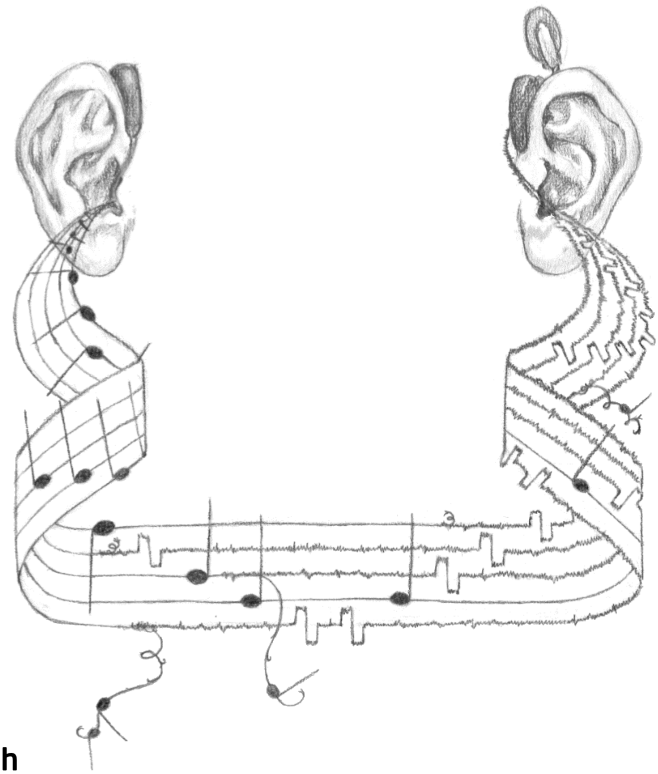
CONTRIBUTIONS TO
HEARING RESEARCH

Volume 40

Niclas A. Janssen

Binaural Streaming in Cochlear Implant Patients



DTU Health Tech
Department of Health Technology

Binaural Streaming in Cochlear Implant Patients

Ph.D. thesis by
Niclas A. Janssen

Preliminary version: June 18, 2019

HEARING SYSTEMS

Technical University of Denmark

2019

This Ph.D. dissertation is the result of a research project carried out at the Hearing Systems Group, Department of Health Technology, Technical University of Denmark.

The project was partly financed by the Centre of Excellence in Hearing and Speech Sciences (CHeSS; 1/6), the Oticon Foundation (1/3), Oticon Medical (1/6), and by the Technical University of Denmark (1/3).

Supervisors

Assoc. Prof. Dr. Jeremy Marozeau
Hearing Systems Group
Department of Health Technology
Technical University of Denmark
Kgs. Lyngby, Denmark
Dr. Lars Bramsløw
Eriksholm Research Centre
Snekkersten, Denmark
Dr. Søren Riis
Oticon Medical
Smørum, Denmark

Abstract

Cochlear implants (CI), a form of neural prostheses, are used to aid patients with severe-to-profound hearing loss by representing sounds in a way fundamentally different compared to acoustic hearing. The healthy auditory system constantly analyses the complex sound environment around us. It forms sound objects from acoustic components, that, when linked over time, can form a so-called stream, such as a voice or melody. Normal-hearing listeners perform this auditory streaming with their two ears, unaware of the binaural processes that integrate the information from both ears into common sound objects. Especially for bimodal CI patients with a CI in one ear and hearing aid (HA) on the other ear, one sound can be perceived very differently across ears. Bilateral CI patients, implanted in both ears, may also perceive sounds differently in each ear. For CI patients the potential differences in a sound's representation across ears may influence how they process binaural sounds. The studies in this thesis center on the question of whether bimodal CI patients can use their devices effectively together and whether they can build common streams from binaural sounds like normal-hearing listeners would. In a first study, bimodal CI patients were interviewed and only a minority indicated to perceive sounds from one source as a simple, uniform sound object. Most participants reported to perceive something more complex, non-uniform, instead, with some describing their percepts as two entirely separate sounds with strong differences in pitch and loudness across ears. The second study dealt with the development of a psychoacoustic experiment to test the listeners' abilities to form streams from sounds delivered separately to the two ears. The paradigm was validated with normal-hearing listeners. A third study centered on bilateral CI patients, as these could, like bimodal CI patients, show differences in their binaural streaming. The bimodal CI patients were the focus of a fourth study. Results suggest that at least some bilateral CI listeners can form streams binaurally, but none of the bimodal CI patients tested, mirroring the patient's reports from the first study. This suggests that the extended daily use of the bimodal devices altered the binaural processes of the bimodal CI patients so that they do not integrate binaural stimuli into a common stream based on perceptual similarity as normal-hearing listeners would. These results could help to guide the development of CI-candidacy criteria, clinical fitting of the devices, and new strategies for their simulation.

Resumé

Cochlear implantater (CI), en slags neurale proteser, bruges til at hjælpe patienter med alvorligt høretab ved at repræsentere lyde på en fundamentalt anderledes måde i forhold til akustisk hørelse. Det normale auditive system analyserer hele tiden det komplekse lydmiljø omkring os. Det danner lydobjekter fra akustiske komponenter, der, når de er sammenhørende over tid, kan danne en såkaldt strøm, som f.eks. en stemme eller en melodi. Normalthørende personer kan udføre denne auditive *streaming* med begge ører, helt ubevidst om de binaurale processer, der integrerer informationen fra begge ører og danner fælles lydobjekter. Især for bimodale CI-brugere med et CI på det ene øre og et høreapparat (HA) på det andet øre, kan en lyd opfattes meget forskelligt på hvert øre. Bilaterale CI-brugere, implanteret i begge ører, kan også opleve lyde på hvert øre forskelligt. De potentielle forskelle i oplevelser på hvert øre hos CI-patienter kan påvirke hvordan de behandler binaural lyde. Undersøgelserne i denne afhandling centrerer omkring spørgsmålet om hvorvidt bimodale CI-brugere kan bruge deres to enheder effektivt sammen, og om de kan forme sammenhængende strømme fra binaural lyde på sammen måde som normalthørende personer. I det første studie blev bimodale CI-brugere interviewet. Kun få brugere indikerede at opleve lyd fra en kilde som et simpelt, ensartet lydobjekt. De fleste reporterede at opleve noget mere komplekst og ikke-ensartet hvoraf nogle beskrev deres opfattelse som to helt separate lydstrømme og med store forskelle i tonehøjde og lydstyrke på hvert øre. Den andet studieomhandlede udviklingen af et psykoakustisk eksperiment for at teste CI-brugers evne til at forme strømme af lyde, der leveres separat til hvert øre. Paradigmet blev valideret med normalthørende. Det tredje studie omhandlede bilaterale CI-brugere, idet de kunne, som bimodale CI-brugere, vise forskelle i deres evne til at udføre auditiv *streaming*. Bimodale CI-brugere var emnet i det fjerde studie. Resultaterne tyder på, at idet mindste nogle bilaterale CI-brugere kan danne lydstrømme binauralt, mens ingen af de bimodale CI-brugere kan. Disse resultaterne afspejler det første interview studie. Overordnet tyder dette på, at vedvarende, daglig brug af bimodal stimulering ændrer binaural processer hos bimodale CI-brugere, således at de ikke integrerer binaural stimuli til én lydstrøm med ens opfattelse, som normale hørelyttere ellers kan. Resultater kan bidrage til at guide udviklingen af kriterier for CI-kandidatur, klinisk montering af enhederne og nye stimuleringsstrategier af enhederne.

