.M., Michel, C.M., Thiran, J.P., Maeder, P.P., Clarke, S., Meuli, R. A., 2007. Multisensory interactions within human primary cortices revealed by BOLD dynamics. *Cerebr. Cortex* 17 (7), 1672–1679.
- Massaro, D.W., Light, J., 2004. Using Visible Speech to Train Perception and Production of Speech for Individuals with Hearing Loss.
- Meister, H., Fuersen, K., Schreitmüller, S., Walger, M., 2016. Effect of acoustic fine structure cues on the recognition of auditory-only and audiovisual speech. *J. Acoust. Soc. Am.* 139 (6), 3116–3120.
- Merzenich, M.M., Van Vleet, T.M., Nahum, M., 2014. Brain plasticity-based therapeutics. *Front. Hum. Neurosci.* 8, 385.
- Micco, A.G., Kraus, N., Koch, D.B., McGee, T.J., Carrell, T.D., Sharma, A., Nicol, T., Wiet, R.J., 1995. Speech-evoked cognitive P300 potentials in cochlear implant recipients. *Am. J. Otol.* 16 (4), 514–520.
- Michel, C.M., Koenig, T., 2018. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *Neuroimage* 180, 577–593.
- Michel, C.M., Murray, M.M., 2012. Towards the utilization of EEG as a brain imaging tool. *Neuroimage* 61 (2), 371–385.
- Michel, C.M., Murray, M.M., Lantz, G., Gonzalez, S., Spinelli, L., de Peralta, R.G., 2004. EEG source imaging. *Clin. Neurophysiol.* 115 (10), 2195–2222.
- Michel, C.M., Koenig, T., Brandeis, T., Gianotti, L.R., Wackermann, J., 2009. Electrical Neuroimaging. Cambridge University Press.
- Miller, J., 1982. Divided attention: evidence for coactivation with redundant signals. *Cognit. Psychol.* 14 (2), 247–279.
- Molholm, S., Ritter, W., Murray, M.M., Javitt, D.C., Schroeder, C.E., Foxe, J.J., 2002. Multisensory auditory–visual interactions during early sensory processing in humans: a high-density electrical mapping study. *Cognit. Brain Res.* 14 (1), 115–128.
- Moody-Antonio, S., Takayanagi, S., Masuda, A., Auer Jr., E.T., Fisher, L., Bernstein, L.E., 2005. Improved speech perception in adult congenitally deafened cochlear implant recipients. *Otol. Neurotol.* 26 (4), 649–654.
- Murray, M.M., Molholm, S., Michel, C.M., Heslenfeld, D.J., Ritter, W., Javitt, D.C., et al., 2005. Grabbing your ear: rapid auditory–somatosensory multisensory interactions in low-level sensory cortices are not constrained by stimulus alignment. *Cerebr. Cortex* 15 (7), 963–974.
- Murray, M.M., Brunet, D., Michel, C.M., 2008. Topographic ERP analyses: a step-by-step tutorial review. *Brain Topogr.* 20 (4), 249–264.
- Näätänen, R., Picton, T., 1987. The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24 (4), 375–425.
- Obleser, J., Elbert, T., Lahiri, A., Eulitz, C., 2003. Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. *Cognit. Brain Res.* 15 (3), 207–213.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia* 9 (1), 97–113.
- Park, L.R., Preston, E., Noxon, A.S., Dillon, M.T., 2021. Comparison of test methods to assess the implanted ear alone for pediatric cochlear implant recipients with single-sided deafness. *Cochlear Implants Int.* 1–8.
- Perrin, F., Pernier, J., Bertrand, O., Echallier, J.F., 1989. Spherical splines for scalp potential and current density mapping. *Electroencephalogr. Clin. Neurophysiol.* 72 (2), 184–187.
- Prince, P., Paul, B.T., Chen, J., Le, T., Lin, V., Dimitrijevic, A., 2021. Neural correlates of visual stimulus encoding and verbal working memory differ between cochlear implant users and normal-hearing controls. *Eur. J. Neurosci.* 54(3), 5016–5037.
- R Core Team, 2020. R: a language and environment for statistical computing. R Found. Stat. Comput. <https://www.R-project.org/>.
- Raab, D.H., 1962. Division of psychology: statistical facilitation of simple reaction times. *Trans. N. Y. Acad. Sci.* 24 (5 Series II), 574–590.
- Radecke, J.O., Schierholz, I., Kral, A., Lenarz, T., Murray, M.M., Pascale, S., 2022. Distinct Multisensory Perceptual Processes Guide Enhanced Auditory Recognition Memory in Older Cochlear Implant Users. *NeuroImage: Clinical*, 102942.

- Ratcliff, R., 1979. Group reaction time distributions and an analysis of distribution statistics. *Psychol. Bull.* 86 (3), 446.
- Ross, B., Tremblay, K., 2009. Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hear. Res.* 248 (1–2), 48–59.
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., Barone, P., 2007. Evidence that cochlear-implemented deaf patients are better multisensory integrators. *Proc. Natl. Acad. Sci. USA* 104 (17), 7295–7300.
- Rouger, J., Fraysse, B., Deguine, O., Barone, P., 2008. McGurk effects in cochlear-implemented deaf subjects. *Brain Res.* 1188, 87–99.
- Rouger, J., Lagleyre, S., Démonet, J.-F., Fraysse, B., Deguine, O., Barone, P., 2012. Evolution of crossmodal reorganization of the voice area in cochlear-implemented deaf patients. *Hum. Brain Mapp.* 33 (8), 1929–1940.
- Sandmann, P., Eichele, T., Buechler, M., Debener, S., Jäncke, L., Dillier, N., Hugdahl, K., Meyer, M., 2009. Evaluation of evoked potentials to dyadic tones after cochlear implantation. *Brain* 132 (7), 1967–1979.
- Sandmann, P., Kegel, A., Eichele, T., Dillier, N., Lai, W., Bendixen, A., Debener, S., Jancke, L., Meyer, M., 2010. Neurophysiological evidence of impaired musical sound perception in cochlear-implant users. *Clin. Neurophysiol.* 121 (12), 2070–2082.
- Sandmann, P., Dillier, N., Eichele, T., Meyer, M., Kegel, A., Pascual-Marqui, R.D., Marcar, V.L., Jäncke, L., Debener, S., 2012. Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain* 135 (2), 555–568.
- Sandmann, P., Plotz, K., Hauthal, N., Vos, M. de, Schönfeld, R., Debener, S., 2015. Rapid bilateral improvement in auditory cortex activity in postlingually deafened adults following cochlear implantation. *Clin. Neurophysiol.* 126 (3), 594–607.
- Schiel, F., 1999. Automatic phonetic transcription of non-prompted speech. Open Access LMU.
- Schierholz, I., Finke, M., Schulte, S., Hauthal, N., Kantzke, C., Rach, S., Büchner, A., Dengler, R., Sandmann, P., 2015. Enhanced audio–visual interactions in the auditory cortex of elderly cochlear-implant users. *Hear. Res.* 328, 133–147.
- Schierholz, I., Finke, M., Kral, A., Büchner, A., Rach, S., Lenarz, T., Dengler, R., Sandmann, P., 2017. Auditory and audio–visual processing in patients with cochlear, auditory brainstem, and auditory midbrain implants: an EEG study. *Hum. Brain Mapp.* 38 (4), 2206–2225.
- Schierholz, I., Schönermark, C., Ruigendijk, E., Kral, A., Kopp, B., Büchner, A., 2021. An event-related brain potential study of auditory attention in cochlear implant users. *Clin. Neurophysiol.* 132(9), 2290–2305.
- Schreitmüller, S., Frenken, M., Bentz, L., Ortmann, M., Walger, M., Meister, H., 2018. Validating a method to assess lipreading, audiovisual gain, and integration during speech reception with cochlear-implanted and normal-hearing subjects using a talking head. *Ear Hear.* 39 (3), 503–516.
- Senkowski, D., Pomper, U., Fitzner, I., Engel, A.K., Kral, A., 2014. Beta-band activity in auditory pathways reflects speech localization and recognition in bilateral cochlear implant users. *Hum. Brain Mapp.* 35 (7), 3107–3121.
- Sharma, A., Dorman, M.F., Spahr, A.J., 2002. A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear.* 23 (6), 532–539.
- Stevenson, R., Ghose, D., Fister, J.K., Sarko, D.K., Altieri, N.A., Nidiffer, A.R., Kurela, L. R., Siemann, J.K., James, T.W., Wallace, M.T., 2014. Identifying and quantifying multisensory integration: a tutorial review. *Brain Topogr.* 27 (6), 707–730.
- Strelnikov, K., Rouger, J., Démonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., Barone, P., 2010. Does brain activity at rest reflect adaptive strategies? Evidence from speech processing after cochlear implantation. *Cerebr. Cortex* 20 (5), 1217–1222.
- Strelnikov, K., Rouger, J., Démonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., Barone, P., 2013. Visual activity predicts auditory recovery from deafness after adult cochlear implantation. *Brain* 136 (12), 3682–3695.
- Stropahl, M., Debener, S., 2017. Auditory cross-modal reorganization in cochlear implant users indicates audio-visual integration. *Neuroimage: Clinica* 16, 514–523.
- Stropahl, M., Plotz, K., Schönfeld, R., Lenarz, T., Sandmann, P., Yovel, G., De Vos, M., Debener, S., 2015. Cross-modal reorganization in cochlear implant users: auditory cortex contributes to visual face processing. *Neuroimage* 121, 159–170.
- Stropahl, M., Bauer, A.-K.R., Debener, S., Bleichner, M.G., 2018. Source-modeling auditory processes of EEG data using EEGLAB and brainstorm. *Front. Neurosci.* 12, 309.
- Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M., 2011. Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011.
- Tremblay, C., Champoux, F., Lepore, F., Théoret, H., 2010. Audiovisual fusion and cochlear implant proficiency. *Restor. Neurol. Neurosci.* 28 (2), 283–291.
- Ulrich, R., Miller, J., Schröter, H., 2007. Testing the race model inequality: an algorithm and computer programs. *Behav. Res. Methods* 39 (2), 291–302.
- Vaughan Jr., H.G., 1982. The neural origins of human event-related potentials. *Ann. N. Y. Acad. Sci.* 388 (1), 125–138.
- Viola, F.C., De Vos, M., Hine, J., Sandmann, P., Bleeck, S., Eyles, J., Debener, S., 2012. Semi-automatic attenuation of cochlear implant artifacts for the evaluation of late auditory evoked potentials. *Hear. Res.* 284 (1–2), 6–15.
- Weglage, A., Müller, V., Layer, N., Abdellatif, K., Lang-Roth, R., Walger, M., Sandmann, P., 2022. Side-of-implantation effect on functional asymmetry in the auditory cortex of single-sided deaf cochlear-implant users. *Brain Topogr.* 1–22.
- Wesemann, W., Schiefer, U., Bach, M., 2010. Neue DIN-Normen zur Sehschärfestimmung. *Ophthalmologie* 107 (9), 821–826.
- Wesker, T., Meyer, B., Wagener, K., Anemüller, J., Mertins, A., Kollmeier, B., 2005. Oldenburg logatome speech corpus (OLLO) for speech recognition experiments with humans and machines. In: Ninth European Conference on Speech Communication and Technology.
- Widmann, A., Schröger, E., Maess, B., 2015. Digital filter design for electrophysiological data—a practical approach. *J. Neurosci. Methods* 250, 34–46.
- Williams, N., 2017. The borg rating of perceived exertion (RPE) scale. *Occup. Med.* 67 (5), 404–405.
- Williges, B., Wesarg, T., Jung, L., Geven, L.I., Radeloff, A., Jürgens, T., 2019. Spatial speech-in-noise performance in bimodal and single-sided deaf cochlear implant users. *Trends. Hear.* 23, 2331216519858311.
- Woods, D.L., Knight, R., Scabini, D., 1993. Anatomical substrates of auditory selective attention: behavioral and electrophysiological effects of posterior association cortex lesions. *Cognit. Brain Res.* 1 (4), 227–240.
- Zeng, F.-G., 2011. Advances in auditory prostheses. In: *Auditory Prostheses*. Springer, pp. 1–11.

Supplementary Material

We added a supplementary figure to illustrate the “raw” data for each condition, as we displayed only the condition A and the difference wave (AV-V) in the manuscript:



GFP of the conditions A, V, and AV for CI-CHD users (red), CI-SSD users (green) and NH listeners (blue). It is important to note that the GFP only provides positive values because it represents the standard deviation across all electrodes separately for each time point. The ERP topographies at the GFP peaks (N1(A) = CI-CHD: 147 ms, CI-SSD: 136 ms, NH: 118 ms; N1(V) = CI-CHD: 149 ms, CI-SSD: 134 ms, NH: 153 ms; N1(AV) = CI-CHD: 147 ms, CI-SSD: 144 ms, NH: 127 ms ; P2(A) = CI-CHD: 305 ms, CI-SSD: 288 ms, NH: 256 ms; P2(V) = CI-CHD: 259 ms, CI-SSD: 265 ms, NH: 268 ms; P2(AV) = CI-CHD: 253 ms, CI-SSD: 250 ms, NH: 264 ms) are given separately for each group (displayed on the right).

"THE INFLUENCE OF BACKGROUND NOISE AND NOISE REDUCTION BY 'FORWARDFOCUS' ON AUDIOVISUAL SPEECH PERCEPTION IN BIMODAL CI USERS: AN ERP STUDY"

The manuscript was submitted on November 7, 2022 to *Clinical Neurophysiology*, and is currently being reviewed. Therefore, unlike in the other publications, the manuscript is inserted the way it was submitted to *Clinical Neurophysiology*. This chapter is based on the following manuscript:

Layer, N., Abdel-Latif K. H., Radecke J-O., Müller, V., Weglage, A., Lang-Roth, R., Walger, M. & Sandmann, P. (2022). The influence of background noise and noise reduction by "ForwardFocus" on audiovisual speech perception in bimodal CI users: an ERP study. *Clinical Neurophysiology*, in revision.

Author contributions:

Natalie Layer: Conceptualisation; funding acquisition; literature review; creation, editing and calibration of stimuli; programming of the experiment; pilot measurements; participant recruitment; data acquisition; data analysis; interpretation and presentation of results to research group; discussion of findings and further procedures; writing initial manuscript draft; incorporating suggestions of co-authors; submission to journal; presentation of results at national conference. **Khaled H. Abdel-Latif:** Assistance in programming of the adaptive SNR measurement; partially writing of part "Adaptive signal-to-noise (SNR) measurement"; group discussions; proofreading of manuscript. **Jan-Ole Radecke:** Discussion of results; assistance in writing discussion on P2 results; proofreading of manuscript. **Verena Müller:** Advisory role (CI technician); assistance in participant enrolment; proofreading of manuscript. **Anna Weglage:** Assistance in data acquisition; group discussions; proofreading of manuscript. **Ruth Lang-Roth:** Conceptualisation; proof-reading of manuscript. **Martin Walger:** Conceptualisation; proof-reading of manuscript. **Pascale Sandmann:** Supervisor of Natalie Layer; conceptualisation; funding acquisition; suggestions; discussions; assistance in writing manuscript and proofreading of manuscript.

SYNOPSIS

OBJECTIVES Even though it is possible to achieve good speech comprehension with a CI in quiet listening situations (Fetterman and Domico, 2002; Holden et al., 2013; Krueger et al., 2008), hearing in noisy environments remains highly challenging

for CI users (Müller-Deile et al., 1995; Wilson and Dorman, 2008). Furthermore, it has been shown that hearing difficulties in noisy environments are reflected in the subjective listening effort (Han and Dimitrijevic, 2020; Hughes and Galvin, 2013; Winn, 2016). However, behavioural studies have demonstrated that noise reduction algorithms, such as "ForwardFocus", can improve speech intelligibility in demanding auditory situations (Mauger et al., 2014; Müller-Deile et al., 2008; Wolfe et al., 2015; Hey et al., 2019). CI users show strong audiovisual interactions (Radecke et al., 2022; Butera et al., 2018; Layer et al., 2022), and it is currently not well understood how the presence of background noise and the activation of a noise reduction algorithm, specifically "ForwardFocus", affect audiovisual speech perception in CI users. The aim of the current ERP study is to extend previous results by incorporating objective ERP measures and audiovisual speech conditions, as audiovisual conditions better reflect real-life hearing situations (Sumby and Pollack, 1954; Ross et al., 2007).

METHODS We recorded an EEG with 64 electrodes of bimodal CI users (CI on one ear, hearing aid on the other ear) who were provided with a CP1000 processor of the manufacturer Cochlear Limited. The CI users performed a syllable-identification task with the syllables /ba/, /da/, /ga/, presented from the front (0°) and stationary noise presented from the left and right (90° and 270°). The syllables were presented in three background noise conditions: 1) with background noise and "ForwardFocus" on (BN+FF+), 2) with background noise and "ForwardFocus" off (BN+FF-), and 3) without background noise and "ForwardFocus" off (BN-FF-). In addition, the syllables were presented either in auditory-only (A), visual-only (V) and audiovisual (AV) conditions. The task of the participants was to push a corresponding button as fast as possible on a touchscreen when a syllable was identified. In addition, we asked the participants to rate the task difficulty and the listening effort after each experimental block. Finally, we assessed the speech-in-noise performance using the clinical Oldenburg sentence test (OLSA; Wagener et al., 1999).

ANALYSIS The behavioural responses (hit rates, response times) and the cortical processing of auditory and audiovisual syllables were compared between the three noise conditions. The ERP data were analysed using a set of fronto-central sensory channels and a source analysis for a spatio-temporal evaluation of the ERPs, focusing on the N1 and the P2 time windows. For both analytical procedures, we statistically analysed differences between modality conditions (A, AV) and background noise conditions (BN+FF+, BN+FF-, BN-FF-). In the source analysis, we identified five activity-based regions of interest (insular cortex, lateral-occipital cortex, superior-frontal cortex, superior-temporal cortex, transverse-temporal cortex), which were also taken into account for the statistical analysis. In addition, we statistically analysed the subjective ratings of the task difficulty and the listening effort for each background noise condition. Finally, we compared the speech-in-noise performance in the OLSA with "ForwardFocus" switched on or off.

RESULTS The results revealed a multisensory benefit which was reflected in shorter response times and greater P2 amplitudes for audiovisual stimuli compared to purely auditory stimuli in expanded cortical regions, in particular the superior temporal

cortex, the lateral occipital cortex, and the insular cortex. Furthermore, we found that speech processing was generally hampered in the presence of background noise, as indicated by longer response times, greater listening effort, higher task difficulty, and a delayed activation in the auditory cortex at N1 latency, both for auditory and audiovisual conditions. But, we found evidence for successful noise reduction by the "ForwardFocus" algorithm in speech-in-noise conditions. This effect was reflected by shorter response times, reduced listening effort, lower task difficulty, and better speech intelligibility in the OLSA. In addition, our results provide objective evidence for beneficial effects of using "ForwardFocus" in speech-in-noise conditions, by showing enhanced cortical responses within the superior frontal cortex at P2 latency particularly in audiovisual listening situations.

DISCUSSION The results further confirm subjective and objective difficulties in hearing in CI users when competing background noise is present. Nevertheless, the results indicate that "ForwardFocus" has positive effects which are reflected in the behavioural and the ERP results. "ForwardFocus" allows for better speech intelligibility and reduced listening effort in complex auditory scenes by enabling the reallocation of attentional resources to the relevant auditory speech cues and by leading to a stronger perceived spatial congruency of auditory and available visual speech cues. Based on these results, we recommend using "ForwardFocus" to CI patients in appropriate listening situations.

The Influence of Background Noise and Noise Reduction by “ForwardFocus” on audiovisual Speech Perception in Bimodal CI Users: An ERP Study

Natalie Layer^{a,*}, Khaled H. A. Abdel-Latif^b, Jan-Ole Radecke^{c,d}, Verena Müller^a, Anna Weglage^a, Ruth Lang-Roth^a, Martin Walger^{a,b}, Pascale Sandmann^a

^aUniversity of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Otorhinolaryngology, Head and Neck Surgery, Audiology and Pediatric Audiology, Cochlear Implant Center, Germany

^bJean-Uhrmacher-Institute for Clinical ENT Research, University of Cologne, Germany

^cDept. of Psychiatry and Psychotherapy, University of Lübeck, Germany

^dCenter for Brain, Behaviour and Metabolism (CBBM), University of Lübeck, Germany

Abstract

Objective: Hearing with a cochlear implant (CI) is difficult in noisy environments, but the use of noise reduction algorithms, specifically ForwardFocus, can improve speech intelligibility. The current event-related potentials (ERP) study examined the electrophysiological correlates of this perceptual improvement.

Methods: CI users performed a syllable-identification task in auditory and audiovisual conditions, with syllables presented from the front and stationary noise presented from the sides. Brainstorm was used for spatio-temporal evaluation of ERPs.

Results: CI users revealed an audiovisual benefit as reflected by shorter response times and greater activation in temporal and occipital regions at P2 latency. However, in auditory and audiovisual conditions, background noise hampered speech processing, leading to longer response times and delayed auditory-cortex-activation at N1 latency. Nevertheless, activating ForwardFocus resulted in shorter response times, reduced listening effort and enhanced superior-frontal-cortex-activation at P2 latency, particularly in audiovisual conditions.

Conclusions: ForwardFocus enhances speech intelligibility in audiovisual speech conditions by allowing the reallocation of attentional resources to relevant auditory speech cues and/or by improving the perceived spatial congruency of auditory and visual speech signals.

Significance: This study shows for CI users that background noise and ForwardFocus differentially affect spatio-temporal cortical response patterns, both in auditory and audiovisual speech conditions.

*Corresponding author at: University of Cologne, Faculty of Medicine and University Hospital Cologne, Department of Otorhinolaryngology, Head and Neck Surgery, Audiology and Paediatric Audiology, Cochlear Implant Center, Kerpenener Strasse 62, 50937 Cologne, Germany, Email address: natalie.layer@uk-koeln.de (Natalie Layer)

Highlights

- Behavioural audiovisual improvements are reflected by enhanced activation in temporal and occipital regions at P2 latency
- Background noise delays responses within the auditory cortex (N1) and reduces superior frontal cortex activity (P2)
- Noise reduction via ForwardFocus improves audiovisual speech performance by allowing the reallocation of attentional resources

Keywords

Cochlea implant, ForwardFocus, noise reduction, bimodal hearing loss, evoked potential, audiovisual speech perception

Funding

This research was funded through a grant provided by Cochlear Research & Development Ltd under reference number IIR-2306.

Disclosure statement

The authors report no other potential or actual conflict of interest. The authors alone are responsible for the content and writing of this article.

1 Introduction

A cochlear implant (CI) is a neuroprosthetic device that can (partially) restore hearing in patients with sensorineural hearing loss through direct electrical stimulation of auditory nerve fibres (Zeng, 2011; House & Urban, 1973; Simmons et al., 1965). Hearing with a CI is highly different from natural acoustic hearing because the sounds transmitted by a CI are limited in terms of spectral and temporal information and have a lower dynamic range (Drennan & Rubinstein, 2008). Nevertheless, the CI users can learn to interpret these electrical speech signals as meaningful information (Green et al., 2008; Sandmann et al., 2015), and they can reach satisfactory levels of open-set speech recognition in optimal listening conditions within a few months after implantation (Fetterman & Domico, 2002; Hey et al., 2016; Holden et al., 2013; Krueger et al., 2008). However, in difficult hearing situations, as for example in speech conditions with interfering background noise, the CI users' speech comprehension typically deteriorates as a function of increasing background noise (Müller-Deile et al., 1995; Wilson & Dorman, 2008). Thus, CI users are critically more susceptible to

background noise in comparison to normal-hearing listeners, despite having many years of experience with the CI (Davidson et al., 2010; Fu & Nogaki, 2005; Radecke et al., 2022). In addition, it is often reported that the presence of additional background noise increases the subjective listening effort (Han & Dimitrijevic, 2020; Hughes & Galvin, 2013; Winn, 2016), which has been referred to as the cognitive resources and the attention necessary for comprehending speech (e.g., Gosselin & Gagne, 2011; Downs 1982). Therefore, it is not surprising that hearing-impaired individuals complain about hearing problems and high listening effort, particularly in noisy environments (Zhao et al., 1997; for review McCormack & Fortnum, 2013).

Improved speech intelligibility in background noise has been recognised as an important goal in CI users. To address this issue, CI manufacturers are working on developing noise reduction algorithms, as they proved to provide significant benefits in hearing performance in noisy environments (Mauger et al., 2014; Müller-Deile et al., 2008; Wolfe et al., 2015). One example is ForwardFocus, a new noise reduction algorithm available for the Nucleus 7 system's CP1000 and Kanso 2 sound processor of the manufacturer Cochlear (Hey et al., 2019). ForwardFocus can attenuate noise sources arriving from multiple locations in the background, in contrast to a directional adaptive beamformer microphone (BEAM) which adaptively directs its null towards a single dominant noise source. A directional microphone has a fixed null point at 210 degrees to provide the best directivity index, whereas ForwardFocus does not have a fixed null point but applies a digital attenuation to sound sources located on the sides and behind the listeners. ForwardFocus in Nucleus 7 is a user-controlled technology (enabled via app), which means that CI users can activate it whenever they are in a noisy environment. Activating the ForwardFocus algorithm allows the CI users to concentrate on a single speaker from the front while removing noise from behind, where other competing conversations are taking place. In contrast, other directional technologies in Cochlear sound processors (e.g., BEAM or ZOOM) are activated automatically in appropriate situations.

Hey et al. (2019) recently conducted a clinical evaluation of ForwardFocus with experienced CI users. Herein, the Oldenburg sentence test (OLSA; Wagener et al., 1999) was utilised to compare the benefits of ForwardFocus and BEAM in various noise conditions (Oldenburg noise, multiple-talker babble noise). Indeed, the results showed that ForwardFocus significantly improved speech comprehension in co-located frontal and spatially separated listening situations (speech: 0 degrees; noise: 90, 180 and 270 degrees). A follow-up study by Hey et al. (2021) used monosyllabic words, two-digit numbers, and sentences in noise to compare speech intelligibility between different, recent CI sound processors (Cochlear Limited). As expected, the results revealed benefits of using ForwardFocus in situations with speech from the front and noise from 90 degrees. Although these results are promising (Hey et al., 2019; 2021), they are based on subjective measures. An objective evaluation of beneficial ForwardFocus effects is still missing.

An optimal tool for objectively evaluating speech processing are event-related potentials (ERPs) derived from continuous electroencephalography (EEG). ERPs are characterized by an excellent temporal resolution, allowing the assessment of fast brain dynamics (Biasiucci et al., 2019; Michel & Murray, 2012). Especially the auditory N1 (negative potential around 100 ms after stimulus onset) and the P2 (positive potential around 200 ms after stimulus

onset) provide important insight into the cortical processing of auditory stimuli (Finke et al., 2016; Sandmann et al., 2009; Weglage et al., 2022). The generators of these ERP components are proposed, at least partly, within the primary and secondary auditory cortex and they reflect synchronous neural activity in response to acoustic changes (Hari et al., 1987; Näätänen & Picton, 1987; Pantev et al., 1996; Ross & Tremblay, 2009; Tremblay et al., 2014; Vaughan Jr & Ritter, 1970; Wolpaw & Penry, 1975). However, current models propose the existence of a semi-hierarchical and highly parallel organisation of auditory signal propagation (e.g., Kaas & Hackett, 2000). This implies that prominent components of auditory ERPs, such as the N1 and the P2, include generators not only within primary auditory cortices, but also within a distributed network including the superior temporal cortex, fronto-parietal structures, and even the visual cortex. Importantly, these auditory processes can be modulated by top-down effects, such as the direction of attention or the prediction of auditory events originating from the frontal cortex (Dürschmid et al., 2019).

One principal aim of this study is to objectively evaluate the advantages in speech perception provided by ForwardFocus in noisy environments by using ERPs in patients with a CI in one ear and a hearing aid (HA) in the other ear, referred to as bimodal CI users. Bimodal CI patients represent the most prevalent group of CI users in Germany (Hoppe et al., 2018; Holder et al., 2018). In addition, we extend previous evaluations by considering not only auditory but also audiovisual stimuli, since in everyday life communication both auditory and visual information (facial expressions, gestures, articulatory movements) contribute to the intelligibility of speech (Drijvers & Özyürek, 2017; Grant et al., 1998; Sumbly & Pollack, 1954). The role of the visual modality is even more crucial in difficult listening environments, such as speech conditions with interfering background noise (Ross et al., 2007; Sumbly & Pollack, 1954). If unisensory auditory signals are difficult to perceive (e.g., due to background noise), additional visual signals like lip movements can significantly improve the speech perception (van de Rijt et al., 2019). This remarkable increase in audiovisual integration (Bavelier et al., 2006; Kaiser et al., 2003; Mitchell & Maslin, 2007; Rouger et al., 2007; Stevenson et al., 2017; Rouger et al., 2007; Strelnikov et al., 2013), which especially is observed in conditions with poorly perceivable auditory and visual signals, has been referred to as “inverse effectiveness” (Stein & Meredith, 1993). Due to the limited CI input, it is not surprising that CI users show a pronounced visual influence on auditory perception (Desai et al., 2008; Butera et al., 2018) that even fosters subsequent auditory object recognition (Radecke et al., 2022). In addition, on a neurophysiological level, CI users demonstrate a pronounced interplay between the visual and auditory cortices (Giraud et al., 2001; Strelnikov et al., 2013). It seems that some of these multisensory interactions occur already at early cortical processing stages, as indicated by the altered auditory and visual cortex activation within the N1 and the P2 latency range in response to multisensory stimuli, compared to unisensory conditions (Layer et al., 2022; Schierholz et al., 2015, 2017). However, it remains unknown whether the strong audiovisual interactions in CI users are modulated by additional background noise and can be influenced by a noise reduction algorithm like ForwardFocus.

Here, we present an EEG study in which bimodal CI users were tested with syllables in auditory, visual, and audiovisual conditions. The auditory syllables (0 degrees) were presented with or without stationary background noise (90 and 270 degrees). The CI users were examined in the best-aided listening condition, using both the CI and the (contralateral) hearing aid and with the ForwardFocus noise reduction algorithm switched either on or off

in the speech-in-noise condition. We expected a *noise effect*, as indicated by prolonged N1 ERP latencies, delayed response times, reduced hit rates, and enhanced subjective listening effort due to the presence of additional background noise (e.g., Finke et al., 2016; Han & Dimitrijevic, 2020; Hughes & Galvin, 2013; Winn, 2016). Furthermore, we anticipated a *multisensory effect*, in particular enhanced ERP amplitudes and higher hit rates for the audiovisual compared to the auditory-only condition (e.g., Layer et al., 2022; Schierholz et al., 2015). Finally, we expected objective evidence for a *ForwardFocus effect* by finding enhanced speech intelligibility and weaker noise-induced ERP modulations for the condition with “ForwardFocus switched on” when compared to the condition with “ForwardFocus switched off”.

2 Material and methods

The study was carried out in accordance with the Declaration of Helsinki (2013) and was approved by the Ethics Committee of the University of Cologne’s medical faculty (application number: 21-1258_1). All participants provided written informed consent prior to data collection and were reimbursed.

2.1 Participants

Ten post-lingually deafened CI patients took part in this study (5 female, mean age: 62.3 ± 15.1 years, range: 33 – 81). All CI users were right-handed, as assessed by the Edinburgh inventory (Oldfield, 1971). The CI users were all bimodally fitted, which means that they use a CI on one side and a hearing aid (HA) on the contralateral ear. The participants were chosen based on their HA ear, which had a moderate (35-39 dB), a moderately severe (50-64 dB) or a severe (65-79 dB) hearing impairment, as determined by tone audiometry by averaging the values at 500, 1000, 2000, and 4000 Hz (Humes, 2019). All CI users were implanted with an implant of the type Nucleus CI622 with a Slim Straight electrode, and they were supplied with a CP1000 sound processor (Cochlear Limited). These requirements, which we set in order to make the bimodal CI users as homogeneous as possible, made obtaining these patients very difficult, especially given that the vast majority has profound hearing loss. The HA manufacturers and models varied between the participants. Table 1 contains detailed information on the HA systems, as well as demographic data for each participant. All participants had at least one year of CI experience, they were native German speakers with normal or corrected-to-normal vision (measured by the Landolt test according to the DIN-norm; Wesemann et al., 2010), and none of them had a psychiatric disorder history. All participants were tested bimodally, in the best-aided condition (with CI and HA) to be as close as possible to the CI users’ everyday life hearing situation.

Table 1: Demographic information of the CI participants; F = female; M = male; HL = hearing loss; L = left; R = right; HA = hearing aid.

ID	Sex	Age [years]	Age HL onset [years]	CI ear	Experience CI [months]	Experience HA [years]	Aetiology	HL [grade]	HA manufacturer	HA model
1	F	40	22	L	15	16	unknown	Moderate - severe	Widex	UNIQUE220 U-FS

2	F	33	30	R	18	1	sudden HL	Moderate - severe	ReSound	LiNX 3D 5
3	M	63	35	R	31	29	unknown	Moderate - severe	Signia	Pure 312 3X
4	M	74	55	L	24	17	noise exposure at work	Severe	Phonak	Naída Q70-SP
5	M	73	45	L	21	26	unknown	Moderate - severe	T indera	500 M
6	F	72	37	R	23	29	unknown	Moderate - severe	Signia	Motion 13P 2Nx
7	M	61	50	L	12	10	unknown	Severe	Bernafon	VN5 MNR 100
8	M	64	21	L	43	39	unknown	Moderate	Signia	Intuis 3 M
9	F	62	9	R	20	51	unknown	Severe	Phonak	Naída P-UP
10	F	81	65	L	33	13	unknown	Moderate - severe	ReSound	Key 3

2.2 Stimuli

The stimuli used in this study were taken from the freely available OLAVS stimulus pool (Stropahl et al., 2017). They consisted of the syllables /ba/, /da/ and /ga/ and were spoken (audio file) and visualised (muted video file of the lip movements) by the female speaker 'TK04'. According to the documentation of the OLAVS stimulus pool, all stimuli were normalised to -30 dB FS. The syllables were presented in three different sensory conditions: visual-only (V), auditory-only (A) and audiovisual (AV) (see Figure 1A). Pinnacle Studio 22 (version 22.3.0.377) was used to edit the syllables by creating video files of each syllable in each condition: 1) AV: articulatory movements with corresponding auditory syllables, 2) A: black screen (video track turned off) combined with auditory syllables, 3) V: articulatory movements without sound (audio track turned off). Each sensory condition (except for A) started with a static face (duration of 750 ms) and was followed by the video, which lasted 1070 ms (140 ms articulatory movement initiation + 930 ms auditory syllable and completion of articulatory movements). After the presentation of the video file, a response screen appeared with all three syllables (/ba/, /da/, /ga/) as response options. The response buttons were arranged on the screen such that they had the same distance to a marker on the table. An illustration of an audiovisual trial can be found in Figure 1B. The stimuli were presented using the Presentation software (Neurobehavioral Systems, version 21.1) and a computer in combination with a duplicated touch screen monitor (Phillips 222B9T/00; 21.5 inch) placed in front of the participant. The syllables were presented via a loudspeaker placed in front of the participant (0 degrees, 1.15 m distance, height: individually adjusted according to the height of the participants). Please note the loudspeaker presenting the syllables was always positioned above the touchscreen. For conditions with background noise, we used the stationary 'olnoise' which is generated by multiple random superpositions of the Oldenburg sentence test corpus (OLSA; Wagener et al., 1999). The background noise was presented via loudspeakers placed on both the right and left ears (90 and 270 degrees; 1.15 m distance, height: individually adjusted according to the height of the participants on ear level; model: KRK Classic 5). The setting is displayed in Figure 1C. In Audacity (version 3.0.5), we generated a decorrelated version of the noise by splitting the stereo file into two mono files, which were

then recombined with a slight delay to each other. This procedure is advantageous because the tested person does not perceive a constant, simultaneously presented noise from both sides, but rather a noise more akin to noise encountered in everyday situations. The sound intensity level of the syllables was calibrated to 65 dB SPL, and the sound intensity level of the noise was individually adjusted (see next section). In addition, all participants rated the perceived loudness of the syllables using a seven-point loudness rating scale (e.g., Layer et al., 2022; Sandmann et al., 2010). This rating revealed that all CI users perceived the syllables as moderately loud, which corresponds approximately to an intensity level of 60 – 70 dB SPL.

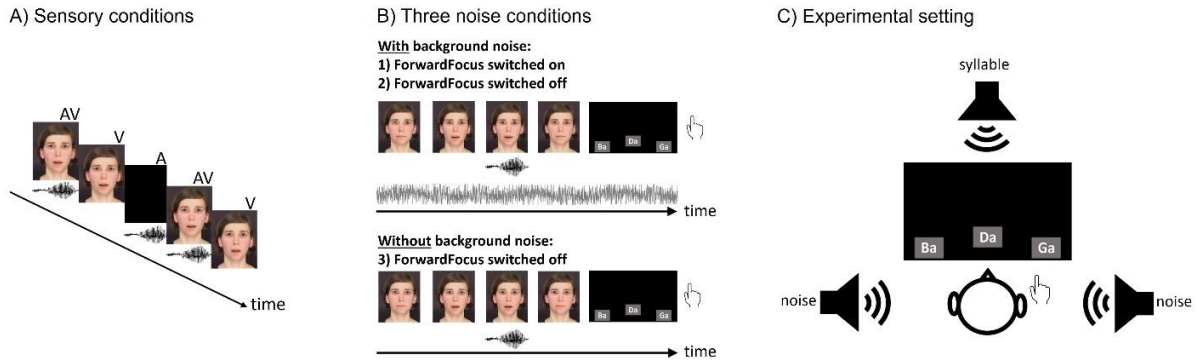


Figure 1: Experimental setting, paradigm and conditions. A) Overview of the sensory conditions (A, V, AV) that are presented pseudo randomly during the experiment. B) Overview of a trial and the three background noise conditions (adapted from Stropahl & Debener (2017)): 1) with background noise - ForwardFocus switched off, 2) with background noise - ForwardFocus switched on, 3) without background noise - ForwardFocus switched off. C) Experimental setting with a participant seated in front of a touchscreen monitor, hearing syllables from the front (0 degrees) and background noise from the sides (90 and 270 degrees), and seeing the stimuli on the monitor. The task was to select the heard syllable on the monitor via touch.

2.3 Adaptive signal-to-noise (SNR) measurement

To obtain an individually adjusted noise level for each participant in the EEG experiment, we programmed an adaptive signal-to-noise (SNR) measurement in MATLAB (version R2020b; Mathworks, Natick, MA). First, a training was carried out by presenting the same three syllables as in the actual experiment (/ba/, /da/, /ga/) in a randomised order in presence of noise. The noise was a stationary noise ('olnoise'; Wagener et al., 1999) and thus was the same noise type as used for the EEG experiment. We started the training by presenting some trials in quiet, then adding noise with different SNRs with the goal that the participant still had a syllable intelligibility of 100 %. This was done manually and was repeated until we were confident that the participants provided consistent responses. We used the end SNR value corresponding to 100 % syllable intelligibility from training as our start value for the adaptive procedure. By that, all participants started the adaptive procedure with the same syllable intelligibility. The adaptive procedure again presented randomly one of the three syllables. It calculated the mean after each twenty responses and adjusted the upcoming SNR corresponding to the goal syllable intelligibility (which was set to 80 %) by increasing or decreasing the SNR value by 2 dB steps (in total 77 ± 41 trials required). The duration of the adaptive measurement was approximately 10 - 15 minutes. This individual SNR was then used for the main EEG experiment, and it ensured that the performance was comparable across the participants. We selected the 80 % criterion because it avoided ceiling effects and

it allowed to measure a potential benefit from ForwardFocus. Furthermore, it is critical for the EEG analysis to ensure that the participants can recognise the syllables well enough, because only correct trials are analysed, and a certain number of correct trials is required to perform an EEG analysis. The individual SNR values can be found in Table 2.

2.4 Procedure

The participants were seated comfortably in an electromagnetically shielded and dimly lit room. Each sensory condition began with a static face of the speaker (duration of 750 ms), which was followed by a syllable (A, V, or AV) and a response screen with three response options to choose from (see Figure 1). The CI users were instructed to select as fast as possible the syllable which was perceived on the touch screen by pressing the corresponding area on the screen. Furthermore, we placed a marker on the table and instructed the participant to place their hand on this area while watching the video and to return to this point after registering the response. This was done for relieving the hand and to ensure that the distance was the same to each button. We created two versions of the presentation script, enabling a variation in the order of the response buttons on the screen. For each sensory condition (A, V, AV) and background noise condition (*without* background noise-ForwardFocus off, *with* background noise-ForwardFocus off, *with* background noise-ForwardFocus on), 84 trials were presented, except for the V condition, which was only presented in the condition without background noise. This resulted in a total number of 588 trials (84 repetitions x 2 sensory conditions (A, AV) x 3 background noise conditions) + 84 visual sensory conditions (V) without background noise only, which will be reported elsewhere). The sensory conditions were pseudo-randomised such that no two trials of the same sensory condition and syllable appeared in a row. We created two experimental blocks for every background noise condition, lasting about six to seven minutes each. The order of blocks was counterbalanced across the participants to avoid order effects of the background noise conditions in the data. The EEG recording time was approximately 38 minutes, excluding breaks. After each block, a short break was given and the CI users were asked to rate the subjective listening effort and the task difficulty on a scale from 0 to 5 (scale: 0=no effort; 5=very effortful/very difficult). We presented a short practise block of 12 trials before beginning the EEG recording to ensure that the task was understood.

Regarding the CI settings, the ForwardFocus was enabled, and all other noise-processing algorithms were deactivated. The processor was linked to a tablet (Apple inc. version 2020) with the Nucleus App installed. We used this app to activate or deactivate ForwardFocus whenever it was required during the experiment.

2.5 Further behavioural measures

Apart from the behavioural measures from the EEG task (response times (RTs), hit rates) and the rating of the listening effort and task difficulty after each experimental block, we obtained results from a working memory task, a behavioural control task and from a number of audiometric tests.

To obtain a measure for working memory capacity in our participants, we administered the German version of the size-comparison span (SICSPAN) test, which was developed by Sörqvist et al. (2010). We decided to include this test, as speech in noise is often associated

with enhanced recruitment of cognitive processes, such as working memory (Rönnerberg et al., 2013). As reported in a previous study with CI users (Finke et al., 2016), we analysed the results of the SICSPAN test by taking the total sum of remembered words.

The goal of the behavioural control task was to demonstrate that reaction times were equal for each button, thereby ruling out the possibility that one button was pushed faster than another, potentially confounding the behavioural results of the EEG experiment. This short reaction time experiment (approximately one minute duration) was constructed in a similar fashion as the EEG task, including three buttons (one of them red, the others grey). These buttons were placed at the same height and distance as the response buttons, and the task of the participants was to press the red button as fast as possible. The order of the appearance of the red button was randomised. Similar to the EEG paradigm, the CI users were instructed to return to a marked area on the table after registering a response. The experimental setting of this experiment can be viewed in Supplementary Figure 1A.

To obtain a detailed picture of the hearing abilities of each CI patient, we used pure-tone audiometry at 500, 1000, 2000, and 4000 Hz to calculate the pure tone average for these four frequencies (PTA4) to know the hearing threshold and the German Oldenburg sentence test (OLSA; Wagener et al., 1999) in noise. Pure-tone audiometry was conducted in free-field with 1) the CI-ear only (contralateral ear without HA and with ear-plug inserted), 2) the HA-ear only (with and without HA, contralateral ear without CI but with ear-plug inserted). Similar to the experimental setting during EEG recordings, the OLSA was conducted in noise with both ears, with the CI users positioned in such a way that the sentences came from the front and the noise came from the side of the CI-ear, which was similar to the setting in the EEG experiment (90 degrees). We made two runs, once with ForwardFocus activated and once with ForwardFocus deactivated. Again, we made sure that the order was balanced among participants. The scores in the audiometric tests can be found in Table 2.

Table 2: Further behavioural results for each participant. The hearing threshold is the average over 500 Hz, 1 kHz, 2 kHz, 4 kHz. For OLSA in noise, the SNR (signal-to-noise ratio) is given, keeping the noise constantly at 65 dB SPL and varying the intensity of the sentences. The 'SNR' column displays the SNR values calculated from the adaptive measurement to obtain the SNR for 80 % correct syllable identification for the EEG experiment. For the SICSPAN, the total sum of remembered words is given in the present table (maximal amount = 63 words).

ID	Hearing threshold (PTA4)			OLSA in noise (dB SNR)		Individual SNR (dB)	SICSPAN
	HA-ear without HA	HA-ear with HA	CI-ear	ForwardFocus off	ForwardFocus on		
1	62.5	31.75	33	-2.7	-5.2	-10	45
2	60.75	35.25	33.25	-3.4	-7.5	-12	47
3	68	38.25	31.25	-6.2	-13.2	-11	30
4	72.75	44	33.25	7.7	4	-2	44

ID	Hearing threshold (PTA4)			OLSA in noise (dB SNR)		Individual SNR (dB)	SICSPAN
5	62.25	49.25	38.75	2.2	-2.1	-11	41
6	68	32.5	36.75	0.2	-4.5	-8	41
7	73.5	50	36.75	9.5	2.2	-5	37
8	49.75	38.75	31	2.9	1.5	-13	44
9	31.25	75	33.5	-3.4	-4.5	-6	55
10	57.5	39.25	33.5	4.3	1.9	-3	46

2.6 EEG recording

The EEG data were collected continuously using 64 AG/AgCl ActiCap slim electrodes and a BrainAmp system (BrainProducts, Gilching, Germany), and a customised electrode cap with an electrode layout (EasyCap, Herrsching, Germany) in accordance with the 10-10 system. Two electrodes were placed beneath and beside (vertical and horizontal eye movements, respectively) the left eye to record an electrooculogram (EOG). A nose-tip electrode was used as reference, and a midline electrode placed slightly anterior to Fz was used as ground. The data were recorded using a sampling rate of 1000 Hz and an online analog filter between .02 and 250 Hz. Electrode impedances were reduced below 10 k Ω prior to data collection.

2.7 Data analysis

Behavioural data analysis was performed using R (version 1.2.5033; R Core Team (2020), Vienna, Austria). Sensor ERP analyses were conducted in MATLAB 9.8.0.1323502 (R2020a; Mathworks, Natick, MA) and R. Source analyses were performed in Brainstorm (Tadel et al., 2011) and R. The following R packages were used: ggplot2 (version 2.3.3) for creating plots; dplyr (version 1.0.4), tidyverse (version 1.3.0) and tidyr (version 1.1.3) for data formatting; ggpubr (version 0.4.0) and rstatix (version 0.7.0) for statistical analyses.

2.7.1 Behavioural data analysis

First, false alarms and missing responses were removed from the dataset. Outliers were defined as trials that exceeded the individual mean by more than three standard deviations for each condition and were removed from the dataset. We focused on the moving face (starting 750 ms post-stimulus onset/after the static face), as the responses to static faces had previously been investigated (Stropahl et al., 2015). Then, for each CI user, RTs and hit rates were calculated for each condition (A, AV) and each background noise condition (*without* background noise, *with* background noise-ForwardFocus off, *with* background noise-ForwardFocus on). To analyse the performance for each condition, a 2 x 3 repeated measures ANOVA was used separately for the RTs and the hit rates, with sensory condition (AV, A) and background noise condition (*without* background noise-ForwardFocus off, *with* background noise-ForwardFocus off, *with* background noise-ForwardFocus on) as the within-subjects factors.

To assess differences in task difficulty and subjective listening effort between the three noise conditions, a one-way repeated measures ANOVA was used. To analyse whether the RTs in the behavioural control task differed between the button positions, we performed a one-way ANOVA. A Greenhouse-Geisser correction was used whenever the sphericity assumption was violated in the ANOVA results. Furthermore, in the case of significant main effects or interactions, post-hoc t-tests were performed, which were corrected for multiple comparisons using a Bonferroni-Holm correction ($p \leq .05$). Finally, to compare the SNR values obtained in the OLSA in noise, we calculated a two-sample t-test to assess differences between switching ForwardFocus on and off.

2.7.2 EEG pre-processing

EEGLAB was used for conducting the pre-processing of the EEG data (version v2021.0.0; Delorme & Makeig, 2004), which is running in the MATLAB environment (Mathwork, Natick, MA). The raw data were downsampled (500 Hz) and filtered with a FIR-filter with a high pass cut-off frequency of 0.1 Hz and a maximum possible transition bandwidth of 1 Hz (cut-off frequency multiplied by two), and a low pass cut-off frequency of 40 Hz and a transition bandwidth of 2 Hz. The Kaiser-window was used for both filters (Kaiser- $\beta = 5.653$, max. stopband attenuation = -60 dB, max. passband deviation = 0.001; Widmann et al., 2015) to maximise energy concentration in the main lobe by averaging out noise in the spectrum and minimising information loss at the window's edges. Electrodes near the speech processor and transmitter coil were detached (mean: 3.1 electrodes; range: 2 – 4). Then, the datasets were epoched into 2 s dummy epoch segments, and pruned of unique, non-stereotype artefacts with a threshold criterion of four standard deviations. An independent component analysis (ICA) was computed (Bell & Sejnowski, 1995), and the resulting ICA weights were applied to the epoched original data (1 – 40 Hz, -200 to 2000 ms relative to the onset of the static face)). Independent components representing vertical and horizontal ocular movements, electrical heartbeat activity, and other non-cerebral activity sources were excluded (Jung et al., 2000). Independent components exhibiting CI artefacts were identified and removed based on the location of the centroid (CI side) and the time course of the component activity, showing a pedestal artefact around 890 ms which reflects the auditory stimulus onset. In the following step, we used a spherical spline interpolation (Perrin et al., 1989) to interpolate the previously removed channels near the speech processor and transmitter coil. Only those trials yielding correct responses were kept for further analyses.

2.7.3 ERP data analysis - sensor level

We compared event-related potentials (ERPs) of all sensory conditions (AV, A) between the three background noise conditions (*without* background noise-ForwardFocus off, *with* background noise-ForwardFocus off, *with* background noise-ForwardFocus on). We identified the N1 and P2 time windows based on visual inspection of the great grand averages and the single subject data. We selected an electrode cluster consisting of fronto-central electrodes (Fz, FC1, Fz, FC2, Cz) for the peak detection. This electrode cluster is commonly used for auditory and audiovisual experiments (e.g., Finke et al., 2016; Weglage et al., 2022). The N1 time window for peak and latency detection was between 250 ms and 350 ms (please note: auditory stimulus onset was 140 ms after the onset of the visual articulatory movement onset), and the P2 time window was identified between 380 ms and 480 ms. The selected

electrode cluster as well as the ERPs including the time windows of the N1 and P2 are displayed in Figure 3A. The corresponding peak latencies can be found in Table 3. We calculated for each condition a 2 x 3 repeated measures ANOVA separately for the N1 and P2 peak means and latencies, with sensory condition (AV, A) and background noise condition (*without* background noise-ForwardFocus off, *with* background noise-ForwardFocus off, *with* background noise-ForwardFocus on) as the within-subjects factors. Whenever the sphericity assumption was violated, a Greenhouse-Geisser correction was applied. Furthermore, when there were significant main effects or interactions, post-hoc t-tests were performed, and multiple comparisons were corrected using the Bonferroni-Holm correction ($p \leq .05$).

Table 3: ERP peak mean latencies for each noise and sensory condition. The peak mean latencies for the N1 and the P2 are given with their corresponding standard deviation. Note that the onset of the auditory part of the syllable was 140 ms which came after the onset of the visual articulatory movement. When subtracting these 140 ms from the peak latency values reported here, one will get the typical approximately 100-200 ms and 200-300 ms ranges for the elicitation of the N1 and the P2, respectively. Without background noise-ForwardFocus off = BN-FF-, with background noise-ForwardFocus off = BN+FF-, with background noise-ForwardFocus on = BN+FF+.

Noise condition	Sensory condition	Peak latency [ms]	
		N1	P2
BN+FF-	A	314 ± 29	415 ± 25
BN+FF-	AV	314 ± 25	419 ± 17
BN+FF+	A	313 ± 24	428 ± 17
BN+FF+	AV	309 ± 23	412 ± 18
BN-FF-	A	294 ± 16	402 ± 20
BN-FF-	AV	293 ± 24	412 ± 14

2.7.4 ERP data analysis - source level

We performed an ERP source analysis for the sensory and background noise conditions by using Brainstorm (Tadel et al., 2011; Stropahl et al., 2018). Similar to previous studies with CI patients (Bottari et al., 2020; Layer et al., 2022; Stropahl & Debener, 2017), we chose the method of dynamic statistical parametric mapping (dSPM, Dale et al. 2000)). Even though the spatial resolution remains low, dSPM works more precisely in identifying deeper sources than standard norm methods (Lin et al., 2006). To approximate the locations of electrical activity recorded on the scalp, minimum-norm inverse maps with constrained dipole orientations are used. This method is effective for locating small cortical areas such as the auditory cortex (Stropahl et al., 2018). To estimate single-subject based noise standard deviations at each location, individual noise covariances are calculated from single-sensory condition pre-stimulus onset baseline intervals (-50 to 0 ms) (Hansen et al., 2010). The boundary element method (BEM), which is implemented in OpenMEEG, was used as a head model. The BEM provides three realistic layers as well as representative anatomical information (Gramfort et al., 2010). Based on the normalisation within the dSPM algorithm, the final activity data is then displayed as absolute values with arbitrary units. We chose regions of interest (ROIs) based on the inbuilt function to compute a t-test against baseline (Bonferroni corrected, $p = .001$) for which we used the grand averaged data (average over

the sensory conditions A and AV and all noise conditions) to determine where most of the activity is concentrated on the cortex. These activity-based areas were matched with predefined regions from the 'Desikan-Killiany' atlas which is implemented in Brainstorm. The resulting final ROIs are 'insula', 'lateraloccipital', 'superiorfrontal', 'superiortemporal' and 'transversetemporal'. EEG source analysis can be very accurate (approximately 2 cm in ideal conditions; Klamer et al., 2015), however this is only ensured when using individual head models for each participant, which is not the case in this study. Therefore, we decided to choose larger ROIs for the statistical analyses, as provided by the inbuilt ROIs in Brainstorm. Supplementary Figure 2 depicts the maximal activity (A) and the ROIs which were selected based on this maximal activity (B).