Acknowledgements

My Ph.D. studies have been a time to focus on the intriguing area of cochlear implants and gave me the opportunity to develop both in my scientific and technical skills, as well as personally. Hence, I would first like to sincerely thank my three supervisors, Lars Bramsløw, Søren Riis, and Jeremy Marozeau, who gave me this opportunity, a lot of helpful feedback, and discussions. I am also thankful for the funding provided by the Centre for Excellence in Hearing and Speech Sciences (CHeSS), Oticon Medical, and the Oticon Foundation. I would further like to thank my colleagues at DTU Hearing Systems for their support, inspiring exchange, and their efforts to create a great work environment over the years. Special thanks to Michal Fereczkowski for helping me to a good start at DTU and his feedback on the survey questions, Pernille Holtegaard, who also helped with the questions and translated the bimodal questionnaire to Danish, François Guérit for his work on a toolbox for CI stimulation, Tanmayee Uday Pathre for collecting part of the normal-hearing listener data, and Charlotte Navntoft for her help with the Danish abstract. It would also have been impossible to conduct any studies without the help of the volunteers who participated in the experiments. Thank you very much for your time, effort, as well as the insightful discussions around the experiment and life with cochlear implants. For their help in finding more CI participants I would like to thank Cecilia Fernandez Samar, Per Cayé-Thomasen, and Line Husted Baungaard from Gentofte Hospital, Jesper Hvass Schmidt from Odense Hospital, Kirsten Volmer Worsøe from Roskilde communication center, and Dorte Lundgaard from Hellerup communication center. Still, the studies on bilateral and bimodal CI patients could not have been conducted without the possibility to spend some time at the German Hearing Center in Hannover. For this opportunity, I would like to thank Andreas Büchner and all the people working there, with a special thank you to Eugen Kludt for his help.
I would also not be in this position if it had not been for my parents' love and support over the years. Danke, dass ihr immer für mich da seid! Likewise, I would like to thank my girlfriend, Thirsa, for her love, support, and the design of the cover picture. Dankjewel, Liefste!

Sincerely,
Niclas A. Janßen

Related publications

Journal papers

- Janssen, N. A., A. Büchner, L. Bramsløw, S. Riis, and J. Marozeau (**2019a**). "Self-assessment of binaural integration in bimodal cochlear implant patients," J. Speech Lang. Hear. Res. (under revision)
- Janssen, N. A., L. Bramsløw, S. Riis, and J. Marozeau (**2019b**). "A modification of the scale illusion into a detection-task for assessment of binaural streaming," J. Acoust. Soc. Am. **145** (6), EL457-EL462

Conference papers

- Janssen, N. A., L. Bramsløw, S. Riis, and J. Marozeau (**2018**). ”The scale illusion detection task: Objective assessment of binaural fusion in normal-hearing listeners,” Proceedings of the International Symposium on Auditory and Audiological Research, 6, 207-213.

Other publications

- Janssen, N. A., A. Büchner, L. Bramsløw, S. Riis, and J. Marozeau (**2019**). The Bimodal Integration Questionnaire, from chapter 2, Self-Assessment of Binaural Integration in Bimodal Cochlear Implant Patients, is available for download under https://doi.org/10.5281/zenodo.2574015.

Contents

1 General introduction

The human sense of hearing is paramount to our life, providing us not only with an omnidirectional way to sense our surroundings but also, perhaps most importantly, giving us the ability to communicate verbally with each other. For this reason, an impairment of the auditory system may require fundamental changes to how we live our lives. A hearing impairment can make it difficult to lead a conversation in a busy, noisy place, so that an everyday occurrence becomes a challenging task. Insecurity, distress, and social isolation can follow (Erikson-Mangold and Carlsson, 1991). Fortunately, technology can be used to aid people with hearing-impairment and even deafness in the form of hearing aids (HA; Kates, 2008) and cochlear implants (CI; Zeng et al., 2008). Apart from the question how well this technology can improve speech reception and other outcome measures for hearing-impaired patients, it can be fascinating to explore how this technology changes a person's perception of the world. Especially with CIs, which represent sounds in a fundamentally different way compared to normal hearing or HAs, the changes may be hard to understand for anyone who is not utilizing such a device. Nevertheless, one can use rational thinking and experiments to gain knowledge about what one may call another sense of hearing altogether.

Day-to-day life exposes us to complex sound environments and the auditory system continuously analyzes these, forming sound objects from overlapping acoustic components and linking these objects over time into so-called streams, such as a voice or melody. The healthy auditory system performs this based on signals collected by the two ears (Bregman, 1990). Our auditory system must process the input from our two ears and associate the components that come from the same sound source. If it did not perform this integration, we would perceive every sound twice, separated for left and right ear. The perceptual changes introduced by a cochlear implant may affect the perception of sounds differently in each ear and, hence, may also affect the binaural processing. This applies whether a patient is aided with a CI unilaterally, possibly with a HA on the other ear, or utilizes bilateral CIs. Notably, the outcome of cochlear implantation not only depends on the devices, but equally on how well the patient's auditory system adapts to the new stimulation (Polonenko et al., 2019; Reiss et al., 2014b).

Differences between HAs and CIs are fundamental in both stimulation and perception, and may become most striking considering the so-called bimodal CI patients, who are aided with a CI in one ear and a HA on the other. In the bimodal case, the CI's settings are generally adjusted first, independently of the HA. Then the HA is adjusted in its settings, although if and how can vary (Scherf and Arnold, 2014). So far, some fitting strategies for bimodal patients aim to align percepts across ears in loudness, but do not address potential differences in percepts, such as pitch, timbre, or timing. Hence, this configuration raises questions about how well the two kinds of devices work together. These studies' aim to understand whether bimodal CI patients can combine stimuli from their two ears into single percepts like normal-hearing listeners. Such an ability may be essential for a natural perception and can affect listening performance and effort (Chang et al., 2016). While several studies have assessed this for the grouping of short dichotic stimuli (binaural fusion), the abilities of CI listeners to form streamed percepts from binaural stimuli have not received as much attention. Changes in the patients' binaural processes could explain much of the difficulties CI patients exhibit when it comes to tasks relying on binaural hearing and streaming, such as the localization of sounds and speech understanding in a noisy environment (Goupell et al., 2013; Kan et al., 2013; Seeber, 2004; Staisloff et al., 2016; Yoon et al., 2013). The motivation for the studies described in this thesis lies in finding out if these binaural processes work differently in bimodal CI patients.

This chapter begins with a description of how the healthy auditory system analyses the acoustic environment, followed by a brief overview of how HAs and CIs present sounds to patients and which limitations CI patients might experience. After that, a description of how CIs can affect the auditory system's ability to analyze the acoustic environment is given, along with a section dedicated to the situation of bimodal CI patients.This introduction concludes with an overview of this thesis' chapters and the four studies they describe.

1.1 Auditory scene analysis and streaming in normal-hearing listeners

Everyday situations provide listeners with a complex acoustic scenery, originating from a variety of sound sources and sound reflections that lead to reverberation. The healthy human auditory system can decode these signals, even though there is no explicit information about the number of sound sources present in the complex mixture. Colin Cherry famously described the problem to listen selectively to one source out of such a complex mixture as the cocktail party problem (Cherry, 1953). Later, the concept of auditory scene analysis was established, describing the formation of auditory objects from complex simultaneous sounds and the formation of streams, auditory objects linked over time (Bregman, 1990; Shinn-Cunningham et al., 2017). Such streams are formed based on grouping of simultaneous (Micheyl and Oxenham, 2010b), as well as sequential stimuli (Moore and Gockel, 2002). Simultaneous grouping relies upon a spectral decomposition of the sound, to determine which frequency-components belong to which sound object. These components can then be integrated into a stream over time through sequential grouping.
The formation of auditory objects is generally ruled by similarity or proximity in perceptual aspects (Bregman, 1990). Conversely, differences in any such perceptual aspect can be used as a cue to segregate sounds (Moore and Gockel, 2002). An auditory object or stream can for example be formed of components with synchronous onset or duration (Bregman, 1990), similar frequency content (Bregman and Campbell, 1971; Noorden, 1975), timbre (Singh and Bregman, 1997), temporal envelope (Cusack and Roberts, 2000; Iverson, 1995; Vliegen et al., 1999b), spatial characteristics (David et al., 2015; Sach and Bailey, 2004; Stainsby et al., 2011), intensity (Noorden, 1975, 1977), and/or temporal coherence (Christiansen et al., 2014; Shamma et al., 2011, 2013). Stimuli that match or are sufficiently similar in such grouping cues can be integrated into one common stream, and stimuli with larger differences in these aspects are likely to be segregated into separate sound objects (Bregman, 1990). For changes over time, perceptual continuity rules their integration, so that more similar, slowly-varying stimuli are likely to form one stream, whereas more different, fast-varying stimuli are likely to be segregated.
Listeners often face difficulties in focusing their attention selectively on one out of multiple concurrent streams (McDermott, 2009). Studies have demonstrated

that selective attention and the segregation of auditory objects and streams are closely intertwined, as the segregation of concurrent streams does not only depend on differences in perceptual aspects between components of the sound but also attention (Bregman, 1990; Moore and Gockel, 2002; Noorden, 1975). The temporal coherence model described in Shamma et al. (2011, 2013) even suggests that attention to an auditory feature is the starting point for auditory object formation, from which further matching features are grouped.
The processes that give rise to the formation of an object or stream are also known to be not necessarily occuring instantaneously, instead auditory streams have been reported to arise gradually over a time of several seconds of listening, referred to as a build-up effect (Anstis and Saida, 1985; Bregman, 1978; Bregman, 1990).
Auditory stream segregation stands at the foundation of how the human brain builds up a percept of the acoustic scene. Similarities and differences among the perceptual quantities of sounds' components are used for grouping and segregation of auditory objects and streams. Consequently, changes to these percepts introduced by listening devices may also fundamentally change a listener's percept of the acoustic environment.

1.2 Hearing aids

Hearing impairment is often assessed using audiograms, measuring the thresholds of sound pressure levels at which a person begins to perceive a pure tone at various frequencies in the range of hearing and comparing them to the average threshold of normal-hearing listeners. A hearing loss can then be defined as a sound pressure level difference to the normal-hearing average (Sataloff and Sataloff, 2005). Hearing aids (HA) address hearing impairment by changing the sounds delivered to the impaired ear, foremost by frequency-band specific amplification, aiming to compensate for the difference in hearing thresholds and lack of compression, but also by processing the sounds to reduce background noise, enhance certain features, or audibility in general (Kates, 2008). The hearing aids' usefulness can reach its limit for severe-to-profound losses, that surpass 70 dB above the threshold for normal-hearing listeners. This also includes patients who only have residual hearing limited to low frequencies. Patients with such extensive hearing losses, areas of poor neural survival along the basilar membrane, inner hair cell loss, or poor speech discrimination are typically considered candidates for cochlear implantation (Flynn et al., 1998; Moore and Alcantara, 2001; Vickers et al., 2016).

1.3 Cochlear implants

CIs can partly restore hearing in patients with hearing losses so severe that patients no longer benefit from HAs (Flynn et al., 1998) and can also be used to give a sense of hearing back to deaf patients (Zeng et al., 2008). Different configurations are possible: Bilateral implantation, unilateral implantation with a normal-hearing ear on the other side (so-called single-sided deafness patients; SSD) or unilateral implantation combined with impaired hearing aided by a HA, a middle ear implant, or bone conduction HA on the other ear (so-called bimodal CI patients). In addition to that, there exist hybrid CI devices that aid residual acoustic hearing in the implanted ear via a HA-component in the CI, also referred to as electro-acoustic stimulation (EAS) (Offeciers et al., 2005; Zeng et al., 2008).

CIs bypass large sections of the auditory system, namely the outer, middle, and inner ear, with an array of electrodes surgically inserted along the basilar membrane of the cochlea. In the healthy auditory system, about 3500 inner hair cells are distributed along the roughly 3.5 cm of the basilar membrane. The sounds' energy is coupled into the cochlea via the oval window and creates travelling waves along the membrane. Consequently, the ends of the inner hair cells (stereocilia) are bend, which leads to electric spikes in the auditory nerve fibers of the spiral ganglia. Due to a tonotopic organization of the cochlea, higher frequencies lead to a maximum excitation closer to the oval window, while lower frequencies lead to a maximum excitation towards the opposing apical end of the basilar membrane. Each point along the basilar membrane has a maximum sensitivity for a characteristic frequency (Oxenham, 2018; Plack, 2018).

The CI's electric stimuli stimulate the auditory nerve fibers in the spiral ganglia directly (Zeng et al., 2008), circumventing inner and outer hair cells. In the healthy cochlea, about 12000 outer hair cells are activated along with the inner hair cells, amplifying the vibrations along the basilar membrane through changes in their length (Oxenham, 2018; Plack, 2018). These outer hair cells are also subject to inhibitory inputs from medial olivocochlear (MOC) efferent nerve fibers, through which top-down processes affect the effective amplification dynamically (Aronoff et al., 2015; Lopez-Poveda et al., 2016). Since the CI bypasses the hair cells and stimulates the auditory nerve directly, the top-down processes cannot affect the percepts, in contrast to acoustic hearing with enough surviving outer hair cells and a HA.

A radio-frequency (RF) link connects the implant to an external sound processor. The electrode array within the cochlea is connected to the RF link's coil-housing, which is surgically placed between the skin and the bone of the skull. The external sound processor features microphones to pick up sound signals and processes them before transmitting the signals to the implant's electrodes (Zeng et al., 2008).

The encoding of pitch information necessitates a fine spectral resolution and ideally, a large coverage of stimulation over the auditory nerve cells located along the extend of the basilar membrane. But the CI electrode-array insertion depths achieved clinically do not generally allow to reach the most apical regions of the cochlea, where the auditory nerve fibers tuned to the lowest frequencies reside (Oxenham, 2018). Two pitch percepts have been reported in CI listeners: Place pitch, elicited by changes in the location of the stimulating electrode, and temporal or rate pitch, elicited by changes in the pulse rate of the stimulation (Oxenham, 2008).

CIs feature typically twelve to twenty-two electrodes distributed over a length of about 13 mm to 26 mm. Some electrode arrays feature shorter designs down to as little as six electrodes distributed over 6 mm, with the aim to preserve residual acoustic hearing (Zeng et al., 2008). The common stimulation strategy for CIs replaces the characteristic tonotopic encoding of pitch information in the healthy cochlea by dividing a spectral range, typically about 200 Hz to 8000 Hz, into bands allocated to the available electrodes. Like in the tonotopic pitch encoding, more apically located electrodes are stimulated by sounds of lower frequency and more basally located electrodes by sounds of higher frequency. Generally, sounds are analyzed separately for each electrode's frequency-band. The sound's envelope is extracted and multiplied with a high-rate biphasic pulse-train carrier signal (Zeng et al., 2008).