Source activities for each ROI, modality and background noise condition were exported from Brainstorm for each CI user. The peak means and latencies were then extracted for each time window of interest (N1: 250 – 350 ms, P2: 350 – 480 ms; relative to the onset of the visual articulatory movements). The source ERPs for three regions, including the time windows of the N1 and P2, are displayed in Figure 4A. We calculated for each ROI a 2 x 3 repeated measures ANOVA separately for the N1 and P2 peak means and latencies, with sensory condition (AV, A) and background noise condition (*without* background noise-ForwardFocus off, *with* background noise-ForwardFocus off, *with* background noise-ForwardFocus on) as the within-subjects factors. Whenever the sphericity assumption was violated, a Greenhouse-Geisser correction was applied. In the case of significant interaction or main effects, post-hoc t-tests were computed and corrected for multiple comparisons using a Bonferroni-Holm correction ($p \leq .05$).

3 Results

To facilitate reading, the background noise conditions are abbreviated in the subsequent sections as follows: *without* background noise-ForwardFocus off = BN-FF-, *with* background noise-ForwardFocus off = BN+FF-, *with* background noise-ForwardFocus on = BN+FF+.

3.1 Behavioural results from the ERP experiment

In general, all participants showed relatively high hit rates of $\geq 81.3\%$ for A and AV, even in the conditions with background noise. The hit rate for V was 71.4 %. The hit rates and response times (RTs) can be found in Table 4.

Table 4: Mean hit rates (in %) and mean response times (in ms) for each sensory and background noise condition (*without* background noise-ForwardFocus off = BN-FF-, *with* background noise-ForwardFocus off = BN+FF-, *with* background noise-ForwardFocus on = BN+FF+).

Noise condition	Sensory condition	Hit rates	Response times
BN+FF-	A	87.9 $\pm$ 7.7	1873 $\pm$ 168
BN+FF-	V	-	-
BN+FF-	AV	95.3 $\pm$ 3.7	1761 $\pm$ 130
BN+FF+	A	83.8 $\pm$ 7.5	1810 $\pm$ 141
BN+FF+	V	-	-

Noise condition	Sensory condition	Hit rates	Response times
BN+FF+	AV	94.1 ± 4.7	1715 ± 89
BN-FF-	A	81.3 ± 10.8	1761 ± 147
BN-FF-	V	71.4 ± 4.0	1749 ± 79
BN-FF-	AV	89.3 ± 11.4	1680 ± 109

The 2 x 3 repeated measures ANOVA for the *hit rates*, with sensory condition (AV, A) and background noise condition (BN+FF-, BN+FF+, BN-FF-) as the within-subjects factors resulted only in a main effect of sensory condition ($F_{1,9} = 18.263$; $p \leq .01$, $\eta p^2 = 0.235$). Follow-up post-hoc t-tests showed that hit rates to AV syllables were significantly higher when compared to A, which is referred to as *multisensory effect* ($t(29) = -5.84$; $p \leq .001$; Figure 2A).

Concerning the *RTs*, the 2 x 3 repeated measures ANOVA, with sensory condition (AV, A) and background noise condition (BN+FF-, BN+FF+, BN-FF-) as the within-subjects factors revealed a main effect of sensory condition ($F_{1,9} = 10.737$; $p \leq .01$, $\eta p^2 = 0.126$) and a main effect of background noise condition ($F_{1.26,11.37} = 5.53$; $p \leq .05$, $\eta p^2 = 0.09$). For the main effect of sensory condition, post-hoc t-tests showed that RTs for AV syllables were significantly faster than for A syllables, which is referred to as *multisensory effect* ($t(29) = 5.66$; $p \leq .001$; Figure 2A). Resolving the main effect of background noise condition, post-hoc t-tests revealed that all conditions differed from each other: BN-FF- vs. BN+FF+ ($t(19) = 2.96$; $p \leq .01$), BN-FF- vs. BN+FF- ($t(19) = 3.6$; $p \leq .01$) and BN+FF+ vs. BN+FF- ($t(19) = 2.94$; $p \leq .01$), which is referred to as *noise effect* and *ForwardFocus effect*. These results are displayed in Figure 2B.

3.2 Other behavioural results

Differences in task difficulty and subjective listening effort between the three background noise conditions (BN+FF-, BN+FF+, BN-FF-) were analysed with a one-way repeated measures ANOVA. For the *subjective listening effort*, the ANOVA revealed a main effect of background noise condition ($F_{2,27} = 4.112$; $p \leq .05$, $\eta p^2 = 0.233$). The post-hoc t-test showed a difference between BN+FF+ and BN+FF- ($t(9) = 6.53$; $p \leq .001$) and between BN+FF- and BN-FF- ($t(9) = 4.25$; $p \leq .01$). In addition, there was a trend for a difference between BN-FF- and BN+FF+ ($t(9) = 2.19$; $p = .056$). So, the subjective listening effort was least effortful for BN-FF-, more effortful for BN+FF+ and most effortful for BN+FF- (*noise effect and ForwardFocus effect*).

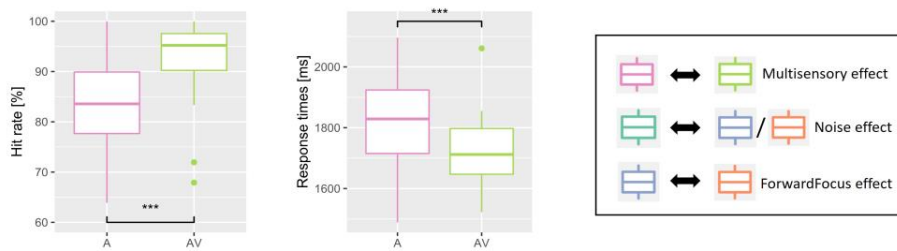
Concerning the *subjective task difficulty*, the ANOVA showed a main effect of background noise condition ($F_{2,27} = 3.344$; $p \leq .05$, $\eta p^2 = 0.119$). Post-hoc t-test revealed a difference between all noise conditions: BN+FF+ and BN+FF- ($t(9) = 2.91$; $p \leq .05$), BN+FF- and BN-FF- ($t(9) = 3.95$; $p \leq .01$) and BN-FF- and BN+FF+ ($t(9) = 3.7$; $p \leq .01$). So, the subjective task difficulty was least difficult for BN-FF-, followed by BN+FF+ and BN+FF- (*noise effect and ForwardFocus effect*) The results of the listening effort and task difficulty can be inspected in Figure 2B.

Concerning the difference between SNR values obtained in the *OLSA in noise* when switching ForwardFocus on and off, the two-sample t-test showed that there was a significant difference between using ForwardFocus and not using ForwardFocus ($t(9) = 5.25$; $p \leq .001$). Thus, activating ForwardFocus in the OLSA lowers the SNRs when compared to performing this test without ForwardFocus (*ForwardFocus effect*).

To assess whether the RTs in the *behavioural control task* differed between the button positions, we performed a one-way ANOVA. The ANOVA did not reveal a button position effect ($F_{2,597} = 0.795$; $p = .452$, $\eta p^2 = 0.003$; see Supplementary Figure 1B).

Behavioural results:

A) Multisensory effect



B) Noise effect + ForwardFocus effect

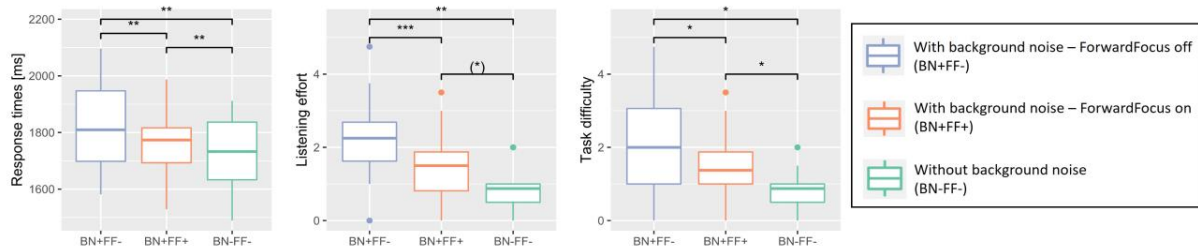


Figure 2: Behavioural results. A) There is a multisensory effect both for hit rates and RTs, showing higher hit rates and shorter RTs for AV compared to A, respectively. All possible effects are defined in the box on the top right. B) There is a noise effect and an additional effect of noise reduction via ForwardFocus. These effects were observed in the response times, the subjective listening effort and the subjective task difficulty. These results imply that in conditions without noise, the response times are faster, listening effort is lower, and task difficulty is lower when compared to the condition without noise. Moreover, switching ForwardFocus on results in a noise reduction effect resulting in faster response times, lower listening effort as well as lower task difficulty compared to the background noise condition in which ForwardFocus is switched off. Significant differences are indicated ((*) $p = 0.056$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

3.3 ERP results - sensor level

We analysed the ERPs on the sensor level by means of a 2 x 3 repeated measures ANOVA separately for the N1 and P2 peak means and latencies. In these ANOVAs, we included the within-subject factors sensory condition (AV, A) and background noise condition (BN+FF-, BN+FF+, BN-FF-).

Concerning the *N1 peak mean*, the ANOVA revealed a main effect of sensory condition ($F_{1,9} = 17.859$; $p \leq .01$, $\eta p^2 = 0.302$). Post-hoc t-test confirmed a difference between A and AV ($t(29) = -6.2$; $p \leq .001$), with AV syllables showing reduced amplitudes compared to A. The

ANOVA on the *N1 peak mean latency* resulted in a main effect of background noise condition ($F_{2,18} = 4.526$; $p \leq .05$, $\eta p^2 = 0.137$). Follow-up t-tests showed a difference between both conditions with background noise compared to the condition without background noise (BN-FF- vs. BN+FF+: $t(19) = 4.13$; $p \leq .01$; BN-FF- vs. BN+FF-: $t(19) = 3.24$; $p \leq .01$) and no difference between BN+FF+ and BN+FF- ($t(19) = 0.423$; $p = .677$). Thus, both conditions with background noise resulted in delayed N1 latencies compared to the condition without background noise, which is referred to as *noise effect* (see Figure 3B).

Regarding the *P2 peak mean*, the ANOVA showed a sensory condition x background noise condition interaction ($F_{2,18} = 4.341$; $p \leq .05$, $\eta p^2 = 0.02$). Post-hoc analyses revealed a difference between A and AV for BN-FF- ($t(9) = -5.11$; $p \leq .001$) and for BN+FF+ ($t(9) = -3.72$; $p \leq .01$), with significantly greater amplitudes for AV compared to A for these two conditions, which is referred to as *multisensory effect* (see 3D). In addition, follow-up analyses showed for A a difference between BN-FF- and the two conditions with background noise (BN-FF- vs. BN+FF+: $t(9) = -3.53$; $p \leq .05$; BN-FF- vs. BN+FF-: $t(9) = -3.29$; $p \leq .05$) and no difference between the conditions with ForwardFocus on or off with background noise ($t(9) = 1.13$; $p = .289$). Thus, the P2 amplitudes were significantly higher for the condition without background noise compared to the two conditions with background noise, which we refer to as *noise effect*. However, for AV there was a difference between all three background noise conditions: BN+FF+ vs. BN+FF- ($t(9) = -2.39$; $p \leq .05$), BN+FF- vs. BN-FF- ($t(9) = -4.83$; $p \leq .01$), and BN-FF- vs. BN+FF+ ($t(9) = -4.41$; $p \leq .01$). As a result, the P2 amplitude was the highest for the condition BN-FF-, followed by BN+FF+ and BN+FF-, which we refer to as *noise and ForwardFocus effect* (see Figure 3C). Concerning the results on the *P2 peak mean latency*, the ANOVA showed a sensory condition x background noise condition interaction ($F_{2,18} = 5.112$; $p \leq .05$, $\eta p^2 = 0.08$). However, resolving the ANOVA yielded no significant effects.

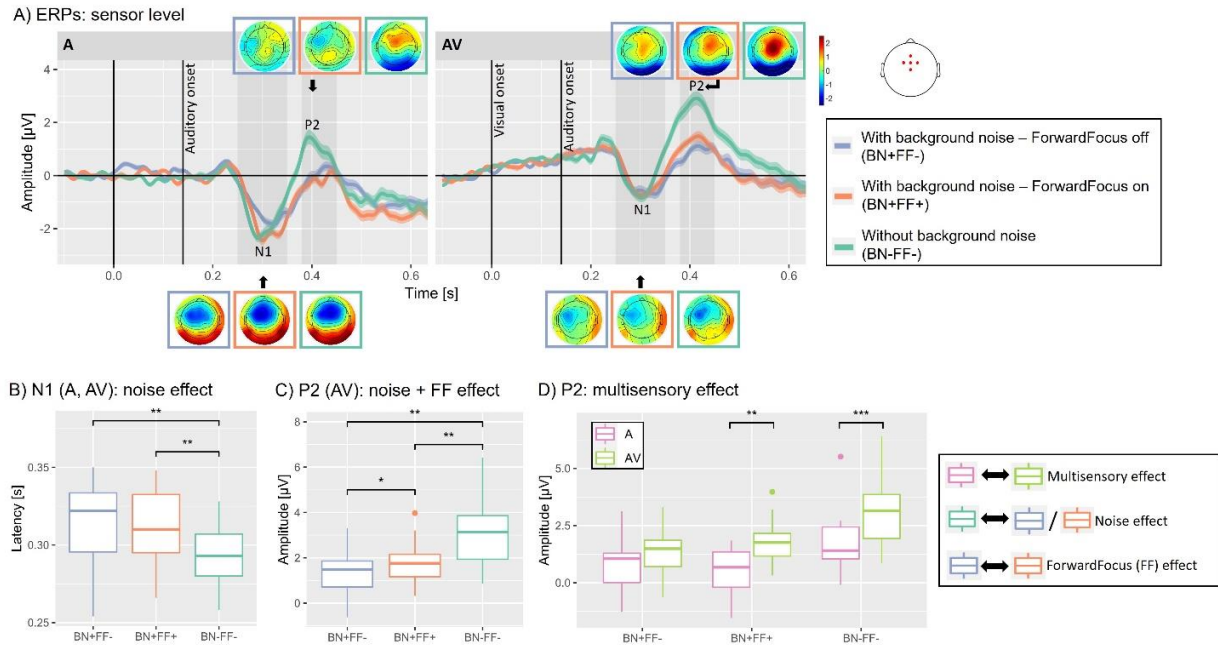


Figure 3: ERP results on the sensor level. A) Sensor ERPs for all noise conditions (BN+FF-, BN+FF+, BN-FF-) and both sensory conditions (A, AV), with shaded standard error and the corresponding topographies for each peak. The grey areas mark the time windows for the N1 and P2 peak mean detection. The peak mean latencies are given in Table 3. B) There is a *noise effect* at N1 latency, which is reflected in shortened latencies for the condition without background noise (BN-FF-) compared to the two conditions with background noise (BN+FF-, BN+FF+). C) There is both a *noise effect* and a *ForwardFocus effect* at P2 latency for AV conditions, showing the largest P2 amplitudes for BN-FF-, followed by BN+FF+ and BN+FF-. For A, there is only a *noise effect* (larger P2 amplitudes for BN-FF- compared to the two conditions with BN). D) At P2 latency, there is a *multisensory effect* for BN-FF- and BN+FF+ only, with AV having larger P2 amplitudes than A. All possible effects are summarised in the box on the right. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

3.4 ERP results - source level

We calculated for each of the five ROIs (insular cortex, lateral occipital cortex, superior temporal cortex, transverse temporal gyri, superior frontal cortex) a 2×3 repeated measures ANOVA separately for the N1 and P2 peak means and latencies, with sensory condition (AV, A) and noise condition (BN+FF-, BN+FF+, BN-FF-) as the within-subject factors.

Regarding the ANOVA for the *transverse temporal gyri*, which comprise the primary and parts of the secondary auditory cortex, we observed a main effect of background noise condition at N1 latency ($F_{2,18} = 11.714$; $p \leq .001$, $\eta p^2 = 0.143$). Post-hoc t-tests revealed a difference between the condition without background noise and the two conditions with background noise (BN-FF- vs. BN+FF+: $t(19) = 2.68$; $p \leq .05$; BN-FF- vs. BN+FF-: $t(19) = 4.07$; $p \leq .01$). However, there was no difference between the two background noise conditions (BN+FF+ vs. BN+FF-: $t(19) = 1.53$; $p = .142$). This means that the N1 is delayed for conditions with background noise compared to the condition without background noise and this is referred to as *noise effect* (see Figure 4C).

The ANOVA for the *superior temporal cortex* did reveal a main effect of sensory condition for the P2 peak mean ($F_{1,9} = 6.744$; $p \leq .05$, $\eta p^2 = 0.09$). Post-hoc t-tests confirmed a

difference between A and AV ($t(29) = -3.39$; $p \leq .01$), with larger P2 amplitudes for AV syllables compared to A syllables, which can be interpreted as *multisensory effect* as well.

Concerning the *lateral occipital cortex*, the ANOVAs again showed only main effects of sensory condition for the peak means of the N1 ($F_{1,9} = 6.344$; $p \leq .05$, $\eta p^2 = 0.17$) and the P2 ($F_{1,9} = 13.638$; $p \leq .01$, $\eta p^2 = 0.272$). No latency effects were observed. For both the N1 and the P2 amplitude, follow-up t-tests revealed that there was a difference between the activity for A and AV syllables (N1: $t(29) = -3.82$; $p \leq .001$; P2: $t(29) = -5.46$; $p \leq .001$). Thus, the activity for AV syllables was higher within the lateral occipital cortex compared to A syllables, which can be inspected in Figure 4B (P2) and is referred to as *multisensory effect*.

For the *insular cortex*, the ANOVAs revealed main effects of sensory condition for the peak means of the N1 ($F_{1,9} = 9.724$; $p \leq .05$, $\eta p^2 = 0.139$) and the P2 ($F_{1,9} = 5.572$; $p \leq .05$, $\eta p^2 = 0.12$). However, no latency effects were observed. For both the N1 and the P2 amplitude, follow-up t-tests revealed that there was a difference between the activity for A and AV syllables (N1: $t(29) = -4.55$; $p \leq .001$; P2: $t(29) = -3.25$; $p \leq .01$). Thus, the activity for AV syllables was higher within the insular cortex compared to A syllables, which can be interpreted as *multisensory effect*.

Finally, the ANOVAs for the *superior frontal cortex* revealed a main effect of sensory condition for the N1 peak mean ($F_{1,9} = 7.243$; $p \leq .05$, $\eta p^2 = 0.16$) and a main effect of background noise condition for the P2 peak mean ($F_{2,18} = 4.128$; $p \leq .05$, $\eta p^2 = 0.075$). Follow-up t-tests on the N1 amplitude revealed a difference between A and AV ($t(29) = -3.43$; $p \leq .01$), with greater amplitudes for AV syllables compared to A syllables. Concerning the post-hoc t-tests for the P2 background noise condition main effect, there was a difference between BN+FF- and BN-FF- ($t(19) = -3.436$; $p \leq .01$), and a difference between BN+FF- and BN+FF+ ($t(19) = -2.46$; $p \leq .05$). There was no difference in P2 activity between BN-FF- and BN+FF+ ($t(19) = -0.273$; $p = .788$). Thus, there was more activity within the superior frontal cortex for the conditions BN-FF- and BN+FF+ when compared to BN+FF-, which is referred to as *noise effect* and *ForwardFocus effect* (see Figure 4D).

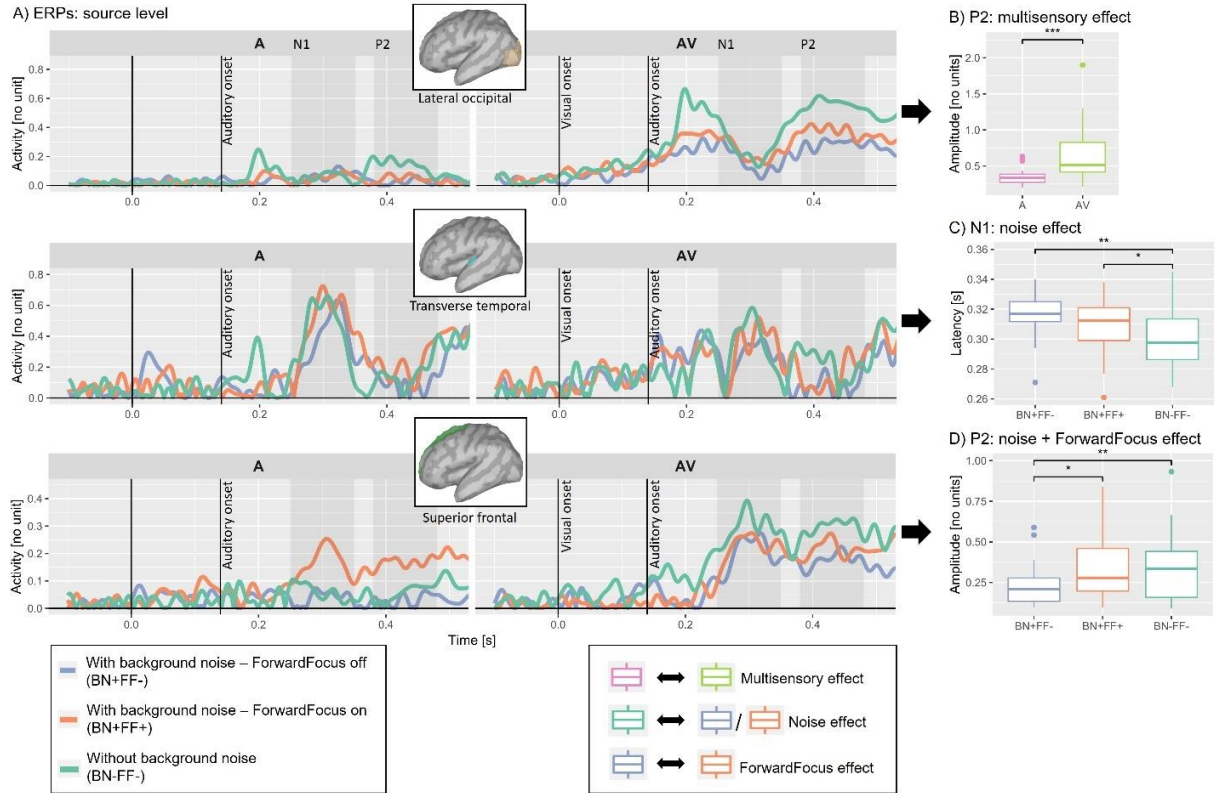


Figure 4: ERP results on the source level. A) The source waveforms for each background noise condition (BN+FF-, BN+FF+, BN-FF-) and the two sensory conditions (A, AV) are displayed for the lateral occipital cortex, the transverse temporal gyri and the superior frontal cortex (from top to down). The black arrows on the right of the source waveforms indicate that the following boxplot corresponds to the assigned ROI. The bottom-right box summarises all possible effects. B) A *multisensory effect* at P2 latency is found for the lateral occipital cortex, as reflected by greater amplitudes for AV compared to A. The same effect has been shown for the superior temporal and the insular cortices, which is why we chose the lateral occipital cortex as representative ROI to display these results. C) There is a *noise effect* at N1 latency within the transverse temporal gyri, as shown by delayed latencies for conditions with noise (BN+FF+ and BN+FF-) compared to the condition without noise (BN-FF-). D) At P2 latency, both a *noise effect* and a *ForwardFocus effect* is detected within the superior frontal cortex, as indicated by highest amplitudes for BN-FF-, followed by BN+FF+ and BN+FF-. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

3.5 Correlations

Based on previous studies that found a relationship between working memory capacity and frontal cortex activation (Rosemann & Thiel, 2018; Giroud et al., 2021; Du et al., 2016; Rudner et al., 2019) in noisy environments, we computed correlations between the SICSPAN test and the P2 activity in the superior frontal cortex. In particular, we used the total score of remembered words (see table 2 for exact scores) and calculated a Pearson correlation with the P2 activity within the superior frontal cortex for each noise condition (BN+FF-, BN+FF+, BN-FF-). The results did not reveal any correlation between these two variables, neither for the condition without background noise (BN-FF-: $p = 0.459$; $r = 0.265$), nor for the two conditions with background noise (BN+FF-: $p = 0.917$; $r = -0.041$; BN+FF+: $p = 0.213$; $r = -0.432$).

4 Discussion

In this study, we assessed the behaviour and the ERPs of bimodal CI users in response to auditory and audiovisual syllables that were presented in three different noise conditions, namely: 1) *without* background noise-ForwardFocus off (BN-FF-), 2) *with* background noise-ForwardFocus off (BN+FF-) and 3) *with* background noise-ForwardFocus on (BN+FF+). The findings revealed a variety of effects, including a *multisensory effect*, a *noise effect*, and a *ForwardFocus effect*. These effects are discussed in detail in the following two sections.

4.1 The benefit of multisensory information

We observed a variety of results, indicating the occurrence of multisensory effects. A multisensory effect means that there are differences in behavioural and/or ERP responses between unisensory (auditory) and multisensory (audiovisual) conditions. In this study, we found evidence for multisensory effects on both the behavioural and ERP level. On the behavioural level, the results showed higher hit rates and shorter response times for AV conditions compared to A conditions (Figure 2A). Regarding the ERPs on the sensor level (channels on the scalp), we observed enhanced P2 amplitudes for audiovisual (AV) syllables compared to auditory-only (A) syllables in conditions without background noise (BN-FF-) and in conditions with background noise and activated ForwardFocus (BN+FF+; Figure 3D). Finally, the ERP source analyses (estimation of cortical ERP generators) revealed enhanced amplitudes in the audiovisual (AV) compared to the auditory-only (A) condition, both for the N1 ERP (lateral occipital cortex and insular cortex; see section 3.4. ERP results – source level) and the P2 ERP (lateral occipital cortex, insular cortex, and superior temporal cortex; see Figure 4B and section 3.4. ERP results – source level).

Behavioural enhancements in multisensory compared to unisensory conditions have previously been shown not only in normal-hearing listeners (Mahoney et al., 2011; Molholm et al., 2002; Radecke et al., 2022) but also in individuals with mild-to-moderate hearing loss (Rosemann & Thiel, 2018; Tye-Murray et al., 2007) and in CI users (Schierholz et al., 2015; Layer et al., 2022; Radecke et al., 2022). Our results of higher hit rates and shortened response times for audiovisual compared to auditory-only conditions confirm previous studies with CI users, reporting audiovisual benefits for basic stimuli (sinusoidal tones and white discs; Schierholz et al., 2015) and syllables (Layer et al., 2022). By extension, for word recognition, it has been observed that CI users show higher accuracies in identifying audiovisual words compared to auditory-only words (Rouger et al., 2007). Given that spoken language communication is an audiovisual experience (Drijvers & Özyürek, 2017; Grant et al., 1998; Sumbly & Pollack, 1954) in which auditory (speech) and visual information (facial expressions, articulatory movements, gestures) must be integrated, it appears natural that audiovisual stimuli are identified significantly faster and more accurately compared to unisensory auditory and visual stimuli. The CI users' enhancement in audiovisual performance can be explained by the "principle of inverse effectiveness", according to which the effectiveness of multisensory enhancement increases proportionally with progressive degradation of unisensory signals (Stein & Meredith, 1993). This can be directly transferred to CI users because the auditory signal, which is typically the most reliable signal in speech, is degraded when compared to normal-hearing listeners (Drennan & Rubinstein, 2008). As visual speech in general is already less informative than auditory speech, inverse

effectiveness predicts that reduced auditory performance with a CI enhances the probability of a multisensory benefit (Stevenson et al., 2017).