Apart from one manufacturer, the temporal pitch percept is not used in clinical CIs. It increases along with the pulse rate at least up to a limit of about 300 Hz (Mcdermott, 2004; Shannon, 1983a; Zeng, 2002). A study by McKay et al. (2000) concluded that these two pitch percepts are generally represented along independent dimensions, perhaps similar to pitch and brightness in acoustic hearing, so that the place pitch in CIs may correspond to a dimension of timbre (Mcdermott, 2004; Oxenham, 2018). CI listeners in Luo et al. (2012) performed better on a melody identification task when temporal and place pitch cues changed congruently than incongruently, suggesting that both cues could be integrated into a common percept, or are at least not completely independent. This becomes clearer considering the results of Lamping et al. (2017), who found that both pulse rate and electrode place induced perceptual dimensions that were linked to pitch and timbre. With only the place pitch cues from the clinical place pitch encoding, CI patients face limitations in pitch perception (Oxenham, 2018).

The loudness of sounds is encoded via variations in the stimulation current level. In a CI fitting process, the device parameters are adjusted to an individual patient. A threshold current level, at which a CI patient begins to perceive the stimulation of a given electrode, and a maximum comfortable current level are defined per electrode. The CI processor maps sounds into the range in between these current levels (Zeng et al., 2008).

The CI represents sounds in a fundamentally different way compared to normal-hearing and hearing aids. Moreover, many parameters can influence the percept that arises from the electrical stimulation, such as the depth of electrode-array insertion, the nerve-fiber survival-state, the distance between the electrode and the modiolus or the spiral ganglia, the patients' duration of deafness and experience with the CI, and the fitting of the device parameters (Oxenham, 2018; Zeng et al., 2008). Since the electrode array resides within the cochlear fluid, the electric stimulation by even a single electrode can spread over a wide area along the cochlea and lead to interactions across channels. In areas with poor or no neural survival, places located further from the electrode can be stimulated due to the current spread, with respective changes in percepts (Oxenham, 2018; Zeng et al., 2008). Even with full neural survival, the spread of excitation leads to the stimulation of a larger ensemble of cells, resulting in a very different neural excitation pattern compared to the healthy auditory system. Hence, a simple pure tone signal is unlikely to evoke percepts that are comparable to the pitch of a pure tone (Landsberger et al., 2016; Liang et al., 1999; Shannon, 1983b). The spread of excitation in current CIs is equivalent to 12 - 24 dB per octave (Bingabr et al., 2008) and limits speech reception (Fredelake and Hohmann, 2012; Jürgens et al., 2019; Zamaninezhad et al., 2017)

Studies of speech reception have shown that beyond about eight channels resp. electrodes, CI patients do not benefit from the additional information presented. Therefore, even for broad-band sounds such as speech, clinical stimulation strategies generally only stimulate a limited number of electrodes (Bingabr et al., 2008; Fishman et al., 1997; Friesen et al., 2001).
To sum up, CIs can provide patients suffering from severe hearing loss with a new means of hearing. These devices represent sounds in a profoundly different way and are optimized to represent speech signals, encoding the envelope of sound signals in a number of bands out of a limited frequency range. They do not encode the sound's fine structure and CI patients also face limitations regarding pitch perception. Hence, the perception of sounds with a CI differs strongly from normal hearing, which may lead to changes in the auditory system.

1.4 Streaming and integration of information across ears in cochlear implant patients

The clinical stimulation strategy provides adequate spectral resolution to understand speech well in quiet ("Speech recognition with primarily temporal cues" 1995; Zeng et al., 2008). Nevertheless, CI patients often struggle to understand one voice among a background of competing sounds or a single melody from concurrent melodies (Galvin et al., 2009a; Hong and Turner, 2006; Nelson et al., 2003).
The spatial cues available through binaural stimulation, namely interaural time differences (ITDs) and interaural level differences (ILDs), can aid the perceptual separation of concurrent talkers in both bilateral and bimodal CI patients (Bernstein et al., 2016), but current CIs do not reproduce spatial cues accurately. This reflects in limited localization abilities (Easwar et al., 2018; Litovsky et al., 2017; Seeber, 2004). These cues are also relevant for stream segregation (Bregman, 1990). But apart from binaural cues, also binaural processes can be affected in CI patients and limit their performance.
A number of studies have investigated whether CI patients are able to form streamed percepts monaurally from sequential (Chatterjee et al., 2006; Cooper and Roberts, 2007; Duran et al., 2012) or simultaneous, concurrent stimuli (Carlyon et al., 2007; Cooper and Roberts, 2010; Deeks and Carlyon, 2004). While

the studies suggest that patients can use either cue for streaming, they generally struggle with situations that require stream segregation with the clinical devices (Galvin et al., 2009a; Hong and Turner, 2006; Nelson et al., 2003). This could be a result of the limited perceptual separation provided via only the place pitch cues clinically, as well as binaural mismatches and changes in binaural processes. It has been demonstrated that CI listeners can benefit from larger separations in both place pitch (Paredes-Gallardo et al., 2018b) and temporal pitch cues (Paredes Gallardo et al., 2018c) to segregate sequential stimuli into separate streams. However, it is not clear whether bimodal and bilateral CI patients form integrated streams from binaurally presented stimuli.

A CI conveys sounds in an entirely different way compared to normal-hearing or a HA, and many factors can influence the percept evoked by the electrical stimulation of the CI (Oxenham, 2018; Zeng et al., 2008). Moreover, CIs, like HAs, are generally subject to a delay from input to output and operate without synchronization across ears (Kates, 2008; Zeng et al., 2008). Hence, whether patients are aided with unilateral or bilateral CIs, their percepts of sound from the same sound source can easily vary across ears in many perceptual aspects that serve as streaming cues, such as pitch, loudness, and timing (Gordon et al., 2017). In the case of bilateral CI patients, for example, the electrodes paired clinically are often not perceived as pitch-matched (Aronoff et al., 2016a). Hence, whether percepts of one sound source are perceived the same across ears ties directly into the question of whether CI patients can utilize the input from their two ears together effectively. If the perceptual proximity across ears is not given, the patients could either fail to form objects binaurally or adapt in various ways to the new stimulation, for example by changes in percepts or in the binaural processes involved in object formation and streaming.

Yoon et al. (2013) investigated the effect of spectral mismatches across ears in a study simulating bilateral CIs with normal-hearing participants and found that spectral mismatches affected the binaural benefit attributed to the squelch effect and redundancy negatively for sentence recognition. Likewise, Aronoff et al. (2015) found that the CI processor's spectral and temporal compression could prevent bimodal CI listeners' from grouping simultaneous binaural stimuli (binaural fusion).

However, such studies using vocoders to simulate the effects of CIs cannot account for possible adaptations in CI listeners due to extended exposure to the stimulation. A range of adaptive behavior has been described for bimodal

CI patients' pitch percepts (Reiss et al., 2014b). The pitch percepts either (1.) reduced the interaural mismatch through adapting pitch percepts, (2.) aligned pitch percepts to a common low pitch percept across multiple electrodes, or (3.) showed no adaptation in pitch percepts over time. In the case of a reduction in interaural mismatches, the perceptual adaptation could also lead to a more normal grouping behavior, while constant pitch percepts could either lead to a lack or adaptations of binaural grouping.