The multisensory improvement observed in the behavioural data is also reflected in significantly enhanced P2 amplitudes for conditions *with* (BN+FF+) and *without* background noise (BN-FF-) (Figure 3D). Thus, the CI users show a multisensory increase in P2 amplitude both in easier listening conditions (speech without background noise) and in the more difficult (speech-in-noise) listening condition with the activated ForwardFocus. Interestingly, the multisensory P2 effect was less pronounced for the perceptually most difficult stimulus condition “*with* background noise-ForwardFocus off” (BN+FF-), suggesting that the multisensory increase in P2 amplitude is reduced in conditions with poorer quality of the auditory perceptual input. Several reasons may account for the condition-specific pattern of multisensory P2 amplitude modulation. First, the P2 ERP component (Crowley & Colrain, 2004) has been linked to higher-order perceptual processing, including attention (Luck & Hillyard, 1994). Given that the allocation of attention is necessary for the identification and classification of stimuli, the current finding of enhanced P2 responses for multisensory conditions, compared to auditory-only conditions, may reflect the improved detection of informative stimuli (García-Larrea et al., 1992). Following this line of argumentation, the observation of increased audiovisual P2 amplitudes for easy (BN-FF-) and relieved (BN+FF+) speech conditions suggest that the CI users can better allocate their attention to the informative features of the stimuli in the absence of noise or by activating ForwardFocus in noise. However, this reason cannot fully explain the condition-specific P2 amplitude modulation, given that behavioural results revealed no differences in the multisensory effect between the three noise conditions (BN-FF-, BN+FF-, BN+FF+).

A second reason for the condition-specific pattern of P2 amplitude modulation could be related to the distribution of attention between the visual and auditory modality (Mishra & Gazzaley, 2012). In this line of reasoning, the P2 amplitude is decreased in the case of divided attention. Without background noise, CI users can allocate their attention towards the informative auditory input in audiovisual conditions. However, as soon as noise is added (BN+FF-), CI users divide their attention between the auditory and visual input of the audiovisual syllable, as the auditory information is no longer intelligible enough. This noise-induced effect on attention distribution, which may result in the reduction of the multisensory P2 amplitude, however, appears to be less pronounced when ForwardFocus is switched on (BN+FF+), because the auditory features seem to be still informative enough (due to the noise reduction algorithm), and only a smaller amount of visual information needs to be considered. Thus, due to the improved auditory input by means of ForwardFocus, the CI users may be less forced to divide their attention between the two modalities, which can explain the enhanced multisensory P2 amplitude in the BN+FF+ compared to the BN+FF- condition. It seems reasonable that additional visual input can inform the interpretation of ambiguous auditory syllables that are compromised by noise. In post-lingually deafened CI users, it has been proposed that an increased attentional allocation to visual information can improve the perception of the impoverished auditory CI input (Butera et al., 2018). To sum up, when compared to the condition without background noise, adding background noise might increase cross-modal divided attention and consequently a smaller P2 amplitude is observed. Activating ForwardFocus (BN+FF+), however, seems to reduce the effect of cross-

modal divided attention in speech-in-noise conditions, which leads to greater audiovisual P2 amplitudes when compared to ForwardFocus switched off (BN+FF-).

A third explanatory approach for the condition-specific pattern of P2 amplitude modulation is that the spatial congruency between the auditory and visual signal is harmed by background noise. The cross-modal information is the most congruent in conditions without background noise (BN-FF-), as the sound and the corresponding visual cues come from the same direction (i.e., one speaker from the front), which is reflected by enhanced P2 amplitudes. In conditions *with* background noise, the additional background noise (from lateral speakers) disturbs the congruency, as it is additionally present from the left and the right sides. Among the speech-in-noise conditions, ForwardFocus (BN+FF+) reduces the noise-induced disturbance, compared to ForwardFocus off (BN+FF-). Thus, a greater P2 amplitude may reflect a stronger perceived congruency of auditory and visual information (Radecke et al., 2022), which, in the current experiment, rather reflects a spatial congruency, as ForwardFocus is a spatially selective algorithm. Similarly, a recent study reported that CI users showed an improved perception of environmental sounds when these ambiguous sounds were presented together with congruent visual context (Radecke et al., 2022). Importantly, the CI users' P2 amplitudes were increased for congruent when compared to non-congruent audiovisual stimuli (Radecke et al., 2022), further supporting our interpretation that the P2 amplitude depends on the perceived congruency of auditory and visual input. Taken together, several reasons may account for our observation that the multisensory P2 amplitude is affected by the presence of background noise and the activation of ForwardFocus. We assume that the increased P2 amplitudes in the BN-FF- (*without* background noise-ForwardFocus off) and BN+FF+ conditions (*with* background noise-ForwardFocus on) might indicate the enhanced detection of informative stimuli, the less divided attention between the auditory and visual modality (because of the highly informative, less ambiguous auditory input), and/or the enhanced perceived (spatial) congruency of auditory and visual input.

Consistent with our behavioural observations and the ERP results on the sensor level, our results of the ERP source analysis revealed a multisensory effect in different cortical regions. As expected, we found an increase in amplitude within the *superior temporal cortex* for AV syllables compared to A syllables (see section 3.4. ERP results – source level). This finding can be explained by the important role the superior temporal lobe plays for multisensory processing (Calvert, 2001; Beauchamp, 2005; Driver & Noesselt, 2008; Noesselt et al., 2007). Previous research has identified convergence zones in various brain regions, among them the superior temporal sulcus, where neurons receive and combine afferent inputs from multiple sensory modalities (Driver & Noesselt, 2008). It has been assumed that activation in multisensory regions may modulate the response in sensory-specific brain areas by top-down effects (Driver & Noesselt, 2008). In addition, animal studies have reported up to 20 % visually sensitive neurons in the primary auditory cortex (Bizley & King, 2009), and they have suggested direct connections between unimodal sensory cortices (Cappe & Barone, 2005). Concerning CI users, our results confirm multisensory interactions in the superior temporal lobe by showing enhanced activation for AV syllables compared to auditory-only syllables (see section 3.4. ERP results – source level). Previous studies have suggested that these multisensory interactions in the auditory cortex are even enhanced when compared to normal-hearing listeners (Strelnikov et al., 2013; Giraud et al., 2001; Layer et al., 2022;

Schierholz et al., 2015). The strong binding between the auditory and the visual sensory system in CI users may be a result of sensory deprivation (prior to CI implantation) and/or the degraded CI signal after implantation (Schierholz et al., 2017), and might facilitate auditory rehabilitation after implantation by allowing enhanced audio-visual integration in these individuals (Stropahl & Debener, 2017; Radecke et al., 2022).

In addition to the superior temporal cortex, we found a multisensory effect in the lateral occipital cortex and in the insular cortex (see Figure 4B and section 3.4. ERP results – source level), as indicated by increased amplitudes for AV conditions compared to A conditions in both cortical regions. Regarding the *lateral occipital cortex*, the increased response in conditions with additional visual information likely reflects the processing of the visual component of AV syllables. Indeed, the lateral occipital cortex is known to be recruited not only for visual object recognition (Grill-Spector et al., 2001) but also during visual speech processing (Weiner & Grill-Spector, 2013). Concerning the *insular cortex*, the pronounced activity for AV compared to A syllables confirms previous observations that the insula is involved in the processing of unisensory auditory and audiovisual stimuli (Chen et al. 2015; Eckert et al., 2009; Sterzer & Kleinschmidt, 2010). Previous results suggest that the insula plays a role in perception by mediating states of enhanced sensory alertness and allocating multisensory attention (for review see Nieuwenhuys, 2012; Eckert et al., 2009; Sterzer & Kleinschmidt, 2010). Our results are in line with these reports by showing that CI users, when tested with a speeded syllables-identification task that is performed in conditions with high alertness, show a multisensory modulation of activation within the insular cortex. Importantly, our results suggest that this multisensory enhancement of insular activation can already be observed in (relatively) early steps of cortical speech processing, in particular at N1 and P2 latency.

Taken together, our behavioural and ERP results of CI users provide evidence that multisensory stimuli are identified more quickly and more accurately when compared to unisensory auditory stimuli. The neuronal correlate of this behavioural multisensory enhancement is the increased P2 amplitude (sensor level; BN+FF+, BN-FF-), and the enhanced activation at P2 latency within the superior temporal cortex, the lateral occipital cortex and the insular cortex (source level; BN-FF-, BN+FF-, BN+FF+). These findings indicate that CI users, when exposed to audiovisual speech conditions with and without background noise, show audiovisual interactions in expanded cortical regions.

4.2 Noise reduction via ForwardFocus affects allocation of attentional resources

Our results revealed general effects of background noise (BN+FF-/BN+FF+ vs. BN-FF-). On the behavioural level, the presence of background noise resulted in prolonged response times (Figure 2B), enhanced subjective listening effort and higher task difficulty (Figure 2B). Regarding the ERPs, background noise conditions exhibited delayed N1 latency (Figure 3B) and reduced P2 amplitude (Figure 3C) compared to conditions without background noise (BN+FF-/BN+FF+ vs. BN-FF-). Similar results were observed in the ERP source analysis, showing a noise-induced delay in N1 response within the transverse temporal gyri, also referred to as auditory cortex (BN+FF-/BN+FF+ > BN-FF-; Figure 4C), and a reduction in P2 amplitude within the superior frontal cortex (BN+FF- < BN+FF+/BN-FF-; Figure 4D).

Nevertheless, we found ForwardFocus-specific effects both in the behavioural and the ERP results (BN+FF- vs. BN+FF+). On the behavioural level, we found shorter response times (Figure 2B), reduced subjective listening effort, lower task difficulty (Figure 2B), and better speech intelligibility (see section 3.2 Other behavioural results) when ForwardFocus was switched on (Figure 2B). Benefits of using ForwardFocus were confirmed by objective measures. In particular, the P2 ERP amplitude (sensor level; Figure 3D) and the activation in the superior frontal cortex (ERP source analysis; Figure 4D) showed a multisensory effect (AV > A) which was found specifically for the condition “ForwardFocus switched on (BN+FF+)” but not for the condition “ForwardFocus switched off (BN+FF-)” (Figure 3D).

Our findings regarding the background-noise effect are consistent with previous results in normal-hearing listeners (e.g., Billings et al., 2009), individuals with mild to moderate hearing impairment (Gillis et al., 2022) and CI users (e.g., Finke et al., 2016; Han & Dimitrijevic, 2020), suggesting delayed cortical responses for conditions with compared to without background noise. In addition to these objective measures, our results confirm previous behavioural observations, suggesting that CI users are very susceptible to noisy environments (Davidson et al., 2010; Fu & Nogaki, 2005) in which they experience a high listening effort (Hughes & Galvin, 2013; Winn, 2016). Background noise in general limits the capacity to understand speech, given that the noise can interfere physically (energetic masking) and perceptually (informational masking due competing talkers) with the speech signal (Pollack, 1975). This competition between speech and noise signals makes it harder to extract the speech information (Helfer & Vargo, 2009; Romei et al., 2011), and requires higher cognitive processes, such as attention and memory, in order to understand the meaning of the spoken language (Kent, 1992). The presented data support the view that CI users have particular difficulties to discriminate speech sounds in conditions with interfering background noise. Our study using a speeded syllables-identification task suggests that the presence of additional noise results in slower response times, enhanced subjective listening effort, higher task difficulty, and delayed cortical responses in the auditory cortex.

Despite these negative effects of background noise, we found positive effects of noise reduction via ForwardFocus. Compared to BN+FF-, activating ForwardFocus (BN+FF+) improved the response times, reduced the subjective listening effort, and decreased the perceived task difficulty. In addition, like Hey et al. (2019, 2021), we found improved speech intelligibility in the clinical OLSA speech-in-noise test when ForwardFocus was switched on. Thus, our study provides additional evidence of successful noise reduction by ForwardFocus in an experimental setting with syllables presented from the front and stationary noise from the sides. The novelty of this study is that we demonstrate beneficial effects of ForwardFocus not only by means of subjective but also by *objective* measures. In particular, we found that the P2 ERP amplitude for AV syllables at P2 latency and the activation within the superior frontal cortex were enhanced for the condition “ForwardFocus switched on (BN+FF+)” when compared to the condition “ForwardFocus switched off (BN+FF-)”. One might argue that this activation effect within the superior frontal cortex is at odds because we found the *strongest* activation for the *easiest* condition (BN+FF-, followed by BN+FF+ and BN+FF-), whereas previous studies have reported *pronounced* frontal lobe recruitment for *difficult* listening situations requiring *enhanced* listening effort (Campbell & Sharma, 2013; Rosemann & Thiel, 2018; Wong et al., 2009). A pronounced engagement of frontal brain regions has been associated with the recruitment of cognitive mechanisms, such as working memory, which

seem to be necessary to explicitly process the ambiguous auditory input in effortful speech-in-noise conditions (Rosemann & Thiel, 2018; Giroud et al., 2021; Du et al., 2016; Rudner et al., 2019). However, preceding studies have assigned these cognitive processes (in demanding listening conditions) to other brain regions than in the current study, namely the inferior frontal cortex, which could explain the discrepancy between the current results and those of previous studies. Moreover, in our study we found no significant correlation between the working memory capacity (SICSPAN) and the P2 response amplitude in the superior frontal cortex (see section 3.5. Correlations). This strengthens our interpretation that the *reduced* activation in the superior frontal cortex, which was found for the *most difficult* noise condition (BN+FF-), cannot be directly explained by altered recruitment of cognitive working memory processes.

An alternative explanation for the different activation in the superior frontal cortex between ForwardFocus off (BN+FF-) and ForwardFocus on (BN+FF+) might relate to the allocation of attentional resources. Our P2 ERP source analysis revealed a generator in the posterior part of the superior frontal cortex (Supplementary Fig. 2), a region that has been reported to be involved in audiovisual divided attention (Vohn et al., 2007). The same brain region incorporates parts of the supplementary eye field (SEF) and the frontal eye field (FEF). Both areas have been frequently related to visuo-spatial attention (Marshall et al., 2015; Kastner et al., 1999; Corbetta & Shulman, 2002). Furthermore, it has been proposed that visual information about a speaker's spatial position can be used as a spatial cue to facilitate auditory speech processing in the presence of distractions (Haykin & Chen, 2005; Kidd et al., 2005). Therefore, visual inputs are critical for multisensory speech recognition, especially in noisy environments (Senkowski et al., 2008). With regards to our study, we speculate that CI users – when tested in audiovisual conditions with interfering noise and deactivated ForwardFocus (BN+FF-) - take visual stimulus features more into account, and that they increasingly divide their attention between auditory and visual stimulus features (Butera et al., 2018; García-Larrea et al., 1992) when compared to conditions with ForwardFocus switched on (BN+FF+). In addition, the presence of noise in conditions with ForwardFocus off (BN+FF-) may limit the CI users' ability to focus on the syllables, and it might compromise the spatial specificity of the perceived auditory speech signal (in front) due to distracting noise from the left and right side. However, when ForwardFocus is switched on, the noise-induced distraction may be reduced, which might result in enhanced spatial specificity of the perceived auditory speech signal (A condition), and in stronger perceived spatial congruency of auditory and visual stimuli (AV condition; compare Radecke et al., 2022). Thus, our result of an enhanced P2 amplitude in speech-in-noise conditions for ForwardFocus on (BN+FF+) compared to ForwardFocus off (BN+FF-) can be explained by the allocation of attentional resources, in particular the less divided attention between auditory and visual stimulus features (because ForwardFocus allows more informative auditory input), and the stronger perceived spatial congruency of auditory and visual stimuli (AV condition). In other words, our results indicate that ForwardFocus can reduce the competing influence of noise, such that the CI users can better focus on the syllables and allocate their attention to the relevant auditory speech cues. This effect is supported by the fact that we did not find a significant activation difference in the superior frontal cortex between the condition without background noise (BN-FF-) and the condition with background noise and activated ForwardFocus (BN+FF+), implying that in these easier listening conditions, the CI users have better perceived spatial congruency of auditory and visual stimuli compared to the more

difficult listening condition with background noise and deactivated ForwardFocus (BN+FF-). Thus, our results suggest that an activated ForwardFocus in combination with the availability of audiovisual speech cues can help to focus the attentional resources on the informative features of the syllables in situations with competing background noise.

To sum up, the results confirm noise-induced alterations in behavioural and ERP results, as evidenced by prolonged response times, enhanced listening effort, and a delayed auditory-cortex activation (N1 latency). These findings support previous conclusions that the N1-P2 complex is a helpful tool to better understand the impact of background noise on speech processing (Billings et al., 2011; Obleser & Kotz, 2011). By extension, our study shows that the objective tool of the assessment of the N1-P2 complex allows the evaluation of noise reduction algorithms. Apart from the frequently reported noise effects, we provide both behavioural and electrophysiological evidence of successful noise reduction via ForwardFocus. Activating ForwardFocus results in shorter response times, reduced listening effort, lower task difficulty, better speech intelligibility, and an enhanced activation within the superior frontal cortex at P2 latency. We assume that these effects of ForwardFocus can be explained by attentional resources that can be allocated more effectively when multisensory information and activated ForwardFocus are combined. To summarise, this study provides both subjective and objective evidence that the use of ForwardFocus in situations with interfering noise (coming from distinct sources than the auditory signal) is highly beneficial for CI users in ecologically valid audiovisual speech conditions.

6 Conclusion

The present study shows multisensory benefits in CI users and suggests that these individuals – when exposed to audiovisual speech conditions with and without background noise – show audiovisual interactions in expanded cortical regions, in particular the superior temporal cortex, the lateral occipital cortex and the insular cortex. The study also provides supportive evidence that speech processing in CI users is hampered by background noise, leading to prolonged response times, enhanced listening effort, higher task difficulty, and a delayed activation in the auditory cortex at N1 latency, both for auditory and audiovisual conditions. Nevertheless, the results confirm benefits of using ForwardFocus in background noise, both on the subjective and the objective ERP level, as evidenced by shorter response times, reduced listening effort, lower task difficulty, better speech intelligibility, and an enhanced activation within the superior frontal cortex in audiovisual speech conditions at P2 latency. We propose that the combination of available multisensory information (AV syllables) and activated noise reduction via ForwardFocus results in a reallocation of attentional resources, leading to enhanced cortical activation at P2 latency, even when an interfering noise is present in addition to the relevant speech signal. This is of high clinical relevance, because the results can help optimising therapeutic procedures in the long term (for example by developing audiovisual training programs) and promoting the usage of ForwardFocus in noisy, real world listening situations.

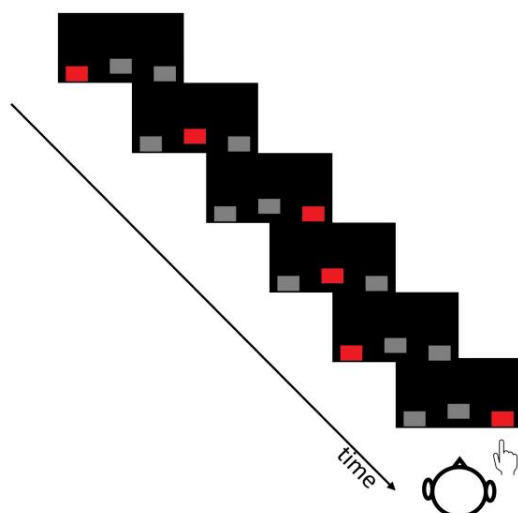
Acknowledgements

We thank Cochlear Research & Development Ltd (reference number: IIR-2306) for providing funding for this research. Especially, we thank Horst Hessel for his helpful comments and suggestions for the project and the manuscript. We also want to thank Pauline Burkhardt and Lina Wiesel for their assistance in data collection. Finally, we thank all participants for taking part in this study.

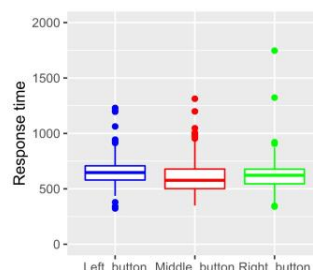
Supplementary Figures

Behavioural control task

A) Paradigm

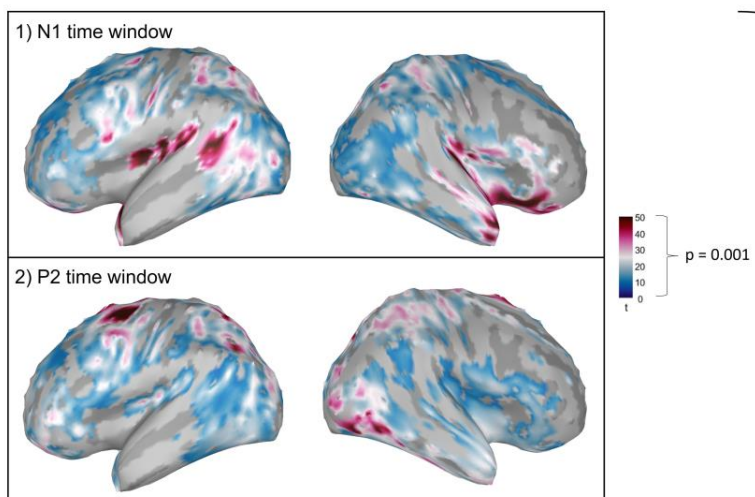


B) Results: RTs

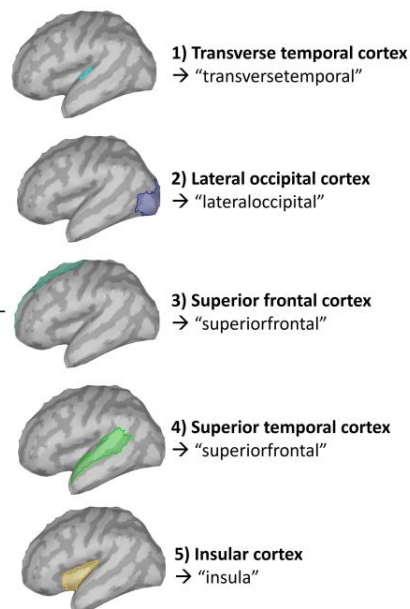


Supplementary Figure 1: Behavioural control task. A) Paradigm of the control task. There are three buttons on the touchscreen and the task is to push the red button, of which the location pseudo randomly varied as fast as possible. B) Results: there is no significant difference in RTs between the location of the red button. Significant differences are indicated (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

A) Results from global t-test:



B) Selected ROIs:



Supplementary Figure 2: Selection of ROIs for source analysis. A) The maximal activity for the grand averaged data (averaged over noise and sensory conditions) is displayed for the N1 (1) and P2 (2) time windows. The maximal activity has been calculated by means of an inbuilt t-test against the baseline directly in Brainstorm (Bonferroni-corrected, p-value threshold: $p = 0.001$). B) Based on this maximal activity, we chose the best-fitting ROIs within the "Desikan-Killiany" atlas in Brainstorm, which we used for follow-up statistical analyses.

References

- Alain, C., Du, Y., Bernstein, L. J., Barten, T., & Banai, K. (2018). Listening under difficult conditions: An activation likelihood estimation meta-analysis. *Human Brain Mapping, 39*(7), 2695–2709.
- Bavelier, D., Dye, M. W., & Hauser, P. C. (2006). Do deaf individuals see better? *Trends in Cognitive Sciences, 10*(11), 512–518.
- Beauchamp, M. S. (2005). See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology, 15*(2), 145–153.
- Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation, 7*(6), 1129–1159.
- Bertoli, S., Smurzynski, J., & Probst, R. (2005). Effects of age, age-related hearing loss, and contralateral cafeteria noise on the discrimination of small frequency changes: Psychoacoustic and electrophysiological measures. *Journal of the Association for Research in Otolaryngology, 6*(3), 207–222.
- Biasiucci, A., Franceschiello, B., & Murray, M. M. (2019). Electroencephalography. *Current Biology, 29*(3), R80–R85.
- Billings, C. J., Bennett, K. O., Molis, M. R., & Leek, M. R. (2011). Cortical encoding of signals in noise: Effects of stimulus type and recording paradigm. *Ear and Hearing, 32*(1), 53.
- Billings, C. J., Tremblay, K. L., Stecker, G. C., & Tolin, W. M. (2009). Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hearing Research, 254*(1-2), 15–24.
- Bizley, J. K., & King, A. J. (2009). Visual influences on ferret auditory cortex. *Hearing research, 258*(1-2), 55–63.
- Bottari, D., Bednaya, E., Dormal, G., Villwock, A., Dzhelyova, M., Grin, K., Pietrini, P., Ricciardi, E., Rossion, B., & Röder, B. (2020). EEG frequency-tagging demonstrates increased left hemispheric involvement and crossmodal plasticity for face processing in congenitally deaf signers. *NeuroImage, 223*, 117315.
- Butera, I. M., Stevenson, R. A., Mangus, B. D., Woynaroski, T. G., Gifford, R. H., & Wallace, M. T. (2018). Audiovisual temporal processing in postlingually deafened adults with cochlear implants. *Scientific Reports, 8*(1), 1–12.
- Calvert, G. A. (2001). Crossmodal processing in the human brain: insights from functional neuroimaging studies. *Cerebral cortex, 11*(12), 1110–1123.
- Campbell, J., & Sharma, A. (2013). Compensatory changes in cortical resource allocation in adults with hearing loss. *Frontiers in Systems Neuroscience, 7*, 71.
- Cappe, C., & Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. *European Journal of Neuroscience, 22*(11), 2886–2902.

- Cardin, V., Orfanidou, E., Rönnerberg, J., Capek, C. M., Rudner, M., & Woll, B. (2013). Dissociating cognitive and sensory neural plasticity in human superior temporal cortex. *Nature Communications*, 4(1), 1–5.
- Chen, T., Michels, L., Supekar, K., Kochalka, J., Ryali, S., & Menon, V. (2015). Role of the anterior insular cortex in integrative causal signaling during multisensory auditory–visual attention. *European Journal of Neuroscience*, 41(2), 264–274.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201–215.
- Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, 115(4), 732–744.
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, 26(1), 55–67.
- Davidson, L. S., Geers, A. E., & Brenner, C. (2010). Cochlear implant characteristics and speech perception skills of adolescents with long-term device use. *Otology & Neurotology: Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 31(8), 1310.
- Debener, S., Hine, J., Bleack, S., & Eyles, J. (2008). Source localization of auditory evoked potentials after cochlear implantation. *Psychophysiology*, 45(1), 20–24.
- Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9–21.
- Desai, S., Stickney, G., & Zeng, F.-G. (2008). Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. *The Journal of the Acoustical Society of America*, 123(1), 428–440.
- Downs, D. W. (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders*, 47(2), 189–193.
- Drennan, W. R., & Rubinstein, J. T. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *Journal of Rehabilitation Research and Development*, 45(5), 779.
- Drijvers, L., & Özyürek, A. (2017). Visual context enhanced: The joint contribution of iconic gestures and visible speech to degraded speech comprehension. *Journal of Speech, Language, and Hearing Research*, 60(1), 212–222.
- Driver, J., & Noesselt, T. (2008). Multisensory interplay reveals crossmodal influences on ‘sensory-specific’ brain regions, neural responses, and judgments. *Neuron*, 57(1), 11–23.

- Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2016). Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nature Communications*, 7(1), 1–12.
- Dürschmid, S., Reichert, C., Hinrichs, H., Heinze, H.-J., Kirsch, H. E., Knight, R. T., & Deouell, L. Y. (2019). Direct evidence for prediction signals in frontal cortex independent of prediction error. *Cerebral Cortex*, 29(11), 4530–4538.
- Eckert, M. A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., & Dubno, J. R. (2009). At the heart of the ventral attention system: the right anterior insula. *Human brain mapping*, 30(8), 2530–2541.
- Elliott, R. (2003). Executive functions and their disorders: Imaging in clinical neuroscience. *British Medical Bulletin*, 65(1), 49–59.
- Fetterman, B. L., & Domico, E. H. (2002). Speech recognition in background noise of cochlear implant patients. *Otolaryngology-Head and Neck Surgery*, 126(3), 257–263.
- Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., & Sandmann, P. (2016). On the relationship between auditory cognition and speech intelligibility in cochlear implant users: An ERP study. *Neuropsychologia*, 87, 169–181.
- Fu, Q.-J., & Nogaki, G. (2005). Noise susceptibility of cochlear implant users: The role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology*, 6(1), 19–27.
- García-Larrea, L., Lukaszewicz, A.-C., & Mauguière, F. (1992). Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia*, 30(8), 723–741.
- Gillis, M., Decruy, L., Vanthornhout, J., & Francart, T. (2022). Hearing loss is associated with delayed neural responses to continuous speech. *European Journal of Neuroscience*, 55(6), 1671–1690.
- Giraud, A.-L., Price, C. J., Graham, J. M., Truy, E., & Frackowiak, R. S. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, 30(3), 657–664.
- Giroud, N., Keller, M., & Meyer, M. (2021). Interacting effects of frontal lobe neuroanatomy and working memory capacity to older listeners' speech recognition in noise. *Neuropsychologia*, 158, 107892.
- Glennon, E., Svirsky, M. A., & Froemke, R. C. (2020). Auditory cortical plasticity in cochlear implant users. *Current Opinion in Neurobiology*, 60, 108–114.
- Gosselin, P. A., & Gagne, J. P. (2011). Older adults expend more listening effort than young adults recognizing speech in noise.
- Gramfort, A., Papadopoulos, T., Olivi, E., & Clerc, M. (2010). OpenMEEG: Opensource software for quasistatic bioelectromagnetics. *Biomedical Engineering Online*, 9(1), 1–20.

- Grant, K. W., Walden, B. E., & Seitz, P. F. (1998). Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *The Journal of the Acoustical Society of America*, 103(5), 2677–2690.
- Green, K., Ramsden, R., Julyan, P., & Hastings, D. (2008). Cortical plasticity in the first year after cochlear implantation. *Cochlear Implants International*, 9(2), 103–117.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41(10-11), 1409–1422.
- Hahlbrock, K.-H. (1970). *Sprachaudiometrie: Grundlagen und praktische Anwendung einer Sprachaudiometrie für das deutsche Sprachgebiet*. Georg Thieme Verlag.
- Han, J.-H., & Dimitrijevic, A. (2020). Acoustic change responses to amplitude modulation in cochlear implant users: Relationships to speech perception. *Frontiers in Neuroscience*, 14, 124.
- Hansen, P., Kringelbach, M., & Salmelin, R. (2010). *MEG: An introduction to methods*. Oxford university press.
- Hari, R., Pelizzzone, M., Mäkelä, J., Hällström, J., Leinonen, L., others. (1987). Neuromagnetic responses of the human auditory cortex to on-and offsets of noise bursts. *Audiology*, 26, 31–43.
- Haykin, S., & Chen, Z. (2005). The cocktail party problem. *Neural computation*, 17(9), 1875–1902.
- Helfer, K. S., & Vargo, M. (2009). Speech recognition and temporal processing in middle-aged women. *Journal of the American Academy of Audiology*, 20(04), 264–271.
- Hey, M., Böhnke, B., Mewes, A., Munder, P., Mauger, S. J., & Hocke, T. (2021). Speech comprehension across multiple CI processor generations: Scene dependent signal processing. *Laryngoscope Investigative Otolaryngology*, 6(4), 807–815.
- Hey, M., Hocke, T., Böhnke, B., & Mauger, S. J. (2019). ForwardFocus with cochlear implant recipients in spatially separated and fluctuating competing signals—introduction of a reference metric. *International Journal of Audiology*, 58(12), 869–878.
- Hey, M., Hocke, T., Mauger, S., & Müller-Deile, J. (2016). A clinical assessment of cochlear implant recipient performance: Implications for individualized map settings in specific environments. *European Archives of Oto-Rhino-Laryngology*, 273(11), 4011–4020.
- Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., Gotter, B. D., Vanderhoof, S. S., Mispagel, K., Heydebrand, G., others. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear and Hearing*, 34(3), 342.
- Holder, J. T., Reynolds, S. M., Sunderhaus, L. W., & Gifford, R. H. (2018). Current profile of adults presenting for preoperative cochlear implant evaluation. *Trends in hearing*, 22, 2331216518755288.

- Hoppe, U., Hocke, T., & Digeser, F. (2018). Bimodal benefit for cochlear implant listeners with different grades of hearing loss in the opposite ear. *Acta Oto-Laryngologica*, 138(8), 713–721.
- House, W. F., & Urban, J. (1973). Long term results of electrode implantation and electronic stimulation of the cochlea in man. *Annals of Otology, Rhinology & Laryngology*, 82(4), 504–517.
- Hughes, K. C., & Galvin, K. L. (2013). Measuring listening effort expended by adolescents and young adults with unilateral or bilateral cochlear implants or normal hearing. *Cochlear Implants International*, 14(3), 121–129.
- Humes, L. E. (2019). The world health organization's hearing-impairment grading system: An evaluation for unaided communication in age-related hearing loss. *International Journal of Audiology*, 58(1), 12–20.
- Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., Mckeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2), 163–178.
- Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences*, 97(22), 11793–11799.
- Kaiser, A. R., Kirk, K. I., Lachs, L., & Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 46(2), 390–404.
- Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22(4), 751–761.
- Kent R. 1992. Auditory processing of speech. Katz J., Stecker N.A., Henderson D., Central Auditory Processing: A Transdisciplinary View. St. Louis: Mosby, 93–106.
- Khalighinejad, B., Patel, P., Herrero, J. L., Bickel, S., Mehta, A. D., & Mesgarani, N. (2021). Functional characterization of human heschl's gyrus in response to natural speech. *Neuroimage*, 235, 118003.
- Kidd Jr, G., Arbogast, T. L., Mason, C. R., & Gallun, F. J. (2005). The advantage of knowing where to listen. *The Journal of the Acoustical Society of America*, 118(6), 3804–3815.
- Klamer, S., Elshahabi, A., Lerche, H., Braun, C., Erb, M., Scheffler, K., & Focke, N. K. (2015). Differences between MEG and high-density EEG source localizations using a distributed source model in comparison to fMRI. *Brain Topography*, 28(1), 87–94.
- Krueger, B., Joseph, G., Rost, U., Strau-Schier, A., Lenarz, T., & Buechner, A. (2008). Performance groups in adult cochlear implant users: Speech perception results from 1984 until today. *Otology & Neurotology*, 29(4), 509–512.

- Layer, N., Weglage, A., Müller, V., Meister, H., Lang-Roth, R., Walger, M., Murray, M. M., & Sandmann, P. (2022). The timecourse of multisensory speech processing in unilaterally stimulated cochlear implant users revealed by ERPs. *NeuroImage: Clinical*, 34, 102982.
- Lin, F.-H., Witzel, T., Ahlfors, S. P., Stufflebeam, S. M., Belliveau, J. W., & Hämäläinen, M. S. (2006). Assessing and improving the spatial accuracy in MEG source localization by depth-weighted minimum-norm estimates. *Neuroimage*, 31(1), 160–171.
- Luck, S. J., & Hillyard, S. A. (1994). Electrophysiological correlates of feature analysis during visual search. *Psychophysiology*, 31(3), 291–308.
- Mahoney, J. R., Li, P. C. C., Oh-Park, M., Verghese, J., & Holtzer, R. (2011). Multisensory integration across the senses in young and old adults. *Brain Research*, 1426, 43–53.
- Marshall, T. R., O'Shea, J., Jensen, O., & Bergmann, T. O. (2015). Frontal eye fields control attentional modulation of alpha and gamma oscillations in contralateral occipitoparietal cortex. *Journal of Neuroscience*, 35(4), 1638–1647.
- Mauger, S. J., Warren, C. D., Knight, M. R., Goorevich, M., & Nel, E. (2014). Clinical evaluation of the nucleus 6 cochlear implant system: Performance improvements with SmartSound iQ. *International Journal of Audiology*, 53(8), 564–576.
- McCormack, A., & Fortnum, H. (2013). Why do people fitted with hearing aids not wear them? *International Journal of Audiology*, 52(5), 360–368.
- Merzenich, M. M., Van Vleet, T. M., & Nahum, M. (2014). Brain plasticity-based therapeutics. *Frontiers in Human Neuroscience*, 8, 385.
- Michel, C. M., & Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *Neuroimage*, 61(2), 371–385.
- Mishra, J., & Gazzaley, A. (2012). Attention distributed across sensory modalities enhances perceptual performance. *Journal of Neuroscience*, 32(35), 12294–12302.
- Mitchell, T. V., & Maslin, M. T. (2007). How vision matters for individuals with hearing loss. *International Journal of Audiology*, 46(9), 500–511.
- Molholm, S., Ritter, W., Murray, M. M., Javitt, D. C., Schroeder, C. E., & Foxe, J. J. (2002). Multisensory auditory–visual interactions during early sensory processing in humans: A high-density electrical mapping study. *Cognitive Brain Research*, 14(1), 115–128.
- Müller-Deile, J., Kiefer, J., Wyss, J., Nicolai, J., & Battmer, R. (2008). Performance benefits for adults using a cochlear implant with adaptive dynamic range optimization (ADRO): A comparative study. *Cochlear Implants International*, 9(1), 8–26.
- Müller-Deile, J., Schmidt, B., & Rudert, H. (1995). Effects of noise on speech discrimination in cochlear implant patients. *The Annals of Otology, Rhinology & Laryngology. Supplement*, 166, 303–306.

- Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology*, 24(4), 375–425.
- Nieuwenhuys, R. (2012). The insular cortex: a review. *Progress in brain research*, 195, 123–163.
- Noesselt, T., Rieger, J. W., Schoenfeld, M. A., Kanowski, M., Hinrichs, H., Heinze, H. J., & Driver, J. (2007). Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *Journal of Neuroscience*, 27(42), 11431–11441.
- Obleser, J., & Kotz, S. A. (2011). Multiple brain signatures of integration in the comprehension of degraded speech. *Neuroimage*, 55(2), 713–723.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The edinburgh inventory. *Neuropsychologia*, 9(1), 97–113.
- Pantev, C., Eulitz, C., Hampson, S., Ross, B., & Roberts, L. (1996). The auditory evoked “off” response: Sources and comparison with the “on” and the “sustained” responses. *Ear and Hearing*, 17(3), 255–265.
- Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology*, 72(2), 184–187.
- Pollack I. (1975). Auditory informational masking. *J. Acoust. Soc. Am.* 57, S5 10.1121/1.1995329
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Radecke, J. O., Schierholz, I., Kral, A., Lenarz, T., Murray, M. M., & Sandmann, P. (2022). Distinct multisensory perceptual processes guide enhanced auditory recognition memory in older cochlear implant users. *NeuroImage: Clinical*, 33, 102942.
- Romei, L., Wambacq, I. J., Besing, J., Koehnke, J., & Jerger, J. (2011). Neural indices of spoken word processing in background multi-talker babble. *International Journal of Audiology*, 50(5), 321–333.
- Rosemann, S., & Thiel, C. M. (2018). Audio-visual speech processing in age-related hearing loss: Stronger integration and increased frontal lobe recruitment. *Neuroimage*, 175, 425–437.
- Ross, B., & Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research*, 248(1-2), 48–59.
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., & Foxe, J. J. (2007). Do you see what i am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral Cortex*, 17(5), 1147–1153.

- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., & Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences*, 104(17), 7295–7300.
- Rudner, M., Seeto, M., Keidser, G., Johnson, B., & Rönnberg, J. (2019). Poorer speech reception threshold in noise is associated with lower brain volume in auditory and cognitive processing regions. *Journal of Speech, Language, and Hearing Research*, 62(4S), 1117–1130.
- Sandmann, P., Eichele, T., Buechler, M., Debener, S., Jäncke, L., Dillier, N., Hugdahl, K., & Meyer, M. (2009). Evaluation of evoked potentials to dyadic tones after cochlear implantation. *Brain*, 132(7), 1967–1979.
- Sandmann, P., Kegel, A., Eichele, T., Dillier, N., Lai, W., Bendixen, A., Debener, S., Jancke, L., & Meyer, M. (2010). Neurophysiological evidence of impaired musical sound perception in cochlear-implant users. *Clinical Neurophysiology*, 121(12), 2070–2082.
- Sandmann, P., Plotz, K., Hauthal, N., Vos, M. de, Schönfeld, R., & Debener, S. (2015). Rapid bilateral improvement in auditory cortex activity in postlingually deafened adults following cochlear implantation. *Clinical Neurophysiology*, 126(3), 594–607.
- Schierholz, I., Finke, M., Kral, A., Büchner, A., Rach, S., Lenarz, T., Dengler, R., & Sandmann, P. (2017). Auditory and audio-visual processing in patients with cochlear, auditory brainstem, and auditory midbrain implants: An EEG study. *Human Brain Mapping*, 38(4), 2206–2225.
- Schierholz, I., Finke, M., Schulte, S., Hauthal, N., Kantzke, C., Rach, S., Büchner, A., Dengler, R., & Sandmann, P. (2015). Enhanced audio-visual interactions in the auditory cortex of elderly cochlear-implant users. *Hearing Research*, 328, 133–147.
- Senkowski, D., Saint-Amour, D., Gruber, T., & Foxe, J. J. (2008). Look who's talking: the deployment of visuo-spatial attention during multisensory speech processing under noisy environmental conditions. *Neuroimage*, 43(2), 379–387.
- Simmons, F. B., Epley, J. M., Lummis, R. C., Guttman, N., Frishkopf, L. S., Harmon, L. D., & Zwicker, E. (1965). Auditory nerve: electrical stimulation in man. *Science*, 148(3666), 104–106.
- Sörqvist, P., Ljungberg, J. K., & Ljung, R. (2010). A sub-process view of working memory capacity: Evidence from effects of speech on prose memory. *Memory*, 18(3), 310–326.
- Stein, B., & Meredith, M. (1993). *The merging of the senses*. Cambridge, MA: A Bradford book. MIT Press.
- Stevenson, R. A., Bushmakin, M., Kim, S., Wallace, M. T., Puce, A., & James, T. W. (2012). Inverse effectiveness and multisensory interactions in visual event-related potentials with audiovisual speech. *Brain Topography*, 25(3), 308–326.
- Stevenson, R., Sheffield, S. W., Butera, I. M., Gifford, R. H., & Wallace, M. (2017). Multisensory integration in cochlear implant recipients. *Ear and Hearing*, 38(5), 521.

- Sterzer, P., & Kleinschmidt, A. (2010). Anterior insula activations in perceptual paradigms: often observed but barely understood. *Brain Structure and Function*, 214(5), 611-622.
- Strelnikov, K., Rouger, J., Demonet, J. F., Lagleyre, S., Fraysse, B., Deguine, O., & Barone, P. (2013). Visual activity predicts auditory recovery from deafness after adult cochlear implantation. *Brain*, 136(12), 3682-3695.
- Stropahl, M., Bauer, A.-K. R., Debener, S., & Bleichner, M. G. (2018). Source-modeling auditory processes of EEG data using EEGLAB and brainstorm. *Frontiers in Neuroscience*, 12, 309.
- Stropahl, M., & Debener, S. (2017). Auditory cross-modal reorganization in cochlear implant users indicates audio-visual integration. *NeuroImage: Clinical*, 16, 514-523.
- Stropahl, M., Plotz, K., Schönfeld, R., Lenarz, T., Sandmann, P., Yovel, G., De Vos, M., & Debener, S. (2015). Cross-modal reorganization in cochlear implant users: Auditory cortex contributes to visual face processing. *Neuroimage*, 121, 159-170.
- Stropahl, M., Schellhardt, S., & Debener, S. (2017). McGurk stimuli for the investigation of multisensory integration in cochlear implant users: The Oldenburg audio visual speech stimuli (OLAVS). *Psychonomic Bulletin & Review*, 24(3), 863-872.
- Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The Journal of the Acoustical Society of America*, 26(2), 212-215.
- Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., & Leahy, R. M. (2011). Brainstorm: A user-friendly application for MEG/EEG analysis. *Computational Intelligence and Neuroscience*, 2011.
- Tremblay, K., Ross, B., Inoue, K., McClannahan, K., & Collet, G. (2014). Is the auditory evoked P2 response a biomarker of learning? *Frontiers in Systems Neuroscience*, 8, 28.
- Tye-Murray, N., Sommers, M. S., & Spehar, B. (2007). Audiovisual integration and lipreading abilities of older adults with normal and impaired hearing. *Ear and Hearing*, 28(5), 656-668.
- Uddin, L. Q., Nomi, J. S., Hébert-Seropian, B., Ghaziri, J., & Boucher, O. (2017). Structure and function of the human insula. *Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society*, 34(4), 300.
- van de Rijt, L. P., Roye, A., Mylanus, E. A., Van Opstal, A. J., & Van Wanrooij, M. M. (2019). The principle of inverse effectiveness in audiovisual speech perception. *Frontiers in human neuroscience*, 13, 335.
- Vaughan Jr, H. G., & Ritter, W. (1970). The sources of auditory evoked responses recorded from the human scalp. *Electroencephalography and Clinical Neurophysiology*, 28(4), 360-367.
- Vohn, R., Fimm, B., Weber, J., Schnitker, R., Thron, A., Spijkers, W., ... & Sturm, W. (2007). Management of attentional resources in within-modal and cross-modal divided attention tasks: An fMRI study. *Human brain mapping*, 28(12), 1267-1275.

- Wagener, K., Brand, T., & Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache. I-III: Design, Optimierung und Evaluation des Oldenburger Satztests (development and evaluation of a sentence test for the German language. I-III: Design, optimization and evaluation of the Oldenburg sentence test). *Zeitschrift für Audiologie (Audiological Acoustics)*, 38, 4–15.
- Warrier, C., Wong, P., Penhune, V., Zatorre, R., Parrish, T., Abrams, D., & Kraus, N. (2009). Relating structure to function: Heschl's gyrus and acoustic processing. *Journal of Neuroscience*, 29(1), 61–69.
- Weglage, A., Müller, V., Layer, N., Abdel-Latif, K. H., Lang-Roth, R., Walger, M., & Sandmann, P. (2022). Side-of-implantation effect on functional asymmetry in the auditory cortex of single-sided deaf cochlear-implant users. *Brain Topography*, 1–22.
- Weiner, K. S., & Grill-Spector, K. (2013). Neural representations of faces and limbs neighbor in human high-level visual cortex: evidence for a new organization principle. *Psychological research*, 77(1), 74–97.
- Wesemann, W., Schiefer, U., & Bach, M. (2010). Neue DIN-Normen zur Sehschärfebestimmung. *Der Ophthalmologe*, 107(9), 821–826.
- Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data—a practical approach. *Journal of Neuroscience Methods*, 250, 34–46.
- Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research*, 242(1-2), 3–21.
- Winn, M. B. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing*, 20, 2331216516669723.
- Wolfe, J., Neumann, S., Marsh, M., Schafer, E., Lianos, L., Gilden, J., O'Neill, L., Arkis, P., Menapace, C., Nel, E., others. (2015). Benefits of adaptive signal processing in a commercially available cochlear implant sound processor. *Otology & Neurotology*, 36(7), 1181–1190.
- Wolpaw, J. R., & Penry, J. K. (1975). A temporal component of the auditory evoked response. *Electroencephalography and Clinical Neurophysiology*, 39(6), 609–620.
- Wong, P. C., Ettlinger, M., Sheppard, J. P., Gunasekera, G. M., & Dhar, S. (2010). Neuroanatomical characteristics and speech perception in noise in older adults. *Ear and Hearing*, 31(4), 471.
- Wong, P. C., Jin, J. X., Gunasekera, G. M., Abel, R., Lee, E. R., & Dhar, S. (2009). Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia*, 47(3), 693–703.
- Zhao, F., Stephens, S. D. G., Sim, S. W., & Meredith, R. (1997). The use of qualitative questionnaires in patients having and being considered for cochlear implants. *Clinical Otolaryngology & Allied Sciences*, 22(3), 254–259.

Zeng, F.-G. (2011). Advances in auditory prostheses. In *Auditory prostheses* (pp. 1–11). Springer.

DISCUSSION

6.1 SUMMARY

The projects presented in this thesis cover a broad variety of research questions regarding auditory and especially audiovisual speech processing in CI users from different perspectives. They include new insights for auditory processing in SSD CI users, explore the timecourse of auditory and audiovisual speech processing in CI users with bilateral hearing loss and compare it to a group of NH controls. A further novelty is a systematic comparison of this timecourse between CI users with bilateral hearing loss, CI users with unilateral hearing loss and NH controls. Furthermore, we present results on audiovisual processing and lip-reading abilities in CI users with unilateral hearing loss for the first time. Finally, we investigate the issue of background noise and its influence on audiovisual interactions in CI users with bilateral hearing loss. The following sections summarise the most important results and discusses them in the context of Health Sciences.

6.2 ALTERED AUDITORY PROCESSING IN CI USERS WITH UNILATERAL HEARING LOSS

The project "Side of implantation effect on functional asymmetry in the auditory cortex of single-sided deaf cochlear-implant users" (Chapter 3) aims to further investigate the particularities of auditory processing in SSD CI users. In contrast to research investigating auditory processing in CI users with bilateral hearing loss, research focusing on CI users with unilateral hearing loss is rather scarce. This is because SSD CI users have an intact NH ear, which is why until recently they typically were not implanted with a CI (Zeitler and Dorman, 2019; Arndt et al., 2011; Buechner et al., 2010). For binaural hearing restoration various behavioural advantages have been reported in recent years (Kitterick et al., 2015; Arndt et al., 2017). By contrast, the number of studies on auditory processing in SSD CI users remains relatively low. Some ERP studies demonstrated that the processing with the CI ear is delayed compared to the NH ear (e.g. Finke et al., 2016b). It is unknown, however, whether SSD CI users' NH ears process auditory stimuli in the same way that NH listeners do. Furthermore, it is unclear whether the side of implantation plays a role in achieving implantation benefits in these patients.

The results of our ERP study replicated previous findings (Finke et al., 2016b; Bönitz et al., 2018; Legris et al., 2020), demonstrating prolonged response times, increased

subjective listening effort, and delayed and reduced auditory ERPs for the CI ear compared to the NH ear. This further confirms that auditory processing is more difficult via the CI ear than via the NH ear. However, comparing auditory processing in NH controls by contrasting "normal" and "vocoded" (CI-simulated) stimulation did not show delayed responses. Given that vocoded CI simulations can well approximate the CI experience (Friesen et al., 2001), one can assume that stimulus degradation alone cannot be the sole cause for the delayed responses. Previous longitudinal studies on intra-modal plasticity within the first months after implantation have shown that the N1 amplitude is increasing over time in CI users with bilateral hearing loss and in SSD CI users, but the amplitude stays below the level of NH controls (Sandmann et al., 2015; Legris et al., 2020). Thus, the differences between NH and CI ear processing appear most likely as a result of cortical reorganisation or a combination of cortical reorganisation and the auditory stimulus degradation via the CI.

Furthermore, the oddball paradigm that we used to present our syllables made it possible to objectively distinguish between proficient and non-proficient SSD CI users, as proficient SSD CI users were capable of distinguishing both the easy syllable contrast (/ki/ vs. /ka/) and the more difficult syllable contrast (/ki/ vs. /ti/). Whether a CI user was proficient was determined based on both behavioural (higher hit rates) and ERP measures (faster and/or enhanced N1 and P3b ERPs) compared to non-proficient CI users. This is consistent with previous studies, which found that when acoustic cues are easily distinguishable, ERP responses to deviant sounds are similar between CI users with bilateral hearing loss and NH listeners (Henkin et al., 2009; Turgeon et al., 2014).

Concerning the questions on side-of-implantation, left-implanted CI users showed an enhanced functional auditory-cortex asymmetry compared to right-implanted CI users, reflected in greater right hemispheric activation compared to the left hemisphere. This seems to contradict previous results on CI users bilateral hearing loss who demonstrated a stronger auditory-cortex asymmetry for right-stimulated CI users (Sandmann et al., 2009). The previous study, however, used tones, whereas we used syllables which can potentially cause these contradicting effects, given that the two cortical hemispheres have been shown to be differently engaged for these inherently different types of stimuli (Poeppel, 2003; Belin et al., 2000; Zatorre and Belin, 2001). Furthermore, because auditory deprivation reduces metabolism in the contralateral auditory cortex (Speck et al., 2020) and can induce cortical reorganisation in the auditory cortex (Stropahl et al., 2017a), different activity patterns may have emerged as a result of the confounding effect of hearing loss in the second ear, making CI users with bilateral and unilateral hearing loss not very comparable. However, because similar effects were observable in left-ear stimulated NH controls when they were stimulated with vocoded stimuli, we suggest that the side-of-implantation-effect originates from the limited quality of the CI signal. Comparing the processing between the NH ears of SSD CI users and NH controls revealed, a functional auditory-cortex asymmetry in left-implanted SSD CI users, whereas no difference was detected in NH controls. As far as we are aware, this study is the first one demonstrating a difference in auditory processing between the NH ear of SSD CI users and NH controls. This observations leads to the assumption that cortical reorganisation occurs as a result of unilateral deafness, which appears to extend to the NH ear, indicating that auditory

processing via the NH ear of SSD CI users is not the same as in NH controls. However, because the sample size was relatively low and the participants were stimulated with syllables only, more studies with different stimuli (words, musical sounds) are needed to verify the robustness of the observed side-of-implantation effects.

Taken together, the findings of this project contribute to a better understanding of rehabilitation success in SSD CI users by showing that auditory input degradation caused by the CI and experience-related functional changes within the auditory cortex influence cortical speech processing in SSD CI users.