Another way to assess the integration of binaural stimuli is to use recordings of the auditory brainstem, as, e.g., Gordon et al. (2012) did to investigate the role of place cues. They obtained binaural difference (BD) measurements from bilateral CI listeners, also known as the binaural interaction component (BIC). This measure is based on the difference in between the response from binaural stimulation over the sum of the monaural auditory brainstem responses. The measurements indicated that the BD persisted over misalignments in place of stimulation up to four electrodes and vanished at larger interaural mismatches. Gordon et al. (2012) interpreted this as a sign of underdevelopment and/or underutilization of the tonotopic organization of the auditory brainstem in CI listeners. However, the wider range could also be a result of the spread of excitation over a wider area along the cochlea, i.e., a larger collection of auditory nerve fibers. Further, the BD amplitude was largest when the stimuli were balanced in level.

Apart from such objective measures relying on measurements on the auditory brainstem, several studies have investigated binaural fusion in CI patients with listening experiments, usually relying on listener feedback. Fitzgerald et al. (2015) found that some bilateral CI listeners did not report loudness-matched binaural stimuli as fused and/or centered even after a bilateral loudness-alignment had been performed. Suneel et al. (2017) also found that bilateral CI patients often had difficulties fusing sounds and that localization performance correlated with binaural fusion. In clinical practice, matching loudness binaurally is more complex because of the independent automated gain controls (AGC) used in the devices. For bilateral CI patients, the AGC systems are typically the same across ears. But due to the head shadow effect, the devices across ears face different sound pressure levels, so that the AGCs operate at different working points. Thus, the interaural level cues will be distorted without a link for synchronization between the devices. For most bimodal CI patients the AGC systems, coming from independent manufacturers, may operate very differently, leading to further distortions (Veugen et al., 2016). The observed lack of binaural fusion in these studies could be a sign of maladaptation in the processes for binaural grouping under the extended exposure to mismatches in bilateral percepts or prolonged period of deafness. It fits thus, that Gordon et al. (2017) concluded from studies on pediatric CI patients, that deafness or the lack of access to binaural input disrupted the processes of binaural hearing and that current hearing devices were unable to reverse these changes and/or promote expected development.

Such changes in binaural processes surface in several further studies: Normal-hearing listeners fuse dichotic stimuli over a small pitch-range, up to 0.3 octaves. At higher frequency separations, listeners perceive binaural beatings or tones of separate pitches (Brink et al., 1976). It was found that most CI patients, both bimodal and bilateral, may group dichotic stimuli over a far wider pitch range, up to the equivalent of four octaves or even across the entire range of the electrode-array (Kan et al., 2013; Reiss et al., 2014a, 2017; Steel et al., 2015). Such an abnormal grouping behavior could be detrimental to the patients' listening performance if it leads to a grouping of sounds from separate sound sources. But the studies have also shown some variability both in the patients' behavior and the range over which they appear to group dichotic stimuli. In any case, the widened grouping behavior could also fundamentally affect binaural streaming in CI listeners.

As Gordon et al. (2017) argued, it may be necessary to change clinical practice and implement binaural fitting targets, which match pitch, loudness, and timing across devices for bilateral but also bimodal CI patients, to allow them to integrate sounds and form streams binaurally.

1.5 The hearing of bimodal cochlear implant patients

Bimodal CI patients are stimulated electrically by the CI in one ear and acoustically by the HA on the other ear. Hence, they are subject to fundamentally different representations of the sounds around them when one ear is compared to the other (Offeciers et al., 2005; Zeng et al., 2008). Moreover, their residual hearing is often limited, with a mild-to-moderate hearing loss up to 500 Hz, that transitions to a severe-to-profound hearing loss above 500 Hz (Ching et al., 2007; Gifford et al., 2007).

Even in case of a severe hearing loss, the HA component offers potential advantages over a CI, such as access to low-frequency sounds, fine structure, and fundamental frequency cues, which also can aid stream segregation and speech reception (Dorman and Gifford, 2010; Kong et al., 2005; Seeber, 2004; Williges et al., 2015). These cues help bimodal patients to a better recognition of tone in (Mandarin) speech, but not vowels, while the bimodal configuration also appears to reduce listening effort over the unilateral CI alone (Chang et al., 2016). Despite that, Gifford et al. (2007) concluded that bilateral implantation with fully inserted electrode arrays may offer better speech reception for the patients over bimodal aids. One contributing factor is the limited residual hearing of most bimodal CI patients, often with a mild-to-moderate hearing loss up to 500 Hz, that transitions to a severe-to-profound hearing loss at higher frequencies (Ching et al., 2007; Gifford et al., 2007). Nevertheless, the bimodal configuration gives rise to a significantly better speech reception over the unilateral CI alone, despite generally worse performance with the HA alone compared with the CI alone (Bernstein et al., 2016; Gifford et al., 2007; Hamzavi et al., 2004). Likely, a significant part of the bimodal benefit with CI and HA combined is derived from the low frequencies up to 250 Hz or 300 Hz alone, that increase with access to a broader acoustic range, as demonstrated with CI simulations using vocoders (Sheffield et al., 2016; Zamaninezhad et al., 2017). Still, the lack of symmetric binaural stimulation may at least partly explain the differences between bilateral and bimodal CI patients' performance.

Another factor contributing to binaural processing may be a lack of similarity in the representations of a sound source across ears. Ma et al. (2016) compared different frequency mappings in a unilateral CI simulation with normal-hearing listeners and concluded that individualized frequency-mappings were the best basis for optimum speech perception in noise. However, it remains unclear whether the benefit of this optimization is worth the additional time in clinical practice. Also, the loudness across bimodal devices is not aligned for most device combinations, since only some systems attempt to link and align loudness via matching and linked automated gain controls (AGCs) across ears and devices. Using the same compression systems across HA and CI can yield significantly improved speech reception in noise and reduced listening effort compared to independent systems with different AGC systems (Veugen et al., 2016).

An important question for the CI and HA in the bimodal combination is how to adjust their fitting parameters, so that patients can achieve optimum listening performance. The bimodal fitting is generally seen as the fitting of a HA to a unilateral CI, and both devices are not necessarily fitted by the same professionals (Scherf and Arnold, 2014). Hence, the CI fitting procedure follows the same approach as for any unilaterally fitted CI. Current levels for the various frequency bands, resp. electrodes, are adjusted for the listening threshold and comfortable loudness, according to the guideline provided by the CI's manufacturer.

Ching et al. (2004) developed a fitting procedure for HAs worn contralateral to a CI, optimizing their frequency-dependent gain and compression ratio, aiming to achieve a loudness balance across ears. This was used in Morera et al. (2012) with bimodal listeners showing large individual variation in both speech reception and binaural benefit. While a significant binaural summation effect was found for speech and noise from the front, when the noise was spatially separated from the speech, the advantage varied with interaural asymmetry in monaural listening performance. The binaural squelch effect was only significant when the noise was presented to the HA side, but not when it was presented to the CI side. The bimodal stimulation led to significantly better sound localization and subjectively better sound quality. However, the balancing procedure from Ching et al. (2004) had no overall effect on the outcome.