Studying the behaviour and cortical changes in SSD CI users is very interesting because these patients provide a unique perspective on what hearing with a cochlear implant sounds like (Zeitler and Dorman, 2019). These patients can compare hearing sensations of their NH ear with their CI ear, which makes it difficult for these CI users to adapt to the new sensation because they constantly have the direct comparison to natural hearing. This leads to the experience of a maximally asymmetric hearing sensation that must be learned to be integrated (Gordon et al., 2013; Kral et al., 2013), which often results in longer rehabilitation periods compared to CI users with bilateral hearing loss (Muigg et al., 2020). Because of the extreme differences between the ears, it was assumed that cochlear implantation would have no benefits because the central auditory system would be unable to successfully process the auditory discrepancy between acoustic and electric auditory input (Zeitler and Dorman, 2019). However, a study performed in 2008 was the first one to show remarkable benefits of implantation of SSD CI users in the reduction of tinnitus (van de Heyning et al., 2008), with various other studies to follow. These subsequent studies found significant behavioural improvements in tinnitus reduction, speech-in-noise intelligibility, directional hearing, and quality of life due to the restoration of binaural hearing (Kitterick et al., 2015; Arndt et al., 2017; Dorbeau et al., 2018; Galvin et al., 2018; Grossmann et al., 2016; Távara-Vieira et al., 2015). In recent years, there has been a growing body of reports indicating that implantation in SSD children is beneficial and highly recommended (for review see Benchetrit et al., 2021). Recent studies have even highlighted the consequences for academic performance and social interactions if binaural hearing restoration is not addressed on time (Kay-Rivest et al., 2022; Macías et al., 2019).

Thus, there exists clear behavioural evidence that the central auditory system is indeed capable of matching the highly differing input, even though processing seems to be delayed for the CI ear compared to the NH ear (Finke et al., 2016b; Bönitz et al., 2018), which is also replicated in the presented study. Studying the processing mechanisms underlying these behavioural effects can help to identify additional beneficial effects and their objective correlates. However, more research is needed to investigate the auditory processing differences, like evaluating the robustness of the side-of-implantation effect observed in this study, as well as the implications for auditory processing with the NH ear in these patients. Such studies would contribute to a better understanding of cortical reorganisation in SSD CI users compared to CI users with bilateral hearing loss. Longitudinal studies, in particular, have the potential to provide valuable insights into the changes in cortical processing before and after implantation in order to study the timecourse of rehabilitation. Insights from such studies may aid in predicting CI outcomes and determining, for example, the neural basis for proficient and non-proficient SSD CI users. Finally, such studies

may encourage other countries to relax their implantation criteria, which still vary greatly across the globe (Van de Heyning et al., 2022), in order to allow for the most early intervention possible.

6.3 ALTERED AUDIOVISUAL INTERACTIONS IN CI USERS WITH BILATERAL HEARING LOSS

The study "The timecourse of multisensory speech processing in unilaterally stimulated cochlear implant users revealed by ERPs" (Chapter 2) investigates, as the title already indicates, the timecourse of audiovisual speech processing in CI users with bilateral hearing loss. Previous research has demonstrated that CI users with bilateral hearing loss exhibit both intra-modal and cross-modal changes before and after receiving the CI (for review see Stropahl et al., 2017a). However, cross-modal recruitment of the visual cortex for auditory stimuli has only been shown in PET and fNIRS studies (Giraud et al., 2001c; Chen et al., 2016) which could not provide information about the precise temporal dynamics of this phenomenon, due to their limited temporal resolution. Moreover, numerous studies have demonstrated that the visual aspect of speech is particularly important for CI users, as they develop pronounced lip-reading skills (Desai et al., 2008; Butera et al., 2018; Giraud et al., 2001c; Strelnikov et al., 2015; Rouger et al., 2007; Stropahl et al., 2015). Most ERP studies, however, reported results on auditory-only processing. Only recently, some ERP studies were rather focused to audiovisual processing (Schierholz et al., 2015, 2017; Radecke et al., 2022), as audiovisual stimulation better represents the CI users' real-world listening situations (for review see Stevenson et al., 2017).

In our ERP study, we focused on the question whether CI users have an altered timecourse of audiovisual syllable processing as well as distinct audiovisual interactions compared to a group of matched NH controls. By administering methods from electrical neuroimaging which provide a spatio-temporal perspective (Michel et al., 2009), we were able to confirm differences in speech processing between CI users and NH listeners.

However, there were also some similarities between the groups: both groups showed a multisensory effect which was behaviourally reflected in shorter audiovisual response times compared to the ones of unimodal syllables. This is in line with the results of the topographic and source analyses, revealing distinct P2 topographies and generally reduced amplitudes for visually modulated auditory responses (AV-V) compared to auditory responses (A). Thus, the additional visual input caused a modulation of auditory responses, which has previously been shown for CI users (Schierholz et al., 2015) and NH listeners (Cappe et al., 2010). These condition-specific results show that audiovisual information is both beneficial for CI users and NH listeners. Given that CI users are "better audiovisual integrators" (Rouger et al., 2007), one might be surprised to not find a difference between CI users and NH listeners for audiovisual and visual syllable discrimination. Nevertheless, this is in accordance with an ERP study by Schierholz et al. (2015) using simple, non-linguistic stimuli that did not reveal a difference between those two groups either. Given that CI users were better at identifying visually and audiovisually presented words compared to NH controls (Rouger et al., 2007), syllables are apparently not complex enough, which is

in accordance with the generally high hit rates in this study. Further studies using more complex linguistic stimuli are required to verify behavioural differences (hit rates and response times) between CI users and NH listeners.

However, CI users showed a distinct topographic pattern at N1 latency, with a clearer N1 topography for AV-V compared to A, which indicates a particularly strong visual influence in auditory processing in CI users. In contrast to CI users, NH listeners showed the same presence of a clear N1 topography both for AV-V and A, which leads to the assumption that similar neuronal processes are engaged to evaluate these stimuli in NH listeners. This difference was accompanied by a distinct source configuration in CI users within the auditory and visual cortices at N1 latency. CI users exhibited a delayed activation in the auditory cortex in contrast to NH listeners, which was also shown in a previous ERP study with audiovisual stimuli (Schierholz et al., 2015). Given that a follow-up study (Schierholz et al., 2017) found delayed N1 responses not only in CI users, but also in ABI and AMI patients, which are auditory implants that bypass even more structures of the auditory pathway (see chapter 1, 1.1 hearing disorders, paragraph "types of hearing loss"). These findings suggest that the delayed auditory-cortex responses can be attributed to the auditory implants' limited input quality (Schierholz et al., 2017). Processing delays, in particular, may reflect difficulties processing auditory information, which then require more time to be encoded. In line with this, previous ERP studies with CI users have repeatedly presented delayed N1 responses to stimuli of varying linguistic complexity (Finke et al., 2016a; Beynon et al., 2005; Sandmann et al., 2015; Senkowski et al., 2014).

In addition, CI users showed an enhanced recruitment of the visual cortex for auditory and audiovisual stimuli compared to NH listeners, that was positively correlated with the CI-experience. This complements previous PET results that demonstrated enhanced visual recruitment for auditory speech stimuli in CI users compared to NH controls, which was also correlated with enhanced lip-reading skills and longer CI-experience (Giraud et al., 2001c). Giraud et al. (2001c) suggested that the enhanced visual-cortex recruitment represents a particularly strong synergy between the visual and auditory systems. This synergy may aid in the recovery of speech intelligibility after implantation, when CI users must learn to match novel auditory sensations to corresponding visual speech cues (Strelnikov et al., 2013; Rouger et al., 2012). The presence of a positive relationship between the extent of cross-modal visual-cortex recruitment and CI-experience suggests experience-related reorganisation after receiving the implant. Our study is the first EEG study that confirms these early observations and extends them by investigating the additional temporal aspect. These experience-related cortical changes seem to appear in the initial stages of auditory and audiovisual processing, already at N1 latency. Finally, in accordance with previous studies (Rouger et al., 2007; Stropahl et al., 2015), we confirmed enhanced lip-reading abilities in CI users compared to NH listeners, which may be due to sharpened visual abilities which emerge either during the phase of deafness before implantation or after restoring hearing abilities with a CI. Lip-reading abilities that exist prior to implantation do not reverse after implantation, but rather remain stable and may even improve due to the audiovisual integration which is required to match the limited CI input with visual speech information (Lazard et al., 2012a; Rouger et al., 2012).

In summary, our study confirms that CI users show a distinct timecourse of auditory and audiovisual processing and distinct audiovisual interactions. These differences are most evident at N1 latency, which exhibited a distinct topographic pattern and CI-specific alterations within the auditory and visual cortices compared to NH controls. The enhanced visual cortex-recruitment may be a compensatory strategy to overcome the limited CI input, which causes delays in auditory-cortex responses. This compensatory strategy appears to cause the development of behavioural visual refinements such as lip-reading abilities, which aid in audiovisual speech integration.

We suggest that future ERP studies should investigate the timecourse of audiovisual speech processing also using more complex linguistic stimuli, such as words or sentences. Moreover, given previous results, as well as the findings presented in this project, we suggest that clinical assessments of speech comprehension in CI users need to be adjusted. Currently, these are restricted to auditory-only methods, which is insufficient for gauging the actual speech intelligibility abilities of CI patients, as the important visual aspect of speech is neglected (Stevenson et al., 2017). The presented study confirms that CI users benefit greatly from additional visual input and appear to use different processing strategies than NH listeners. Although the auditory CI input is limited in acoustic quality, the CI users' multisensory processes seem to remain intact.

The observation of strong audiovisual interactions is clinically relevant, as they can aid in the optimisation of auditory rehabilitation after implantation. As previously stated, CI users are required to match novel auditory sensations to corresponding visual information after implantation (Strelnikov et al., 2009). This strongly suggests that audiovisual training may be beneficial for accelerating and optimising the restoration of speech recognition abilities with a CI. According to Strelnikov et al. (2009), audiovisual rehabilitation may even extend to other benefits, such as sound localisation. Previous research has shown that there must be an adequate amount of auditory-cortex activity after CI implantation in response to auditory stimuli (Green et al., 2005; Sandmann et al., 2015; Mortensen et al., 2006), as the strength of auditory-cortex activation and speech recognition performance are positively correlated (Green et al., 2005). Thus, for obtaining successful speech recognition abilities the auditory cortex has to be effectively stimulated by the CI. This is consistent with findings that show that good CI performers show an enhanced recruitment of the auditory and inferior prefrontal cortices in response to speech stimuli compared to poor CI performers (Mortensen et al., 2006).

To achieve an optimal stimulation of the auditory cortex, an effective rehabilitation is essential, and audiovisual training may have the potential to realise faster improvements. Indeed, some studies with NH listeners provided evidence for beneficial effects of audiovisual training. Seitz et al. (2006) discovered that when a visual motion detection task was previously trained in an audiovisual setting, the participants showed significantly better results than when the same task was trained only visually. Similar results have been presented in the context of training with congruent and incongruent object recognition. Audiovisual training resulted in an improved recognition of the same objects in subsequent auditory-only or visual-only settings, but only for congruent pairs (Matusz et al., 2015; Thelen and Murray, 2013).

A further training study by Kawase et al. (2009) investigated NH listeners using vocoded stimuli simulating the signal provided with a CI. They used a word recognition task including auditory or audiovisual training and either provided feedback on the performance during the training or not. The results demonstrated that audiovisual training resulted in better subsequent auditory word recognition than auditory-only training. In addition, providing feedback during training generally resulted in better performance (Kawase et al., 2009). Another study that used vocoded stimuli with NH listeners reported similar results for nonsense words and pictures, where audiovisual training resulted in better performance compared to auditory-only training (Bernstein et al., 2013).

A follow up study of Bernstein et al. (2013) used the same task and compared the training results for NH listeners with CI users (Bernstein et al., 2014). However, CI users did not show the same benefits of audiovisual training as NH controls. But, the CI users in this study had prelingual hearing loss or deafness at birth, causing them to rely heavily on vision, which may explain the unexpected results (Bernstein et al., 2014). By contrast, a recent study with *postlingual* CI users used the same stimuli as in Matusz et al. (2015) and assessed auditory object recognition with previously presented auditory and audiovisual stimuli (Radecke et al., 2022). The task was a two-alternative forced choice task, for which the participants were asked to indicate whether a current stimulus was already presented before. For both CI users and NH controls, the authors confirmed a multisensory benefit for later auditory object recognition (Radecke et al., 2022), and even a greater behavioural audiovisual benefit for CI users compared to NH controls (see supplementary material of Radecke et al., 2022).

Taken together, insights from NH listeners with both normal and vocoded stimulation and initial insights from CI users indicate that audiovisual training may have beneficial effects for subsequent auditory performance with the CI. We suggest that future studies should confirm the efficiency of audiovisual training in CI users by contrasting patients with experienced and inexperienced CI users, or pre-and postlingual CI users. Insights from such studies may help determining whether audiovisual training is best suited for specific cases or for all CI users and whether it can be used for achieving a faster and improved auditory rehabilitation after CI implantation.

6.4 ALTERED AUDIOVISUAL INTERACTIONS IN DIFFERENT CI USER GROUPS

The project "Electrophysiological differences and similarities in audiovisual speech processing in cochlear implant users with unilateral and bilateral hearing loss" (Chapter 4) is a follow-up study of Layer et al. (2022), which is also included in this thesis. As previously mentioned in this chapter, ERP studies on auditory processing in SSD CI users are scarce, with even fewer studies comparing SSD CI users with CI users with bilateral hearing loss. As far as we are aware, our study is the first one that systematically compares auditory and especially audiovisual processing between these two CI user groups and a control group of NH listeners. Currently, there are no audiovisual studies with SSD CI users and there are no reports on lip-reading abilities in this CI user group. Therefore, our study intended to fill these gaps.

As in Layer et al. (2022), we observed a multisensory effect for all experimental groups, which was reflected in the behavioural and ERP data. The response times were significantly reduced for audiovisual syllables compared to the two unimodal syllables. The ERPs were generally reduced (both at N1 and P2 latency) for modulated responses (AV-V) compared to auditory responses (A). As a result, audiovisual information benefited all three groups. Concerning the group differences, topographic analyses revealed that CI users with bilateral hearing loss showed a significant change in N1 voltage topographies for AV-V compared to A. This suggests that additional visual information leads to a more frequent distribution of a clear N1 topography in CI users with bilateral hearing loss, which was neither observed in SSD CI users nor in the NH controls. A condition-specific modulation of the topography points to a particularly strong visual impact on auditory processing for CI users with bilateral hearing loss compared to the other two groups. However, both CI user groups displayed delayed responses in the auditory cortex, indicating difficulties in processing the limited CI input, as previously demonstrated for each CI group separately (e.g. Finke et al., 2016a,b). To probably compensate for the limited CI input (Doucet et al., 2006; Giraud et al., 2001c; Strelnikov et al., 2013), both CI user groups showed an enhanced cross-modal recruitment of the visual cortex and better results in the lip-reading test compared to NH listeners. SSD CI users appear to develop similar strategies and processing patterns to compensate for the limited CI input as CI users with bilateral hearing loss, even though they have an intact contralateral NH ear. Therefore, we conclude that these are universal to all CI user groups, as the observed visual enhancements (enhanced lip-reading skills and visual-cortex activity) appear to be independent of the contralateral ear's hearing abilities.

We propose that future studies should further investigate SSD CI users in the context of audiovisual processing by including more complex linguistic stimuli, like words and sentences. Furthermore, more systematic comparisons of different CI user groups are required to verify that all CI user groups have universal processing strategies to compensate for the degraded CI input. Given the difference in topographic pattern at N1 latency, which was only observed in CI users with bilateral hearing loss, we hypothesise that more differences may emerge for more complex linguistic stimuli. Furthermore, future studies may consider further dividing CI user groups with bilateral hearing loss into bimodally (CI + hearing aid) and bilaterally (CI + CI) supplied CI users and to compare the processing of the CI ear, the contralateral ear and both ears together. Finally, differences between CI user groups may not be limited to the auditory and visual cortices, which were the focus of our research. Therefore, future studies should consider further cortical regions, such as frontal or temporal areas (Campbell and Sharma, 2013; Giraud et al., 2001a), as they may be the source of potential topographic dissimilarities.

Understanding CI user group-specific processing mechanisms may aid in the development of tailored therapeutic solutions to ensure a rapid and successful CI rehabilitation. Our findings, as well as the previous section's discussion of existing audiovisual training studies, indicate that audiovisual training has the potential to benefit all CI user groups. A study by Strelnikov et al. (2011) investigated the effect of audiovisual training on monaural spatial hearing in NH listeners. The results confirmed superior performance when the participants received audiovisual compared to

visual training. Based on these findings, the authors suggest that audiovisual training in the context of spatial hearing may be used to train SSD CI users with their CI after implantation. However, there exist no audiovisual training studies with SSD CI users or audiovisual training studies comparing the effectiveness for different CI user groups. Thus, future research should include audiovisual training in order to find evidence for potentially beneficial effects for an improved and faster rehabilitation with a CI.

6.5 BACKGROUND NOISE INFLUENCES AUDIOVISUAL INTERACTIONS IN CI USERS

The final project of this thesis "The influence of background noise and noise reduction by 'ForwardFocus' on audiovisual speech perception in bimodal CI users: an ERP study" (Chapter 5) investigates the impact of background noise on auditory and audiovisual speech processing in CI users. This is an important issue because background noise makes the already limited CI input even more difficult to understand (Müller-Deile et al., 1995; Wilson and Dorman, 2008). Especially in situations with background noise, the visual aspect of speech (facial movements, gestures) has been shown to be extremely helpful (Ross et al., 2007; Sumby and Pollack, 1954). To relieve CI users in demanding hearing situations, CI manufacturers constantly develop specific algorithms which can be activated (or get activated automatically in appropriate situations), such as noise-reduction algorithms (Mauger et al., 2014; Müller-Deile et al., 2008). A recently introduced noise-reduction algorithm is "ForwardFocus", which has been shown to be effective in behavioural studies (Hey et al., 2019, 2021). Our study aims to investigate the impact of background noise on auditory and audiovisual syllable processing, and to evaluate "ForwardFocus" by means of objective ERPs.

The results reveal that the presentation of audiovisual syllables generally results in shorter response times and enhanced P2 amplitudes in various cortical regions (superior temporal cortex, lateral occipital cortex, insular cortex) compared to auditory-only syllables. The result of a multisensory benefit is consistent with previous reports with CI users (Schierholz et al., 2015; Radecke et al., 2022), which are extended to conditions with noise in this study. These results indicate that both conditions — with and without background noise — cause audiovisual interactions in CI users in expanded cortical areas. However, our results confirm that background noise generally hampers speech processing, which is reflected in prolonged response times, enhanced listening effort, and perceived task difficulty, delayed auditory-cortex activity within the time window of the N1, and reduced activity within the superior frontal cortex at P2 latency (both for auditory and audiovisual conditions). Activating "ForwardFocus", however, is beneficial as it positively influences the aforementioned behavioural measures. Concerning the ERPs, "ForwardFocus" causes enhanced P2 responses in the superior frontal cortex, especially in audiovisual conditions. The results provide further behavioural and electrophysiological evidence showing clear effects of background noise, both on auditory and audiovisual processing in CI users. The P2 has been shown to be modulated by attention (Crowley and Colrain, 2004), and the superior frontal cortex has been described in the context of audiovisual divided attention (Vohn et al., 2007) and visuo-spatial attention (Marshall et al., 2015; Kastner et al., 1999; Corbetta and Shulman, 2002). Given that "ForwardFocus" seems

to work best in audiovisual conditions, it is reasonable to hypothesise that it helps in the reallocation of attentional resources, which results in a stronger perceived spatial congruency of auditory and visual speech cues (Radecke et al., 2022; Butera et al., 2018; García-Larrea et al., 1992).

Overall, multisensory information cannot prevent the delayed auditory-cortex activity which happens independently of whether syllables were presented in auditory-only conditions or audiovisually. However, other cortical regions seem to be differently involved based on whether the information was presented in unisensory or in multisensory conditions and with or without background noise. More studies are required to further disentangle these effects. It would be interesting to compare the presented effects of noise to NH individuals, to see whether noise has distinct effects on auditory and audiovisual processing, and whether some cortical regions are more strongly involved in CI users than in NH listeners, such as the visual cortex in our previous projects.

A limitation of this study is that the background noise was only presented from 90 and 270 degrees, and that we only included one type of noise. Future studies should test multiple noise directions with different types of noise, as these factors may highly influence the observed benefits of "ForwardFocus", given that it is primarily specified to reducing noise from behind. In addition, we would suggest comparing various signal-to-noise ratios within the EEG experiment, which would provide a more fine-grained picture of the benefits of using "ForwardFocus" and a clearer picture of the effects of background noise on audiovisual speech processing. Moreover, including more complex linguistic stimuli, such as words and sentences would probably result in even greater benefits in this experimental setting and would be even more ecologically valid. Nevertheless, by using simple syllables we could already demonstrate a clear effect of background noise and an effect of noise reduction via "ForwardFocus", which extends previous behavioural results (Hey et al., 2019, 2021) by the present objective ERP results. Based on the presented results, we recommend this algorithm to CI users in situations with interfering noise (coming from distinct sources than the speech signal). We suggest that future noise reduction algorithms should be evaluated in both auditory and audiovisual conditions.

The results of multisensory benefits further support that clinical hearing assessments should also include audiovisual testing procedures, as spoken communication is inherently multisensory. Furthermore, given that successful speech recognition in noise is a desirable milestone in rehabilitation which cannot always be guaranteed, the inclusion of training procedures with background noise is critical. Indeed, previous studies with CI patients practising digits in noise (Oba et al., 2011), sentences in noise (Green et al., 2019), and spatially separated speech in noise (Lotfi et al., 2021) have shown that training in noise is beneficial. Training with background noise can have positive implications for speech understanding in quiet, as it may trigger top-down processes that may be helpful in developing listening strategies to compensate for background noise (Oba et al., 2011). A study with NH listeners by Lidestam et al. (2014) showed positive effects of audiovisual training on speech-in-noise identification compared to purely auditory training. Similar audiovisual training procedures may be helpful in audiological settings, as CI users may highly benefit from such training, which was also suggested by the authors (Lidestam et al., 2014).

6.6 OUTLOOK

Further systematic comparisons between different CI user groups and stimuli may allow for the identification of the optimal type of audiovisual training, and by extension audiovisual training with background noise, given that real-world listening situations are multisensory and frequently masked with noise. Practising real-world listening situations may enable a better coping and preparation for the listening experiences in everyday life, to prevent too overwhelming sensations. Furthermore, one could consider developing training procedures that can be completed at home, as already suggested by Green et al. (2019). For example, if training material includes audiovisual stimulation with noise or multi-talker babble, one can train and sharpen hearing at home, as some patients may rather prefer practising at home or in small groups.

For obtaining effective test material, it is necessary to generate appropriate and convincing computer based animations of different speakers. Because the computer-based animation used in studies 1 and 3 ("MASSY"; Modular Audiovisual Speech SYnthesizer (Fagel and Clemens, 2004)) is relatively static, and lacks emotional and facial expressions (head and brow movements), many participants in our studies found this animation unnatural and artificial. Using video recordings of real people is also not a perfect solution, as these people have to be recorded many times. One approach could be to create a convincing video animation, which is based on movements from real persons which can be fed with arbitrary speech material and be parametrised. This would allow for the generation of an infinite number of test materials, optimally with the feature to select different sexes, age groups, and voices (to practise hearing at different frequencies and intensity levels). Such a video animation has recently been proposed by Meister et al. (2022), who are currently evaluating the suitability of this tool for using it in CI users. Such an optimal computer animation may also be used for clinical hearing test procedures. The existing auditory-only tests could be extended by audiovisual information and would give a more accurate evaluation of the speech comprehension of hearing impaired individuals.

6.7 CONCLUSION

Overall, this thesis provides further evidence that the visual part of audiovisual speech information is highly beneficial for all CI user groups, particularly in background noise situations, and that these benefits can be detected using both subjective, behavioural, and objective (electrophysiological) measures. The findings from the experiments presented in this thesis contribute to a better understanding of both auditory and audiovisual cortical processing in CI patients. They may aid in the development of predictive outcome models, audiovisual clinical hearing assessment procedures, as well as therapeutic rehabilitation techniques such as audiovisual training for improving the outcome with a CI.

BIBLIOGRAPHY

- Adrian, E. D. and Matthews, B. H. (1934). The Berger rhythm: potential changes from the occipital lobes in man. *Brain*, 57(4):355–385.
- Adunka, O. and Kiefer, J. (2005). Wie funktioniert der Sprachprozessor von Cochlea-Implantaten? *Laryngo-Rhino-Otologie*, 84(11):841–851.
- Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.-H., Sams, M., Shinn-Cunningham, B. G., et al. (2006). Task-modulated “what” and “where” pathways in human auditory cortex. *Proceedings of the National Academy of Sciences*, 103(39):14608–14613.
- Anderson, C. A., Wiggins, I. M., Kitterick, P. T., and Hartley, D. E. (2017). Adaptive benefit of cross-modal plasticity following cochlear implantation in deaf adults. *Proceedings of the National Academy of Sciences*, 114(38):10256–10261.
- Arndt, S., Aschendorff, A., Laszig, R., Beck, R., Schild, C., Kroeger, S., Ihorst, G., and Wesarg, T. (2011). Comparison of pseudobinaural hearing to real binaural hearing rehabilitation after cochlear implantation in patients with unilateral deafness and tinnitus. *Otology & neurotology*, 32(1):39–47.
- Arndt, S., Laszig, R., Aschendorff, A., Hassepass, F., Beck, R., and Wesarg, T. (2017). Cochleaimplantatversorgung bei einseitiger Taubheit oder asymmetrischem Hörverlust. *HNO*, 65(7):586–598.
- Arndt, S., Wesarg, T., Stelzig, Y., Jacob, R., Illg, A., Lesinski-Schiedat, A., Ketterer, M., Aschendorff, A., and Speck, I. (2019). Einfluss einseitiger Taubheit auf das Hörvermögen des besseren Ohrs. *HNO*, 67(10):739–749.
- Arnold, P. and Hill, F. (2001). Bisensory augmentation: A speechreading advantage when speech is clearly audible and intact. *British Journal of Psychology*, 92(2):339–355.
- Baillet, S. (2013). *Forward and Inverse Problems of MEG/EEG*, pages 1–8. Springer New York, New York, NY.
- Barth, D. S., Goldberg, N., Brett, B., and Di, S. (1995). The spatiotemporal organization of auditory, visual, and auditory-visual evoked potentials in rat cortex. *Brain research*, 678(1-2):177–190.
- Başar, E. (2022). Brain oscillations in neuropsychiatric disease. *Dialogues in clinical neuroscience*.
- Bavelier, D., Dye, M. W., and Hauser, P. C. (2006). Do deaf individuals see better? *Trends in cognitive sciences*, 10(11):512–518.
- Bawazeer, N., Vuong, H., Riehm, S., Veillon, F., and Charpiot, A. (2019). Magnetic resonance imaging after cochlear implants. *Journal of otology*, 14(1):22–25.

- Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature*, 403(6767):309–312.
- Benchetrit, L., Ronner, E. A., Anne, S., and Cohen, M. S. (2021). Cochlear implantation in children with single-sided deafness: a systematic review and meta-analysis. *JAMA Otolaryngology–Head & Neck Surgery*, 147(1):58–69.
- Berger, H. (1929). Über das Elektrenkephalogramm des Menschen. *Archiv f. Psychiatri*, 87:527–570.
- Bernstein, J. G., Goupell, M. J., Schuchman, G. I., Rivera, A. L., and Brungart, D. S. (2016). Having two ears facilitates the perceptual separation of concurrent talkers for bilateral and single-sided deaf cochlear implantees. *Ear and Hearing*, 37(3):289.
- Bernstein, L. E., Auer Jr, E. T., Eberhardt, S. P., and Jiang, J. (2013). Auditory perceptual learning for speech perception can be enhanced by audiovisual training. *Frontiers in neuroscience*, 7:34.
- Bernstein, L. E., Eberhardt, S. P., and Auer Jr, E. T. (2014). Audiovisual spoken word training can promote or impede auditory-only perceptual learning: Prelingually deafened adults with late-acquired cochlear implants versus normal hearing adults. *Frontiers in Psychology*, 5:934.
- Bernstein, L. E. and Liebenthal, E. (2014). Neural pathways for visual speech perception. *Frontiers in neuroscience*, 8:386.
- Beynon, A., Snik, A., Stegeman, D., and Van den Broek, P. (2005). Discrimination of speech sound contrasts determined with behavioral tests and event-related potentials in cochlear implant recipients. *Journal of the American Academy of Audiology*, 16(01):042–053.
- Biasiucci, A., Franceschiello, B., and Murray, M. M. (2019). Electroencephalography. *Current Biology*, 29(3):R80–R85.
- Billings, C. J., Tremblay, K. L., Stecker, G. C., and Tolin, W. M. (2009). Human evoked cortical activity to signal-to-noise ratio and absolute signal level. *Hearing research*, 254(1-2):15–24.
- Blackwood, D. and Muir, W. J. (1990). Cognitive brain potentials and their application. *The British Journal of Psychiatry*, 157(S9):96–101.
- Blamey, P., Artieres, F., Başkent, D., Bergeron, F., Beynon, A., Burke, E., Dillier, N., Dowell, R., Fraysse, B., Gallégo, S., et al. (2013). Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients. *Audiology and Neurotology*, 18(1):36–47.
- Blamey, P. J., Pyman, B. C., Clark, G. M., Dowell, R. C., Gordon, M., Brown, A. M., and Hollow, R. D. (1992). Factors predicting postoperative sentence scores in postlinguistically deaf adult cochlear implant patients. *Annals of Otology, Rhinology & Laryngology*, 101(4):342–348.

- Boisvert, I., Reis, M., Au, A., Cowan, R., and Dowell, R. C. (2020). Cochlear implantation outcomes in adults: A scoping review. *PLoS One*, 15(5):e0232421.
- Bönitz, H., Kopp, B., Buechner, A., Lunner, T., Lyxell, B., and Finke, M. (2018). Event-related neuronal responses to acoustic novelty in single-sided deaf cochlear implant users: Initial findings. *Clinical Neurophysiology*, 129(1):133–142.
- Bosnyak, D. J., Eaton, R. A., and Roberts, L. E. (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebral cortex*, 14(10):1088–1099.
- Bottari, D., Bednaya, E., Dormal, G., Villwock, A., Dzhelyova, M., Grin, K., Pietrini, P., Ricciardi, E., Rossion, B., and Röder, B. (2020). Eeg frequency-tagging demonstrates increased left hemispheric involvement and crossmodal plasticity for face processing in congenitally deaf signers. *Neuroimage*, 223:117315.
- Brunet, D., Murray, M. M., and Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Computational intelligence and neuroscience*, 2011.
- Buechner, A., Brendel, M., Lesinski-Schiedat, A., Wenzel, G., Frohne-Buechner, C., Jaeger, B., and Lenarz, T. (2010). Cochlear implantation in unilateral deaf subjects associated with ipsilateral tinnitus. *Otology & Neurotology*, 31(9):1381–1385.
- Butera, I. M., Stevenson, R. A., Mangus, B. D., Woynaroski, T. G., Gifford, R. H., and Wallace, M. T. (2018). Audiovisual temporal processing in postlingually deafened adults with cochlear implants. *Scientific Reports*, 8(1):1–12.
- Buzsáki, G., Traub, R. D., and Pedley, T. A. (2003). The cellular basis of EEG activity. *Current practice of clinical electroencephalography*, pages 1–11.
- Caldwell, M. T., Jiam, N. T., and Limb, C. J. (2017). Assessment and improvement of sound quality in cochlear implant users. *Laryngoscope investigative otolaryngology*, 2(3):119–124.
- Campbell, J. and Sharma, A. (2013). Compensatory changes in cortical resource allocation in adults with hearing loss. *Frontiers in systems neuroscience*, 7:71.
- Cappe, C., Thut, G., Romei, V., and Murray, M. M. (2010). Auditory–visual multisensory interactions in humans: timing, topography, directionality, and sources. *Journal of Neuroscience*, 30(38):12572–12580.
- Carlson, M. L. (2020). Cochlear implantation in adults. *New England Journal of Medicine*, 382(16):1531–1542.
- Chen, L.-C., Sandmann, P., Thorne, J. D., Bleichner, M. G., and Debener, S. (2016). Cross-modal functional reorganization of visual and auditory cortex in adult cochlear implant users identified with fnirs. *Neural plasticity*, 2016.
- Chen, L.-C., Sandmann, P., Thorne, J. D., Herrmann, C. S., and Debener, S. (2015). Association of concurrent fNIRS and EEG signatures in response to auditory and visual stimuli. *Brain Topography*, 28(5):710–725.

- Ciorba, A., Bianchini, C., Pelucchi, S., and Pastore, A. (2012). The impact of hearing loss on the quality of life of elderly adults. *Clinical interventions in aging*, 7:159.
- Colletti, L., Shannon, R., and Colletti, V. (2012). Auditory brainstem implants for neurofibromatosis type 2. *Current opinion in otolaryngology & head and neck surgery*, 20(5):353–357.
- Cone-Wesson, B. and Wunderlich, J. (2003). Auditory evoked potentials from the cortex: audiology applications. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 11(5):372–377.
- Corbetta, M. and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature reviews neuroscience*, 3(3):201–215.
- Crowley, K. E. and Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical neurophysiology*, 115(4):732–744.
- Cullington, H., Kitterick, P., Darnton, P., Finch, T., Greenwell, K., Riggs, C., Weal, M., Walker, D.-M., Sibley, A., et al. (2022). Telemedicine for adults with cochlear implants in the United Kingdom (CHOICE): Protocol for a prospective interventional multisite study. *JMIR Research Protocols*, 11(4):e27207.
- Cunningham, L. L. and Tucci, D. L. (2017). Hearing loss in adults. *New England Journal of Medicine*, 377(25):2465–2473.
- da Silva, F. L. (2013). EEG and MEG: relevance to neuroscience. *Neuron*, 80(5):1112–1128.
- Davidson, L. S., Geers, A. E., and Brenner, C. (2010). Cochlear implant characteristics and speech perception skills of adolescents with long-term device use. *Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 31(8):1310.
- Davis, P. A. (1939). Effects of acoustic stimuli on the waking human brain. *Journal of neurophysiology*, 2(6):494–499.
- Dazert, S., Thomas, J. P., Loth, A., Zahnert, T., and Stöver, T. (2020). Cochlear implantation: Diagnosis, indications, and auditory rehabilitation results. *Deutsches Ärzteblatt International*, 117(41):690.
- Debener, S., Hine, J., Bleack, S., and Eyles, J. (2008). Source localization of auditory evoked potentials after cochlear implantation. *Psychophysiology*, 45(1):20–24.
- Deep, N. L., Dowling, E. M., Jethanamest, D., and Carlson, M. L. (2019). Cochlear implantation: an overview. *Journal of Neurological Surgery Part B: Skull Base*, 80(02):169–177.
- Desai, S., Stickney, G., and Zeng, F.-G. (2008). Auditory-visual speech perception in normal-hearing and cochlear-implant listeners. *The Journal of the Acoustical Society of America*, 123(1):428–440.

- DeSaSouza, S. (2022). A historical review of the development of cochlear implants. *Cochlear Implants*, pages 1–7.
- Dhanasingh, A. and DeSaSouza, S. (2022). Evolution of cochlear implant technology over the last 35 years. In *Cochlear Implants*, pages 19–34. Springer.
- Diller, G. (2009). (Re) habilitation nach Versorgung mit einem Kochleaimplantat. *HNO*, 57(7):649–656.
- Dorbeau, C., Galvin, J., Fu, Q.-J., Legris, E., Marx, M., and Bakhos, D. (2018). Binaural perception in single-sided deaf cochlear implant users with unrestricted or restricted acoustic hearing in the non-implanted ear. *Audiology and Neurotology*, 23(3):187–197.
- Doucet, M., Bergeron, F., Lassonde, M., Ferron, P., and Lepore, F. (2006). Cross-modal reorganization and speech perception in cochlear implant users. *Brain*, 129(12):3376–3383.
- Downs, D. W. (1982). Effects of hearing aid use on speech discrimination and listening effort. *Journal of Speech and Hearing Disorders*, 47(2):189–193.
- Drennan, W. R. and Rubinstein, J. T. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *Journal of rehabilitation research and development*, 45(5):779.
- Drijvers, L. and Özyürek, A. (2017). Visual context enhanced: The joint contribution of iconic gestures and visible speech to degraded speech comprehension. *Journal of Speech, Language, and Hearing Research*, 60(1):212–222.
- Dürschmid, S., Reichert, C., Hinrichs, H., Heinze, H.-J., Kirsch, H. E., Knight, R. T., and Deouell, L. Y. (2019). Direct evidence for prediction signals in frontal cortex independent of prediction error. *Cerebral Cortex*, 29(11):4530–4538.
- Fagel, S. and Clemens, C. (2004). An articulation model for audiovisual speech synthesis—determination, adjustment, evaluation. *Speech Communication*, 44(1–4):141–154.
- Fellinger, J., Holzinger, D., Gerich, J., and Goldberg, D. (2007). Mental distress and quality of life in the hard of hearing. *Acta Psychiatrica Scandinavica*, 115(3):243–245.
- Ferraro, J. A. (2010). Electrocochleography: a review of recording approaches, clinical applications, and new findings in adults and children. *Journal of the American Academy of Audiology*, 21(03):145–152.
- Fetterman, B. L. and Domico, E. H. (2002). Speech recognition in background noise of cochlear implant patients. *Otolaryngology-Head and Neck Surgery*, 126(3):257–263.
- Finke, M., Büchner, A., Ruigendijk, E., Meyer, M., and Sandmann, P. (2016a). On the relationship between auditory cognition and speech intelligibility in cochlear implant users: An ERP study. *Neuropsychologia*, 87:169–181.

- Finke, M., Sandmann, P., Bönitz, H., Kral, A., and Büchner, A. (2016b). Consequences of stimulus type on higher-order processing in single-sided deaf cochlear implant users. *Audiology and Neurotology*, 21(5):305–315.
- Friesen, L. M., Shannon, R. V., Baskent, D., and Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America*, 110(2):1150–1163.
- Fryauf-Bertschy, H., Tyler, R. S., Kelsay, D. M., Gantz, B. J., and Woodworth, G. G. (1997). Cochlear implant use by prelingually deafened children: the influences of age at implant and length of device use. *Journal of speech, language, and hearing research*, 40(1):183–199.
- Fu, Q.-J. and Nogaki, G. (2005). Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. *Journal of the Association for Research in Otolaryngology*, 6(1):19–27.
- Galvin, J., Fu, Q., Wilkinson, E., Mills, D., Hagan, S., Lupo, E., Padilla, M., and Shannon, R. (2018). Benefits of cochlear implantation for single sided deafness: data from the House Clinic-USC-UCLA FDA trial. *Journal of Hearing Science*, 8(2).
- García-Larrea, L., Lukaszewicz, A.-C., and Mauguière, F. (1992). Revisiting the oddball paradigm. non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia*, 30(8):723–741.
- Gatehouse, S. and Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *British journal of audiology*, 24(1):63–68.
- Gates, G. A., Anderson, M. L., McCurry, S. M., Feeney, M. P., and Larson, E. B. (2011). Central auditory dysfunction as a harbinger of alzheimer dementia. *Archives of Otolaryngology-Head & Neck Surgery*, 137(4):390–395.
- Gillis, M., Decruy, L., Vanthornhout, J., and Francart, T. (2022). Hearing loss is associated with delayed neural responses to continuous speech. *European Journal of Neuroscience*, 55(6):1671–1690.
- Giraud, A. L., Price, C. J., Graham, J. M., and Frackowiak, R. S. (2001a). Functional plasticity of language-related brain areas after cochlear implantation. *Brain*, 124(7):1307–1316.
- Giraud, A.-L., Price, C. J., Graham, J. M., Truy, E., and Frackowiak, R. S. (2001b). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, 30(3):657–664.
- Giraud, A. L., Truy, E., and Frackowiak, R. (2001c). Imaging plasticity in cochlear implant patients. *Audiology and Neurotology*, 6(6):381–393.
- Glennon, E., Svirsky, M. A., and Froemke, R. C. (2020). Auditory cortical plasticity in cochlear implant users. *Current opinion in neurobiology*, 60:108–114.

- Glick, H. and Sharma, A. (2017). Cross-modal plasticity in developmental and age-related hearing loss: clinical implications. *Hearing Research*, 343:191–201.
- Godey, B., Schwartz, D., De Graaf, J., Chauvel, P., and Liegeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clinical neurophysiology*, 112(10):1850–1859.
- Gordon, K. A., Jiwani, S., and Papsin, B. (2013). Benefits and detriments of unilateral cochlear implant use on bilateral auditory development in children who are deaf. *Frontiers in Psychology*, 4:719.
- Gosselin, P. A. and Gagne, J.-P. (2011). Older adults expend more listening effort than young adults recognizing speech in noise.
- Goudey, B., Plant, K., Kiral, I., Jimeno-Yepes, A., Swan, A., Gambhir, M., Büchner, A., Kludt, E., Eikelboom, R. H., Sucher, C., et al. (2021). A multicenter analysis of factors associated with hearing outcome for 2,735 adults with cochlear implants. *Trends in Hearing*, 25:23312165211037525.
- Grant, K. W., Walden, B. E., and Seitz, P. F. (1998). Auditory-visual speech recognition by hearing-impaired subjects: Consonant recognition, sentence recognition, and auditory-visual integration. *The Journal of the Acoustical Society of America*, 103(5):2677–2690.
- Gravel, J. S., White, K. R., Johnson, J. L., Widen, J. E., Vohr, B. R., James, M., Kennalley, T., Maxon, A. B., Spivak, L., Sullivan-Mahoney, M., et al. (2005). A multisite study to examine the efficacy of the otoacoustic emission/automated auditory brainstem response newborn hearing screening protocol.
- Grech, R., Cassar, T., Muscat, J., Camilleri, K. P., Fabri, S. G., Zervakis, M., Xanthopoulos, P., Sakkalis, V., and Vanrumste, B. (2008). Review on solving the inverse problem in EEG source analysis. *Journal of neuroengineering and rehabilitation*, 5(1):1–33.
- Green, K. M., Julyan, P. J., Hastings, D. L., and Ramsden, R. T. (2005). Auditory cortical activation and speech perception in cochlear implant users: effects of implant experience and duration of deafness. *Hearing research*, 205(1-2):184–192.
- Green, T., Faulkner, A., and Rosen, S. (2019). Computer-based connected-text training of speech-in-noise perception for cochlear implant users. *Trends in Hearing*, 23:2331216519843878.
- Grossmann, W., Brill, S., Moeltner, A., Mlynski, R., Hagen, R., and Radeloff, A. (2016). Cochlear implantation improves spatial release from masking and restores localization abilities in single-sided deaf patients. *Otology & Neurotology*, 37(6):658–664.
- Hammes, D. M., Willis, M., Novak, M. A., Edmondson, D. M., Rotz, L. A., and Thomas, J. F. (2002). Early identification and cochlear implantation: critical factors for spoken language development. *Annals of Otology, Rhinology & Laryngology*, 111(5_suppl):74–78.

- Han, J.-H. and Dimitrijevic, A. (2020). Acoustic change responses to amplitude modulation in cochlear implant users: Relationships to speech perception. *Frontiers in Neuroscience*, 14:124.
- Han, J.-H., Lee, J., and Lee, H.-J. (2021). Ear-specific hemispheric asymmetry in unilateral deafness revealed by auditory cortical activity. *Frontiers in Neuroscience*, 15:698718.
- Hari, R., Pelizzzone, M., Mäkelä, J., Hällström, J., Leinonen, L., et al. (1987). Neuro-magnetic responses of the human auditory cortex to on-and offsets of noise bursts. *Audiology*, 26:31–43.
- Heimler, B., Pavani, F., and Amedi, A. (2018). Implications of cross-modal and intra-modal plasticity for the education and rehabilitation of deaf children and adults. *Evidence-Based Practices in Deaf Education*, page 323.
- Heimler, B., Weisz, N., and Collignon, O. (2014). Revisiting the adaptive and maladaptive effects of crossmodal plasticity. *Neuroscience*, 283:44–63.
- Helfer, K. S. and Vargo, M. (2009). Speech recognition and temporal processing in middle-aged women. *Journal of the American Academy of Audiology*, 20(04):264–271.
- Henkin, Y., Tetin-Schneider, S., Hildesheimer, M., and Kishon-Rabin, L. (2009). Cortical neural activity underlying speech perception in postlingual adult cochlear implant recipients. *Audiology and Neurotology*, 14(1):39–53.
- Henkin, Y., Yaar-Soffer, Y., Steinberg, M., and Muchnik, C. (2014). Neural correlates of auditory-cognitive processing in older adult cochlear implant recipients. *Audiology and Neurotology*, 19(Suppl. 1):21–26.
- Hey, M., Böhnke, B., Mewes, A., Munder, P., Mauger, S. J., and Hocke, T. (2021). Speech comprehension across multiple ci processor generations: Scene dependent signal processing. *Laryngoscope investigative otolaryngology*, 6(4):807–815.
- Hey, M., Hocke, T., Böhnke, B., and Mauger, S. J. (2019). ForwardFocus with cochlear implant recipients in spatially separated and fluctuating competing signals—introduction of a reference metric. *International journal of audiology*, 58(12):869–878.
- Hey, M., Hocke, T., Mauger, S., and Müller-Deile, J. (2016). A clinical assessment of cochlear implant recipient performance: implications for individualized map settings in specific environments. *European Archives of Oto-Rhino-Laryngology*, 273(11):4011–4020.
- Hochmair-Desoyer, I., Schulz, E., Moser, L., and Schmidt, M. (1997). The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users. *The American journal of otology*, 18(6 Suppl):S83–S83.
- Holden, L. K., Finley, C. C., Firszt, J. B., Holden, T. A., Brenner, C., Potts, L. G., Gotter, B. D., Vanderhoof, S. S., Mispagel, K., Heydebrand, G., et al. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear and hearing*, 34(3):342.

- Hoppe, U., Liebscher, T., and Hornung, J. (2017). Cochlear implant fitting strategies. *HNO*, 65(7):546–551.
- House, W. F. and Urban, J. (1973). Long term results of electrode implantation and electronic stimulation of the cochlea in man. *Annals of Otology, Rhinology & Laryngology*, 82(4):504–517.
- Hughes, K. C. and Galvin, K. L. (2013). Measuring listening effort expended by adolescents and young adults with unilateral or bilateral cochlear implants or normal hearing. *Cochlear implants international*, 14(3):121–129.
- Humes, L. E. (2019). The world health organization’s hearing-impairment grading system: an evaluation for unaided communication in age-related hearing loss. *International journal of audiology*, 58(1):12–20.
- Humes, L. E., Busey, T. A., Craig, J., and Kewley-Port, D. (2013). Are age-related changes in cognitive function driven by age-related changes in sensory processing? *Attention, Perception, & Psychophysics*, 75(3):508–524.
- Hummel, F. C. and Cohen, L. G. (2005). Drivers of brain plasticity. *Current opinion in neurology*, 18(6):667–674.
- Isaacson, J. and Vora, N. M. (2003). Differential diagnosis and treatment of hearing loss. *American family physician*, 68(6):1125–1132.
- James, C. J., Karoui, C., Laborde, M.-L., Lepage, B., Molinier, C.-É., Tartayre, M., Escudé, B., Deguine, O., Marx, M., and Fraysse, B. (2019). Early sentence recognition in adult cochlear implant users. *Ear and Hearing*, 40(4):905–917.
- Kaas, J. H. and Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences*, 97(22):11793–11799.
- Kaiser, A. R., Kirk, K. I., Lachs, L., and Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *Journal of Speech, Language, and Hearing Research*, 46(2):390–404.
- Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., and Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, 22(4):751–761.
- Kawase, T., Hori, Y., Ogawa, T., Sakamoto, S., Suzuki, Y., and Katori, Y. (2015). Importance of visual cues in hearing restoration by auditory prosthesis. In *Interface Oral Health Science 2014*, pages 119–127. Springer, Tokyo.
- Kawase, T., Sakamoto, S., Hori, Y., Maki, A., Suzuki, Y., and Kobayashi, T. (2009). Bimodal audio–visual training enhances auditory adaptation process. *Neuroreport*, 20(14):1231–1234.
- Kay-Rivest, E., Roland, J. T., and Friedmann, D. R. (2022). *Cochlear Implants for Single-Sided Deafness*, pages 303–313. Springer Nature Singapore, Singapore.

- Kennedy, K. L. and Singh, A. K. (2017). Middle ear cholesteatoma.
- Kent, R. (1992). Auditory processing of speech. *Central Auditory Processing*. St Louis: Mosby, pages 93–103.
- Key, A. P. (2016). Human auditory processing: Insights from cortical event-related potentials. *AIMS Neuroscience*, 3(2):141–162.
- Kim, B. G., Kim, J. W., Park, J. J., Kim, S. H., Kim, H. N., and Choi, J. Y. (2015). Adverse events and discomfort during magnetic resonance imaging in cochlear implant recipients. *JAMA Otolaryngology–Head & Neck Surgery*, 141(1):45–52.
- Kim, M.-B., Shim, H.-Y., Jin, S. H., Kang, S., Woo, J., Han, J. C., Lee, J. Y., Kim, M., Cho, Y.-S., Moon, I. J., et al. (2016). Cross-modal and intra-modal characteristics of visual function and speech perception performance in postlingually deafened, cochlear implant users. *PLoS One*, 11(2):e0148466.
- Kitterick, P. T., Lucas, L., and Smith, S. N. (2015). Improving health-related quality of life in single-sided deafness: a systematic review and meta-analysis. *Audiology and Neurotology*, 20(Suppl. 1):79–86.
- Kral, A., Hartmann, R., Tillein, J., Heid, S., and Klinke, R. (2001). Delayed maturation and sensitive periods in the auditory cortex. *Audiology and Neurotology*, 6(6):346–362.
- Kral, A., Hubka, P., Heid, S., and Tillein, J. (2013). Single-sided deafness leads to unilateral aural preference within an early sensitive period. *Brain*, 136(1):180–193.
- Kral, A., Kronenberger, W. G., Pisoni, D. B., and O'Donoghue, G. M. (2016). Neurocognitive factors in sensory restoration of early deafness: a connectome model. *The Lancet Neurology*, 15(6):610–621.
- Kral, A. and O'Donoghue, G. M. (2010). Profound deafness in childhood. *New England Journal of Medicine*, 363(15):1438–1450.
- Kral, A. and Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends in neurosciences*, 35(2):111–122.
- Kral, A., Tillein, J., Heid, S., Klinke, R., and Hartmann, R. (2006). Cochlear implants: cortical plasticity in congenital deprivation. *Progress in brain research*, 157:283–402.
- Kraus, N. and McGee, T. (1993). Clinical implications of primary and nonprimary pathway contributions to the middle latency response generating system. *Ear and hearing*, 14(1):36–48.
- Kraus, N. and Nicol, T. (2009). Auditory evoked potentials. In *Encyclopedia of neuroscience*. Springer Berlin Heidelberg.
- Krueger, B., Joseph, G., Rost, U., Strau-Schier, A., Lenarz, T., and Buechner, A. (2008). Performance groups in adult cochlear implant users: speech perception results from 1984 until today. *Otology & Neurotology*, 29(4):509–512.

- Lau, E. F., Phillips, C., and Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature reviews neuroscience*, 9(12):920–933.
- Layer, N., Weglage, A., Müller, V., Meister, H., Lang-Roth, R., Walger, M., Murray, M. M., and Sandmann, P. (2022). The timecourse of multisensory speech processing in unilaterally stimulated cochlear implant users revealed by ERPs. *NeuroImage: Clinical*, 34:102982.
- Lazard, D., Giraud, A.-L., Gnansia, D., Meyer, B., and Sterkers, O. (2012a). Understanding the deafened brain: implications for cochlear implant rehabilitation. *European annals of otorhinolaryngology, head and neck diseases*, 129(2):98–103.
- Lazard, D. S., Collette, J.-L., and Perrot, X. (2012b). Speech processing: from peripheral to hemispheric asymmetry of the auditory system. *The Laryngoscope*, 122(1):167–173.
- Lazard, D. S., Vincent, C., Venail, F., Van de Heyning, P., Truy, E., Sterkers, O., Skarzynski, P. H., Skarzynski, H., Schauwers, K., O’Leary, S., et al. (2012c). Pre-, per-and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model over time. *PloS one*, 7(11):e48739.
- Lee, J. S., Lee, D. S., Oh, S. H., Kim, C. S., Kim, J.-W., Hwang, C. H., Koo, J., Kang, E., Chung, J.-K., and Lee, M. C. (2003). Pet evidence of neuroplasticity in adult auditory cortex of postlingual deafness. *Journal of Nuclear Medicine*, 44(9):1435–1439.
- Legris, E., Galvin, J., Roux, S., Aoustin, J.-M., and Bakhos, D. (2020). Development of cortical auditory responses to speech in noise in unilaterally deaf adults following cochlear implantation. *Plos one*, 15(9):e0239487.
- Lehmann, D. (1987). Principles of spatial analysis. *Handbook of electroencephalography and clinical neurophysiology: Methods of analysis of brain electrical and magnetic signals*, 1:309–354.
- Lehmann, D. and Skrandies, W. (1980). Reference-free identification of components of checkerboard-evoked multichannel potential fields. *Electroencephalography and clinical neurophysiology*, 48(6):609–621.
- Lenarz, T. (2018). Cochlear implant-state of the art. *GMS current topics in otorhinolaryngology, head and neck surgery*, 16:Doco4.
- Li, H., Helpard, L., Ekeroot, J., Rohani, S. A., Zhu, N., Rask-Andersen, H., Ladak, H. M., and Agrawal, S. (2021). Three-dimensional tonotopic mapping of the human cochlea based on synchrotron radiation phase-contrast imaging. *Scientific reports*, 11(1):1–8.
- Lidestam, B., Moradi, S., Pettersson, R., and Ricklefs, T. (2014). Audiovisual training is better than auditory-only training for auditory-only speech-in-noise identification. *The Journal of the Acoustical Society of America*, 136(2):EL142–EL147.
- Lim, H. H. and Lenarz, T. (2015). Auditory midbrain implant: research and development towards a second clinical trial. *Hearing research*, 322:212–223.