Vroegop et al. (2018) reviewed studies on bimodal CI patients regarding the HA fittings used and came to the conclusion that the standard HA prescription rules, NAL-NL1, NAL-NL2, and DSL, were a good and common starting point for fittings of a bimodal HA. Yet, they found no clear evidence as to how specific decisions contributed to bimodal performance. Frequency lowering or transposition techniques were found to offer no benefit. In Vroegop et al. (2019a) the Adaptive Phonak Digital Bimodal (APDB) fitting formula was compared to a standard HA fitting formula, the NAL-NL2. Significant differences between the two formulas surfaced at and above 2 kHz: The APDB provided less gain and a higher compression ratio for frequencies upwards from 1 kHz. A loudness balancing did not lead to significant deviations from the initial gain targets of the APDB formula. Nevertheless, a bimodal benefit was found for speech reception in quiet and in noise. However, the gain and compression changes of the APDB formula did not result in significant changes to the performance compared to NAL-NL2. Vroegop et al. (2019b) compared other HA fitting methods for bimodal CI patients. While, again, the bimodal stimulation provided superior listening performance over that of the CI alone, overall no significant differences were found in the tested methods (NAL-NL2, and Audiogram+ with either broadband or narrowband loudness balancing).

Some CI manufacturers have released specific guidelines for the bimodal fitting (Advanced Bionics, 2016; Reyes, 2016), but also these describe only the fitting of the HA component after the CI has been adjusted without any regard to the HA. All current fitting formulas appear to facilitate bimodal benefits, but the differences in the specific approaches do not lead to significant changes in listening performance. While some of the fitting procedures attempt to align loudness across ears, none of them attempt alignments in other aspects, such as pitch, timbre or timing. Additionally, these different approaches for the fitting of the HA in the bimodal setting are not necessarily applied in clinical practice as the HA is even not commonly refitted after CI activation and many CI centers lack experience with HA fitting altogether, as reported by Scherf and Arnold (2014).

The differences in stimulation and percepts (Reiss et al., 2014b) across ears could affect the binaural processing of bimodal CI patients auditory system and, therefore, their ability to form streams from binaural stimulation, indicated by the grouping of dichotic stimuli over a wide pitch range in Reiss et al. (2014a), as discussed before. For pediatric bimodal CI patients, one-sided or asymmetric stimulation has been shown to affect binaural processing (Polonenko et al., 2019). While these patients adapt to binaural stimulation and their performance improves over time, children with longer periods of asymmetric hearing may develop an abnormal independence of the bilateral auditory pathways. This may be an effect of mismatches in interaural percepts.

Hence, differences in percepts across ears, not accounted for by current fitting methods, may prevent bimodal CI patients from taking full advantage of their binaural aids through binaural hearing and binaural streaming.

1.6 Aims and overview of the thesis

The motivation for the studies in this thesis was to understand better how well CI and HA work together in bimodal CI patients. Yet, in order to understand the situation for bimodal CI patients, it may be necessary to first understand that of bilateral CI patients, for whom the sound representations in both ears may be more similar, while they are also subject to any limitations and changes introduced by CIs alone. Previous studies did not specifically investigate whether there are differences in between how normal-hearing listeners and either bimodal or bilateral CI patients integrate stimuli from both ears into streams through binaural processing. The studies of this Ph.D. thesis center on the question of whether cochlear implantation affects the abilities to perform such binaural streaming that integrates stimuli from both ears into a single common sound object. Understanding if and how bimodal CI patients form streams from binaural stimulation could provide vital information regarding guidelines for the CI candidacy criteria, optimization of clinical device fitting, and the underlying stimulation strategies for the devices. Studying this for bilateral as well as bimodal CI patients allows to compare these groups and to examine which potential changes might be a result of cochlear implantation and which could be attributed to the combination of HA and CI across ears in the bimodal CI configuration.

As a starting point, *chapter 2* deals with this question from the point of view of the patients themselves, on the example of bimodal CI patients, which could exhibit strong perceptual differences due to the different sound representations by CI and HA. A survey was carried out assessing whether these patients experience sounds as integrated into a common sound object or stream across HA-aided and CI-aided ears in their daily life. Along with this question, differences in pitch and loudness and device reliability regarding sound localization and preferences for sound quality, and whether this could be linked to the patients' self-reported listening performance were assessed.

To assess binaural streaming without subjective descriptions from the patients, a new experimental method was needed. *Chapter 3* describes the method developed to assess binaural streaming and verifies it with normal-hearing listeners. The *chapters 4 and 5* take up this method and detail how it was adapted for bilateral (chapter 4) and bimodal (chapter 5) CI patients. Bilateral CI patients provide the opportunity to perform testing purely with electrical stimulation, therefore giving rise to perhaps the most similar stimulation and sound representations across ears possible with CIs. Hence, they are both by themselves an intriguing subject of research but also an important step on the way to understand binaural streaming in bimodally aided CI patients. Therefore, the new method was first utilized to assess the bilateral CI patients' capability for binaural streaming (chapter 4), before it was adopted for bimodal CI patients, for whom the interaural sound representations through electric and acoustic stimulation are fundamentally different.

Finally, *chapter 6* summarizes the main findings of each chapter and its implications.

2

Self-assessment of binaural integration in bimodal cochlear implant patients[a]

Abstract

Purpose: For bimodal cochlear implant (CI) patients, aided with a contralateral hearing aid (HA), sounds might be perceived differently in each ear, since both devices represent sounds in different ways. Hence, it might useful to investigate if patients integrate the binaural sensations into a singular sound-source percept and how they perform in listening tasks.
Method: A survey offers a quick and simple, yet ecologically valid way to assess binaural integration. 38 experienced bimodal CI patients were interviewed regarding whether they perceived the sound from one sound-source as one sound object and if this influenced their perception differently across ears in six everyday listening-situations. Listening abilities were assessed using the SSQ5 questionnaire (Mertens and Heyning, 2013). The device manufacturer was not a selection criterion.
Results: Some patients reported integrated percepts, whereas the majority reported to perceive not a simple, single sound object along with differences in pitch, loudness, and sound quality across ears. Frequently, the CI was perceived as higher and louder, with better sound quality and better for localizing sounds. Participants with better integration trended towards lower listening-effort, and better speech reception in noise.
Conclusions: Bimodal CI patients might benefit from being able to integrate electric and acoustic percepts regarding listening-effort and speech reception in noise, but only a minority reported such integration, independent of listening-situation. The connection to

[a] This chapter is based on Janssen et al. (2019a, under revision).

the performance-measures is not clear, but these depend not on binaural integration alone. A lack of binaural integration coincided with larger differences pitch and loudness percepts. Conversely, aligning the sound representations across ears could make it easier for patients to binaurally integrate sounds.

2.1 Introduction

Bimodal cochlear implant (CI) patients utilize a hearing aid (HA) on the contralateral ear. These two devices produce auditory sensations in very different ways. HAs provide amplification to acoustic sounds (among other signal processing) and present these to the impaired auditory system through a loudspeaker at the ear. CIs are based around an electrode array inserted into the cochlea, exciting the auditory nerve fibers in the spiral ganglia more directly via electric pulses. Only a limited number of positions along the cochlea can be stimulated due to the finite number of electrodes. The electrode positions are used to convey pitch information, resulting in a limited pitch resolution (Oxenham, 2018; Zeng et al., 2008). The electrode array is suspended inside the cochlear fluid, so that the CI's electric stimulation is subject to current spread, which excites a larger area of cells along the cochlea. This further limits the ability to represent different and precise pitches and heavily influences the sound percepts (Oxenham, 2018; Zeng et al., 2008). Furthermore, HAs and CIs are often produced by different manufacturers, using different signal processing, and fitted independently and by different professionals without common fitting procedures (Scherf and Arnold, 2014). Both devices are also inducing a delay to the sound, and do not operate synchronized. Hence, they introduce changes in various perceptual aspects of the sounds that can serve as cues for auditory streaming, such as pitch, loudness, and timing (Gordon et al., 2017). Moreover, bimodal CI patients often only have a limited range of mild-to-moderate residual hearing in the frequency range up to 500 Hz, with a transition into severe-to-profound hearing loss above this frequency (Ching et al., 2007; Gifford et al., 2007).