- Limb, C. J. and Roy, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hearing research*, 308:13–26.
- Lin, F. R., Metter, E. J., O'Brien, R. J., Resnick, S. M., Zonderman, A. B., and Ferrucci, L. (2011). Hearing loss and incident dementia. *Archives of neurology*, 68(2):214–220.
- Lin, F. R., Yaffe, K., Xia, J., Xue, Q.-L., Harris, T. B., Purchase-Helzner, E., Satterfield, S., Ayonayon, H. N., Ferrucci, L., Simonsick, E. M., et al. (2013). Hearing loss and cognitive decline in older adults. *JAMA internal medicine*, 173(4):293–299.
- Lotfi, Y., Hasanlifar, M., Moossavi, A., Bakhshi, E., and Ajalloueyan, M. (2021). Effect of “spatially separated speech in noise training” on speech perception in noise in children with bimodal fitting. *The Egyptian Journal of Otolaryngology*, 37(1):1–10.
- Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT press.
- Luck, S. J. and Kappenman, E. S. (2011). *The Oxford handbook of event-related potential components*. Oxford University Press.
- Ludwig, A. A., Meuret, S., Battmer, R.-D., Schönwiesner, M., Fuchs, M., and Ernst, A. (2021). Sound localization in single-sided deaf participants provided with a cochlear implant. *Frontiers in psychology*, 12.
- Lütkenhöner, B. and Steinsträter, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiology and Neurotology*, 3(2-3):191–213.
- Ma, W. J., Zhou, X., Ross, L. A., Foxe, J. J., and Parra, L. C. (2009). Lip-reading aids word recognition most in moderate noise: a bayesian explanation using high-dimensional feature space. *PloS one*, 4(3):e4638.
- Macherey, O. and Carlyon, R. P. (2014). Cochlear implants. *Current Biology*, 24(18):R878–R884.
- Macías, Á. R., Barreiro, S. A. B., González, J. C. F., and de Miguel, Á. R. (2019). Cochlear implant in single-sided deafness children and adults. In *Advances in Rehabilitation of Hearing Loss*. IntechOpen.
- Makeig, S., Debener, S., Onton, J., and Delorme, A. (2004). Mining event-related brain dynamics. *Trends in cognitive sciences*, 8(5):204–210.
- Marshall, T. R., O'Shea, J., Jensen, O., and Bergmann, T. O. (2015). Frontal eye fields control attentional modulation of alpha and gamma oscillations in contralateral occipitoparietal cortex. *Journal of Neuroscience*, 35(4):1638–1647.
- Maslin, M. R., Taylor, M., Plack, C. J., and Munro, K. J. (2015). Enhanced intensity discrimination in the intact ear of adults with unilateral deafness. *The Journal of the Acoustical Society of America*, 137(6):EL408–EL414.
- Matusz, P. J., Thelen, A., Amrein, S., Geiser, E., Anken, J., and Murray, M. M. (2015). The role of auditory cortices in the retrieval of single-trial auditory–visual object memories. *European Journal of Neuroscience*, 41(5):699–708.

- Mauger, S. J., Warren, C. D., Knight, M. R., Goorevich, M., and Nel, E. (2014). Clinical evaluation of the Nucleus® 6 cochlear implant system: Performance improvements with SmartSound iq. *International journal of audiology*, 53(8):564–576.
- McCormack, A. and Fortnum, H. (2013). Why do people fitted with hearing aids not wear them? *International journal of audiology*, 52(5):360–368.
- Meister, H., Winter, I. S., Wächter, M., Sandmann, P., and Abdellatif, K. H. (2022). A virtual reality-based method for examining audiovisual prosody perception. *arXiv e-prints*, pages arXiv–2209.
- Meredith, M. A. and Stein, B. E. (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain research*, 365(2):350–354.
- Merzenich, M. M., Van Vleet, T. M., and Nahum, M. (2014). Brain plasticity-based therapeutics. *Frontiers in human neuroscience*, 8:385.
- Michel, C. M. and Koenig, T. (2018). EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *Neuroimage*, 180:577–593.
- Michel, C. M., Koenig, T., Brandeis, D., Gianotti, L. R., and Wackermann, J. (2009). *Electrical neuroimaging*. Cambridge University Press.
- Michel, C. M. and Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *Neuroimage*, 61(2):371–385.
- Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., and De Peralta, R. G. (2004). EEG source imaging. *Clinical neurophysiology*, 115(10):2195–2222.
- Mitchell, T. V. and Maslin, M. T. (2007). How vision matters for individuals with hearing loss. *International journal of audiology*, 46(9):500–511.
- Mortensen, M. V., Mirz, F., and Gjedde, A. (2006). Restored speech comprehension linked to activity in left inferior prefrontal and right temporal cortices in postlingual deafness. *Neuroimage*, 31(2):842–852.
- Mosnier, I., Bebear, J.-P., Marx, M., Fraysse, B., Truy, E., Lina-Granade, G., Mondain, M., Sterkers-Artières, F., Bordure, P., Robier, A., et al. (2014). Predictive factors of cochlear implant outcomes in the elderly. *Audiology and Neurotology*, 19(Suppl. 1):15–20.
- Muigg, F., Bliem, H. R., Kühn, H., Seebacher, J., Holzner, B., and Weichbold, V. W. (2020). Cochlear implantation in adults with single-sided deafness: generic and disease-specific long-term quality of life. *European Archives of Oto-Rhino-Laryngology*, 277(3):695–704.
- Müller, J. (2005). Gestörtes Hören: Die apparative Versorgung der Schwerhörigkeit: Cochlea-Implantate und Hirnstammimplantate: Aktuelle Entwicklungen der letzten 10 Jahre. *Laryngo-Rhino-Otologie*, 84(S 1):60–73.

- Müller-Deile, J., Kiefer, J., Wyss, J., Nicolai, J., and Battmer, R. (2008). Performance benefits for adults using a cochlear implant with adaptive dynamic range optimization (adro): a comparative study. *Cochlear implants international*, 9(1):8–26.
- Müller-Deile, J., Schmidt, B., and Rudert, H. (1995). Effects of noise on speech discrimination in cochlear implant patients. *The Annals of Otology, Rhinology & Laryngology. Supplement*, 166:303–306.
- Murray, M. M., Brunet, D., and Michel, C. M. (2008). Topographic ERP analyses: a step-by-step tutorial review. *Brain topography*, 20(4):249–264.
- Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical neurophysiology*, 118(12):2544–2590.
- Näätänen, R. and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, 24(4):375–425.
- Naito, Y., Hirano, S., Honjo, I., Okazawa, H., Ishizu, K., Takahashi, H., Fujiki, N., Shiomi, Y., Yonekura, Y., and Konishi, J. (1997). Sound-induced activation of auditory cortices in cochlear implant users with post-and prelingual deafness demonstrated by positron emission tomography. *Acta oto-laryngologica*, 117(4):490–496.
- Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., Fink, N. E., Team, C. I., et al. (2010). Spoken language development in children following cochlear implantation. *Jama*, 303(15):1498–1506.
- Nordvik, Ø., Laugen Heggdal, P. O., Brännström, J., Vassbotn, F., Aarstad, A. K., and Aarstad, H. J. (2018). Generic quality of life in persons with hearing loss: a systematic literature review. *BMC Ear, Nose and Throat Disorders*, 18(1):1–13.
- Oba, S. I., Fu, Q.-J., and Galvin III, J. J. (2011). Digit training in noise can improve cochlear implant users' speech understanding in noise. *Ear and hearing*, 32(5):573.
- O'Donoghue, G. M., Nikolopoulos, T. P., and Archbold, S. M. (2000). Determinants of speech perception in children after cochlear implantation. *The Lancet*, 356(9228):466–468.
- Oh, S.-H., Kim, C.-S., Kang, E. J., Lee, D. S., Lee, H. J., Chang, S. O., Ahn, S.-h., Hwang, C. H., Park, H. J., and Koo, J. W. (2003). Speech perception after cochlear implantation over a 4-year time period. *Acta oto-laryngologica*, 123(2):148–153.
- Olds, C., Pollonini, L., Abaya, H., Larky, J., Loy, M., Bortfeld, H., Beauchamp, M. S., and Oghalai, J. S. (2016). Cortical activation patterns correlate with speech understanding after cochlear implantation. *Ear and hearing*, 37(3):e160.
- Olusanya, B. O., Davis, A. C., and Hoffman, H. J. (2019). Hearing loss grades and the international classification of functioning, disability and health. *Bulletin of the World Health Organization*, 97(10):725.

- Pantev, C., Dinnesen, A., Ross, B., Wollbrink, A., and Knief, A. (2006). Dynamics of auditory plasticity after cochlear implantation: a longitudinal study. *Cerebral cortex*, 16(1):31–36.
- Pantev, C., Eulitz, C., Hampson, S., Ross, B., and Roberts, L. (1996). The auditory evoked “off” response: Sources and comparison with the “on” and the “sustained” responses. *Ear and hearing*, 17(3):255–265.
- Peterson, D. C., Reddy, V., and Hamel, R. N. (2018). Neuroanatomy, auditory pathway.
- Peterson, N. N., Schroeder, C. E., and Arezzo, J. C. (1995). Neural generators of early cortical somatosensory evoked potentials in the awake monkey. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, 96(3):248–260.
- Picton, T. W. (2010). *Human auditory evoked potentials*. Plural Publishing.
- Picton, T. W., Hillyard, S. A., Krausz, H. I., and Galambos, R. (1974). Human auditory evoked potentials. i: Evaluation of components. *Electroencephalography and clinical neurophysiology*, 36:179–190.
- Pisoni, D. B., Kronenberger, W. G., Harris, M. S., and Moberly, A. C. (2017). Three challenges for future research on cochlear implants. *World Journal of Otorhinolaryngology-Head and Neck Surgery*, 3(4):240–254.
- Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as ‘asymmetric sampling in time’. *Speech communication*, 41(1):245–255.
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical neurophysiology*, 118(10):2128–2148.
- Pollack, I. (1975). Auditory informational masking. *The Journal of the Acoustical Society of America*, 57(S1):S5–S5.
- Prince, P., Paul, B. T., Chen, J., Le, T., Lin, V., and Dimitrijevic, A. (2021). Neural correlates of visual stimulus encoding and verbal working memory differ between cochlear implant users and normal-hearing controls. *European Journal of Neuroscience*.
- Probst, R., Grevers, G., and Iro, H. (2018). *Basic otorhinolaryngology: a step-by-step learning guide*. New York: Georg Thieme Verlag.
- Quaresima, V. and Ferrari, M. (2019). Functional near-infrared spectroscopy (fNIRS) for assessing cerebral cortex function during human behavior in natural/social situations: a concise review. *Organizational Research Methods*, 22(1):46–68.
- Radecke, J.-O., Schierholz, I., Kral, A., Lenarz, T., Murray, M. M., and Sandmann, P. (2022). Distinct multisensory perceptual processes guide enhanced auditory recognition memory in older cochlear implant users. *NeuroImage: Clinical*, 33:102942.
- Romei, L., Wambacq, I. J., Besing, J., Koehnke, J., and Jerger, J. (2011). Neural indices of spoken word processing in background multi-talker babble. *International journal of audiology*, 50(5):321–333.

- Ross, B. and Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing research*, 248(1-2):48–59.
- Ross, L. A., Saint-Amour, D., Leavitt, V. M., Javitt, D. C., and Foxe, J. J. (2007). Do you see what I am saying? Exploring visual enhancement of speech comprehension in noisy environments. *Cerebral cortex*, 17(5):1147–1153.
- Ross, M. (2004). The audiogram: Explanation and significance. *Hearing Loss Association of America*, 25(3):29–33.
- Rouger, J., Lagleyre, S., Démonet, J.-F., Fraysse, B., Deguine, O., and Barone, P. (2012). Evolution of crossmodal reorganization of the voice area in cochlear-implanted deaf patients. *Human brain mapping*, 33(8):1929–1940.
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., and Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proceedings of the National Academy of Sciences*, 104(17):7295–7300.
- Salmelin, R. and Baillet, S. (2009). Electromagnetic brain imaging. *Human brain mapping*, 30(6):1753.
- Sandmann, P., Dillier, N., Eichele, T., Meyer, M., Kegel, A., Pascual-Marqui, R. D., Marcar, V. L., Jäncke, L., and Debener, S. (2012). Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain*, 135(2):555–568.
- Sandmann, P., Eichele, T., Buechler, M., Debener, S., Jäncke, L., Dillier, N., Hugdahl, K., and Meyer, M. (2009). Evaluation of evoked potentials to dyadic tones after cochlear implantation. *Brain*, 132(7):1967–1979.
- Sandmann, P., Plotz, K., Hauthal, N., de Vos, M., Schönfeld, R., and Debener, S. (2015). Rapid bilateral improvement in auditory cortex activity in postlingually deafened adults following cochlear implantation. *Clinical Neurophysiology*, 126(3):594–607.
- Schaul, N. (1998). The fundamental neural mechanisms of electroencephalography. *Electroencephalography and clinical neurophysiology*, 106(2):101–107.
- Schierholz, I., Finke, M., Kral, A., Büchner, A., Rach, S., Lenarz, T., Dengler, R., and Sandmann, P. (2017). Auditory and audio–visual processing in patients with cochlear, auditory brainstem, and auditory midbrain implants: An eeg study. *Human brain mapping*, 38(4):2206–2225.
- Schierholz, I., Finke, M., Schulte, S., Hauthal, N., Kantzke, C., Rach, S., Büchner, A., Dengler, R., and Sandmann, P. (2015). Enhanced audio–visual interactions in the auditory cortex of elderly cochlear-implant users. *Hearing research*, 328:133–147.
- Searight, F. T., Singh, R., and Peterson, D. C. (2022). Otitis media with effusion. In *StatPearls [Internet]*. StatPearls Publishing.
- Seitz, A. R., Kim, R., and Shams, L. (2006). Sound facilitates visual learning. *Current biology*, 16(14):1422–1427.

- Senkowski, D., Pomper, U., Fitzner, I., Engel, A. K., and Kral, A. (2014). Beta-band activity in auditory pathways reflects speech localization and recognition in bilateral cochlear implant users. *Human brain mapping*, 35(7):3107–3121.
- Shahin, A. J., Roberts, L. E., Miller, L. M., McDonald, K. L., and Alain, C. (2007). Sensitivity of EEG and MEG to the N1 and P2 auditory evoked responses modulated by spectral complexity of sounds. *Brain topography*, 20(2):55–61.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270(5234):303–304.
- Sharma, A., Dorman, M. F., and Kral, A. (2005). The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. *Hearing research*, 203(1-2):134–143.
- Sharma, A., Dorman, M. F., and Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear and hearing*, 23(6):532–539.
- Sharma, A., Gilley, P. M., Dorman, M. F., and Baldwin, R. (2007). Deprivation-induced cortical reorganization in children with cochlear implants. *International journal of audiology*, 46(9):494–499.
- Simmons, F. B., Epley, J. M., Lummis, R. C., Guttman, N., Frishkopf, L. S., Harmon, L. D., and Zwicker, E. (1965). Auditory nerve: electrical stimulation in man. *Science*, 148(3666):104–106.
- Sorrentino, A. and Piana, M. (2017). *Inverse Modeling for MEG/EEG Data*, pages 239–253. Springer International Publishing, Cham.
- Speck, I., Arndt, S., Thurow, J., Blazhenets, G., Aschendorff, A., Meyer, P. T., and Frings, L. (2020). 18F-FDG PET imaging of the inferior colliculus in asymmetric hearing loss. *Journal of Nuclear Medicine*, 61(3):418–422.
- Stevenson, R., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., Kurela, L. R., Siemann, J. K., James, T. W., and Wallace, M. T. (2014). Identifying and quantifying multisensory integration: a tutorial review. *Brain topography*, 27(6):707–730.
- Stevenson, R., Sheffield, S. W., Butera, I. M., Gifford, R. H., and Wallace, M. (2017). Multisensory integration in cochlear implant recipients. *Ear and hearing*, 38(5):521.
- Stone, J. L., Calderon-Arnulphi, M., Watson, K. S., Patel, K., Mander, N. S., Suss, N., Fino, J., and Hughes, J. R. (2009). Brainstem auditory evoked potentials—a review and modified studies in healthy subjects. *Journal of Clinical Neurophysiology*, 26(3):167–175.
- Strelnikov, K., Rosito, M., and Barone, P. (2011). Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PloS one*, 6(3):e18344.

- Strelnikov, K., Rouger, J., Barone, P., and Deguine, O. (2009). Role of speechreading in audiovisual interactions during the recovery of speech comprehension in deaf adults with cochlear implants. *Scandinavian journal of psychology*, 50(5):437–444.
- Strelnikov, K., Rouger, J., Demonet, J.-F., Lagleyre, S., Fraysse, B., Deguine, O., and Barone, P. (2013). Visual activity predicts auditory recovery from deafness after adult cochlear implantation. *Brain*, 136(12):3682–3695.
- Strelnikov, K., Rouger, J., Lagleyre, S., Fraysse, B., Démonet, J.-F., Déguine, O., and Barone, P. (2015). Increased audiovisual integration in cochlear-implanted deaf patients: independent components analysis of longitudinal positron emission tomography data. *European Journal of Neuroscience*, 41(5):677–685.
- Stropahl, M., Bauer, A.-K. R., Debener, S., and Bleichner, M. G. (2018). Source-modeling auditory processes of EEG data using EEGLAB and brainstorm. *Frontiers in neuroscience*, 12:309.
- Stropahl, M., Chen, L.-C., and Debener, S. (2017a). Cortical reorganization in postlingually deaf cochlear implant users: intra-modal and cross-modal considerations. *Hearing research*, 343:128–137.
- Stropahl, M., Plotz, K., Schönfeld, R., Lenarz, T., Sandmann, P., Yovel, G., De Vos, M., and Debener, S. (2015). Cross-modal reorganization in cochlear implant users: Auditory cortex contributes to visual face processing. *Neuroimage*, 121:159–170.
- Stropahl, M., Schellhardt, S., and Debener, S. (2017b). McGurk stimuli for the investigation of multisensory integration in cochlear implant users: The Oldenburg Audio Visual Speech Stimuli (OLAVS). *Psychonomic bulletin & review*, 24(3):863–872.
- Sumby, W. H. and Pollack, I. (1954). Visual contribution to speech intelligibility in noise. *The journal of the acoustical society of america*, 26(2):212–215.
- Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: a user-friendly application for meg/eeeg analysis. *Computational intelligence and neuroscience*, 2011.
- Tajudeen, B. A., Waltzman, S. B., Jethanamest, D., and Svirsky, M. A. (2010). Speech perception in congenitally deaf children receiving cochlear implants in the first year of life. *Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 31(8):1254.
- Távora-Vieira, D., Marino, R., Acharya, A., and Rajan, G. P. (2015). The impact of cochlear implantation on speech understanding, subjective hearing performance, and tinnitus perception in patients with unilateral severe to profound hearing loss. *Otology & Neurotology*, 36(3):430–436.
- Thelen, A. and Murray, M. M. (2013). The efficacy of single-trial multisensory memories. *Multisensory research*, 26(5):483–502.

- Timm, L., Agrawal, D., Viola, F. C., Sandmann, P., Debener, S., Büchner, A., Dengler, R., and Wittfoth, M. (2012). Temporal feature perception in cochlear implant users. *PloS one*, 7(9):e45375.
- Tremblay, K., Kraus, N., McGee, T., Ponton, C., Otis, B., et al. (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear and hearing*, 22(2):79–90.
- Tremblay, K. L., Ross, B., Inoue, K., McClannahan, K., and Collet, G. (2014). Is the auditory evoked P2 response a biomarker of learning? *Frontiers in systems neuroscience*, 8:28.
- Turgeon, C., Lazzouni, L., Lepore, F., and Ellemberg, D. (2014). An objective auditory measure to assess speech recognition in adult cochlear implant users. *Clinical Neurophysiology*, 125(4):827–835.
- Van de Heyning, P., Gavilán, J., Godey, B., Hagen, R., Hagr, A., Kameswaran, M., Li, Y., Manoj, M., Mlynski, R., O'Driscoll, M., et al. (2022). Worldwide variation in cochlear implant candidacy. *J Int Adv Otol*, 18(3):196–202.
- van de Heyning, P., Vermeire, K., Diebl, M., Nopp, P., Anderson, I., and De Ridder, D. (2008). Incapacitating unilateral tinnitus in single-sided deafness treated by cochlear implantation. *Annals of Otology, Rhinology & Laryngology*, 117(9):645–652.
- van de Rijt, L. P., Roye, A., Mylanus, E. A., van Opstal, A. J., and van Wanrooij, M. M. (2019). The principle of inverse effectiveness in audiovisual speech perception. *Frontiers in human neuroscience*, 13:335.
- Vaughan Jr, H. G. (1982). The neural origins of human event-related potentials. *Annals of the New York Academy of Sciences*, 388(1):125–138.
- Viola, F. C., De Vos, M., Hine, J., Sandmann, P., Bleeck, S., Eyles, J., and Debener, S. (2012). Semi-automatic attenuation of cochlear implant artifacts for the evaluation of late auditory evoked potentials. *Hearing research*, 284(1-2):6–15.
- Vohn, R., Fimm, B., Weber, J., Schnitker, R., Thron, A., Spijkers, W., Willmes, K., and Sturm, W. (2007). Management of attentional resources in within-modal and cross-modal divided attention tasks: An fMRI study. *Human brain mapping*, 28(12):1267–1275.
- Vroomen, J. and Stekelenburg, J. J. (2010). Visual anticipatory information modulates multisensory interactions of artificial audiovisual stimuli. *Journal of Cognitive Neuroscience*, 22(7):1583–1596.
- Wagener, K., Brand, T., and Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache. I-III: Design, Optimierung und Evaluation des Oldenburger Satztests (Development and evaluation of a sentence test for the German language. I-III: Design, optimization and evaluation of the Oldenburg sentence test). *Zeitschrift für Audiologie (Audiological Acoustics)*, 38:4–15.

- Walter, W. G., Cooper, R., Aldridge, V., McCallum, W., and Winter, A. (1964). Contingent negative variation: an electric sign of sensori-motor association and expectancy in the human brain. *nature*, 203(4943):380–384.
- Wettstein, V. G. and Probst, R. (2018). Right ear advantage of speech audiometry in single-sided deafness. *Otology & Neurotology*, 39(4):417.
- Wilcox, T. and Biondi, M. (2015). fNIRS in the developmental sciences. *Wiley Interdisciplinary Reviews: Cognitive Science*, 6(3):263–283.
- Williges, B., Wesarg, T., Jung, L., Geven, L. I., Radeloff, A., and Jürgens, T. (2019). Spatial speech-in-noise performance in bimodal and single-sided deaf cochlear implant users. *Trends in Hearing*, 23:2331216519858311.
- Wilson, B. S. and Dorman, M. F. (2008). Cochlear implants: a remarkable past and a brilliant future. *Hearing research*, 242(1-2):3–21.
- Winkler, I., Denham, S., and Escera, C. (2013). *Auditory Event-related Potentials*, pages 1–29. Springer New York, New York, NY.
- Winn, M. B. (2016). Rapid release from listening effort resulting from semantic context, and effects of spectral degradation and cochlear implants. *Trends in Hearing*, 20:2331216516669723.
- Wolfe, J., Neumann, S., Marsh, M., Schafer, E., Lianos, L., Gilden, J., O'Neill, L., Arkis, P., Menapace, C., Nel, E., et al. (2015). Benefits of adaptive signal processing in a commercially available cochlear implant sound processor. *Otology & Neurotology*, 36(7):1181–1190.
- Wong, K., Kozin, E. D., Kanumuri, V. V., Vachicouras, N., Miller, J., Lacour, S., Brown, M. C., and Lee, D. J. (2019). Auditory brainstem implants: recent progress and future perspectives. *Frontiers in Neuroscience*, page 10.
- World Health Organization (2021). World report on hearing. *Licence: CC BY-NC-SA 3.0 IGO*.
- Zahnert, T. (2011). The differential diagnosis of hearing loss. *Deutsches Ärzteblatt International*, 108(25):433.
- Zatorre, R. J. and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral cortex*, 11(10):946–953.
- Zeitler, D. M. and Dorman, M. F. (2019). Cochlear implantation for single-sided deafness: a new treatment paradigm. *Journal of Neurological Surgery Part B: Skull Base*, 80(02):178–186.
- Zeng, F.-G. (2004). Trends in cochlear implants. *Trends in amplification*, 8(1):1–34.
- Zeng, F.-G., Popper, A. N., and Fay, R. R. (2011). *Auditory prostheses: New horizons*, volume 39. Springer Science & Business Media.

- Zeng, F.-G., Tang, Q., and Lu, T. (2014). Abnormal pitch perception produced by cochlear implant stimulation. *PloS one*, 9(2):e88662.
- Zhao, E. E., Dornhoffer, J. R., Loftus, C., Nguyen, S. A., Meyer, T. A., Dubno, J. R., and McRackan, T. R. (2020). Association of patient-related factors with adult cochlear implant speech recognition outcomes: a meta-analysis. *JAMA Otolaryngology–Head & Neck Surgery*, 146(7):613–620.
- Zhao, F., Stephens, S., Sim, S., and Meredith, R. (1997). The use of qualitative questionnaires in patients having and being considered for cochlear implants. *Clinical Otolaryngology & Allied Sciences*, 22(3):254–259.

COLOPHON

This document was typeset using classicthesis style developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". It is available for L^AT_EX and L^yX at

<https://bitbucket.org/amiede/classicthesis/>

Final Version as of March 20, 2023 (classicthesis v4.6).

EIDESSTATTLICHE ERKLÄRUNG

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen - einschließlich Tabellen, Karten und Abbildungen -, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, sind in jedem Einzelfall als Entlehnung kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertationsschrift noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss der Promotion nicht ohne Genehmigung der / des Vorsitzenden des IPHS-Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Frau PD Dr. Sandmann betreut worden.

Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen.

Übersicht der Publikationen:

Layer, N., Weglage, A., Müller, V., Meister, H., Lang-Roth, R., Walger, M., Murray, M. M., & Sandmann, P. (2022). The timecourse of multisensory speech processing in unilaterally stimulated CI implant users revealed by ERPs. *NeuroImage: Clinical*, 34, 102982. DOI: 10.1016/j.nicl.2022.102982.

Weglage, A., Müller, V., **Layer, N.**, Abdel-Latif, K. H., Lang-Roth, R., Walger, M., & Sandmann, P. (2022). Side-of-implantation effect on functional asymmetry in the auditory cortex of single-sided deaf CI-implant users. *Brain Topography*, 1–22. DOI: 10.1007/s10548-022-00902-3.

Layer, N., Weglage, A., Müller, V., Meister, H., Lang-Roth, R., Walger, M., Murray, M. M. & Sandmann, P. (2022). Electrophysiological differences and similarities in audiovisual speech processing in CI users with unilateral and bilateral hearing loss. *Current Research in Neurobiology*, 100059. DOI: 10.1016/j.crneur.2022.100059.

Layer, N., Abdel-Latif K. H., Radecke J-O., Müller, V., Weglage, A., Lang-Roth, R., Walger, M. & Sandmann, P. (2022). The influence of background noise and noise reduction by “ForwardFocus” on audiovisual speech perception in bimodal CI users: an ERP study. *Clinical Neurophysiology*, in revision.

Ich versichere, dass ich alle Angaben wahrheitsgemäß nach bestem Wissen und Gewissen gemacht habe und verpflichte mich, jedmögliche, die obigen Angaben betreffenden Veränderungen, dem IPHS-Promotionsausschuss unverzüglich mitzuteilen.

15.03.2023

Natalie Layer