Yoon et al. (2013) and Aronoff et al. (2015) demonstrated in CI-simulation studies that spectral mismatches across ears prevented listeners from grouping simultaneous binaural stimuli and affected the binaural benefit arising from the squelch effect and redundancy. Yet, the bimodal combination of HA and CI has been found to improve speech reception in noise (Dorman and Gifford, 2010; Kong et al., 2005; Seeber, 2004; Williges et al., 2015), but the benefit of these bilateral devices varies significantly across patients (Ching et al., 2007; Dunn et al., 2005).

One important aspect explaining the variation in performance of CI listeners may be adaptation to the new stimulation. Reiss et al. (2014b) found large inter-individual differences in term of pitch adaption for bimodal listeners ranging from no change to large pitch changes with time. While adaptation to continued stimulation by a CI can occur, it may vary strongly for an individual patient. In addition to pitch, adaption of the binaural processes can also occur, as results in pediatric CI patients in Polonenko et al. (2019) demonstrate.

This could lead one to question if the bimodal CI patients integrate and exploit the available binaural information in the same way as normal-hearing listeners. One might also wonder how patients will perceive sounds of the same sound-source under combined stimulation from both devices. Patients who integrate the binaural information perceptually like normal-hearing listeners should perceive the sound from one source as one sound object (Bregman, 1990). If bimodal CI patients did not integrate sounds in the same way binaurally, they could perceive something more complex, or even two separate sound objects with different pitch and loudness from their two ears. Consequently, they might struggle in situations where the use of binaural cues is an advantage, such as when localizing sounds and understanding speech in noise, and this could (partly) explain the high variability in patient performance (Litovsky et al., 2017)).

To investigate whether bimodal CI patients integrate the binaural sounds normally, Reiss et al. (2014a) presented various combinations of short loudness-balanced acoustic pure-tone and biphasic pulse-train stimuli as dichotic stimuli. For each combination, the participants reported whether they perceived (1.) one and the same single sound in both ears, (2.) one sound in the left ear only, (3.) one sound in the right ear only, or two sounds with either (4.) the left ear higher in pitch, or (5.) the right ear higher in pitch. Stimuli were interpreted as integrated in case the participants answered same (1.) or if they were reported as lateralized to either side (2., 3.). The results showed abnormally-broad fusion ranges, over up to four octaves. This abnormally-broad binaural integration was weakly correlated to interaural pitch-perception mismatches, which could

indicate a link between both. Reiss et al. (2017), showed a similar wide binaural fusion for bilateral CI patients and pointed out that such a large range for binaural integration of sounds could be detrimental for tasks relying on source segregation, such as the perception of speech in noise.
The integration of binaural input was also investigated with speech stimuli, namely, vowels. In a study by Guérit et al. (2014), bimodal participants were presented with two-formant vowels, one formant per ear. Results indicated that bimodal listeners face difficulties in integrating both sounds into a percept of a single vowel, which NH listeners were capable of in the same study. Reiss et al. (2016) also studied the binaural integration of vowels, where the first formant spectral peak was varied, while the second formant was kept constant. Such stimuli were presented both monaurally, as well as binaurally. The bimodal CI participants showed various kinds of frequency-dependent integration behavior: (1.) similar monaural and binaural percepts, (2.) percepts dominated by one ear, (3.) percepts averaged across ears and (4.) interference caused by the binaural stimulation. However, the integration behavior could not be predicted based upon interaural pitch-perception mismatch or broad pitch-fusion behavior. Still, the variance in integration-behavior is large and illustrates that, indeed, normal binaural integration might be absent in most bimodal CI patients.
Binaural integration could also be studied using objective measures, such as the binaural interaction component (BIC) in auditory brainstem-responses (ABR), the difference in between the response from binaural stimulation to the sum of the monaural responses (Dobie and Norton, 1980). Gordon et al. (2012) used this to show how differences in electrode place, resp. pitch, and level reduced the BIC. However, such measurements are complex, costly, and require a lot of time. Especially the amount of time necessary may prevent their use in clinical practice. It is thus beneficial to explore binaural integration in other ways, which could more easily be used clinically, before resorting to such objective measures.
Our understanding of the differences and difficulties in binaural integration found in the previous studies can be aided by the patients' perspective from their everyday listening experience. Results from this may be more ecologically valid and could foster a better understanding of their performance in listening experiments. A questionnaire provides a relatively easy way to obtain information about their perception of everyday sounds and what specifically patients perceive as different. It also provides the means to let participants indicate

differences in their percepts and whether they experience binaural integration with greater subtlety than a binary answer. This could, for example, be useful if only part of the sounds' representations were integrated, perhaps in a certain spectral range. Moreover, it allows an assessment regardless of which manufacturer's devices the patients use, making it easier to collect a larger number of responses. It has been suggested that the binaural integration of sound-information relies at least partly on the perceptual similarity of sounds in the binaural auditory system (Breebaart et al., 2001; Bregman, 1990; Goupell et al., 2013; Zhou and Durrant, 2003). Thus, we hypothesize (1.) that users who integrate sounds perceive lower perceptual differences across the devices and (2.) rate their performance in everyday listening-tasks better. With fully-integrated percepts, bimodal CI patients might find it less of an effort to listen (Bernstein et al., 2016; Chang et al., 2016; Steel et al., 2015; Veugen et al., 2016), be able to better localize sounds (Kan et al., 2013), and better understand speech in noise, since they can use the combined information from both ears (Gifford et al., 2007; Ma et al., 2016). To investigate whether there is such a relation based on the patients' own perspective, they were given a newly-developed questionnaire for assessing their abilities regarding everyday listening tasks and the perceptual differences across their HA and CI aided ears. The listening performance was assessed in an efficient way via the SSQ5 questionnaire (Mertens and Heyning, 2013).

2.2 Participants

For this study, 38 bimodal CI patients have been recruited for interviewing either at the Technical University of Denmark (14 participants) or at the German Hearing Centre of Hannover, Germany (24 participants). Participants gave their informed written consent and the study was approved by the Science-Ethics Committee for the Capital Region of Denmark (reference H-16036391). No specific ethical approval was required for conduction of a survey at the German Hearing Centre. Participants had at least six months of experience with the bimodal aids. Their age ranged from 35 to 86 with a mean of 61.6 years. One the HA-aided side, participants had sensorineural hearing-loss that ranged from mild to severe, but audiogram data was not generally available for this study. The manufacturer of the devices was not a selection criterion. Ten participants were aided with Advanced Bionics (Stäfa, Switzerland), nineteen with Cochlear (Macquarie University, Australia), five with Med-El (Starnberg, Germany), and three with Oticon Medical CIs (Smørum, Denmark). More information on the participants is split across tab. 2.1 - 2.3.

2.3 Survey design

The aim of the questionnaire was to explore the everyday experience regarding binaural integration of sounds in bimodal CI users in a formal way. Therefore, questions were posed about how patients perceived sounds under combined stimulation by both devices in their daily life and in several everyday listening-situations. The questions were designed in English, the common language of all researchers involved, and then translated to the native languages of the patients to be interviewed. The Danish translation was conducted by a native Danish Audiologist, while the German translation was conducted by the native German author N. A. J. The survey was conducted in personal interviews of the participants, to ensure that they understood the questions and situations described in the same manner.

The questionnaire was split into three sections, denoted A, B, and C. Section A enquires about the development of the hearing impairment, time of experience with the devices and their daily use, since these have been related to changes in binaural processes and can affect the listening performance (Polonenko et al., 2019). Section B consisted of six everyday situations in which the participants should imagine themselves having a conversation to test if common acoustic environments influenced the participants' percepts with the devices. These environments were similar to those common environments described in (Smeds et al., 2015).

These situations were:

- 1.) Quiet (as a reference),
- 2.) A busy restaurant with people talking in the background (multi-talker speech noise),
- 3.) A nature scene with noise from the wind, such as in the forest or at the sea (wind noise),

- 4.) A busy street at rush-hour (traffic noise),
- 5.) Music in the background, e.g., from the radio, and
- 6.) Shopping in a supermarket (a complex mixture with many different noise sources).

For all these six situations, participants had to indicate on a scale from no difference (one sound) to absolutely different (two separate sounds) with numbers from zero to ten whether they integrated the percepts of one single sound source from both devices into one common, uniform percept (question B1). Although this question seems to call for a binary response, either one or two sounds, a continuous scale was used to allow the participant to indicate if there was more subtlety to whether they experience binaural integration, such as if only part of the sounds' representations were integrated.
In addition to this first question, participants had to answer four questions regarding specific differences in percepts across their ears: pitch, loudness, as well as which device was more reliable for localizing sounds and which preferred for sound quality. Pitch and loudness were chosen since interaural differences in these percepts have previously been linked to a lack of binaural integration (Fitzgerald et al., 2015; Gordon et al., 2012; Suneel et al., 2017; Veugen et al., 2016). The device reliability for localization was assessed because localization is based on interaural level difference and time difference cues (Easwar et al., 2018; Fitzgerald et al., 2015) and thus, reliance upon one device signals abnormal behavior and could be a sign of changes in binaural processes, along with the rating regarding binaural integration. The sound quality rating is of interest, since could likewise be affected by differences in percepts across devices. Answers to these questions could be given on a scale running from minus five (HA absolutely higher/louder/better) to plus five (CI absolutely higher/louder/better), to allow participants to express not only if these aspects were balanced or not, but also how strong a perceived difference was experienced.
Questions for a conversation in six everyday situations (section B) were:

- B1) Do you feel there is a difference in perception between your ears in regard to how you perceive the sound from a single sound source? Explanation: When a sound comes from one source, do you perceive two sounds or one?

- B2) Is there a difference between your ears in regard to how the height of tones is perceived?
- B3) Is there a difference between your ears in regard to how loudness is perceived?
- B4) Is there a device you use more to determine where a sound comes from?
- B5) Which device do you prefer in regard to sound quality?

Question B1 was phrased in a way that may not solely have addressed binaural integration. However, if participants were aware of differences in the percepts from one sound source across their ears, they must first have perceived the sound not as one uniform, binaurally integrated sound object.
In section C, listening abilities were assessed in a time-efficient manner using a questionnaire called the SSQ5, which was derived from the larger set of questions of the SSQ, short for Speech, Spatial and Qualities of Hearing Scale, but yielding comparable results with only five instead of fifty questions (Mertens and Heyning, 2013). Participants could answer the SSQ5 questions on scales from absolutely not to absolutely, with numbers from zero to ten.

Speech, Spatial and Qualities of Hearing Scale Questionnaire (SSQ5, section C, Mertens and Heyning, 2013):

- C1) Can you have a conversation with someone when another person is speaking, whose voice is the same pitch as the person you're talking to?
- C2) You are sitting in between two people. One of them starts to speak. Can you tell right away whether it is the person on your left or right, without having to look?
- C3) Can you tell how far away a bus or truck is, from the sound?
- C4) Do everyday sounds that you can hear easily seem clear to you (not blurred)?
- C5) Do you have to concentrate very much when listening to someone or something?

The questionnaire is available for download (https://doi.org/10.5281/zenodo.2574015).

2.4 Results

At the core of this survey is how question B1, aimed at whether participants integrate the interaural stimulation into one single sound object, relates to demographic factors from section A, specific perceptual aspects from section B, and the listening-performance ratings from section C. Hence, we will first assess the ratings from question B1 by itself, before looking at the relations to other questions.

The average answers regarding B1 did not change significantly in the various listening situations, as assessed by a Kruskal-Wallis one-way rank sum test not relying on normally-distributed data ($\chi^2 = 2.11$, df=5, $p = 0.834$; Hollander et al., 1973). Therefore, the results were averaged over these situations for an easier assessment. The difference in integration-ratings (B1) across the two centers, where the study was conducted, was not significant (Kruskal-Wallis one-way rank sum test: $\chi^2 = 2.5483$, df=1, $p = 0.110$). Still, the answers regarding the integration-rating (B1) varied widely, the data was separated for a better illustration of how the participants' answers to the other questions varied in conjunction with B1. To derive a sensible grouping, the participants integration ratings were analyzed using a cluster analysis. The cluster analysis used the Lance-Williams dissimilarity update-formula applying Ward's criterion, i.e., squaring dissimilarities before cluster updating (Murtagh and Legendre, 2014). Later data analysis was again performed without this illustrative grouping. The cluster analysis divided the participants into two main groups, as depicted in fig. 2.1. The first group reported no up to a moderate difference across ears in B1, while the second group rated the interaural difference as moderate-to-high, as can be seen comparing the interaural difference-ratings in the top panels of fig. 2.2 (group reporting better integration) and fig. 2.3 (group reporting worse integration). Three patients of the group reporting better integration reported to perceive exactly one uniform sound, and another three reported minimal differences up to one on the scale. Of the group reporting worse integration, two participants reported to hear two fully-separate sounds and another one reported an almost as high separation at 9.5 on the scale. The other participants reported difference-values in the intermediate region, either more towards one sound or more towards two sounds.

Information from the questions regarding the participants in section A can be found in tab. 2.1 - 2.3. Using Kendall rank correlations, we assessed relations between participants' demographics, devices, degree of hearing-loss and binaural integration. The threshold for a significant correlation is high ($p < 0.00217$), considering a criterium of $p = 0.05$, corrected for 23 comparisons in total, and none of the factors investigated showed a significant correlation to the raking for question B1 (cf. tab. 2): This includes age ($p = 0.326$), age at implantation ($p = 0.365$), time with hearing loss on the HA-aided ($p = 0.980$) and CI-aided sides ($p = 0.450$), time of experience with the HA ($p = 0.545$) and CI ($p = 0.705$), duration without aid on either side (HA: $p = 0.749$; CI: $p = 0.852$), the year participants received the HA ($p = 0.536$) and CI ($p = 0.608$), the brand of the HA ($p = 0.494$) and CI ($p = 0.0194$), the categorical degree of hearing-loss (mild, moderate, severe, profound; $p = 0.412$), and etiology ($p = 0.847$). Note that the p-value for the CI-manufacturer is below the standard criterion of 0.05, but not the significance-level corrected for the number of comparisons. The study design was also not balanced regarding the devices' manufacturers. For this reason, this factor was considered as non-significant (as it might be a false positive).

Section B contained several questions about specific perceptual aspects, namely pitch (B2) and loudness of sounds (B3), reliability for localization of sounds (B4) and preference for sound quality (B5). The results for these are shown in the mid-panels of fig. 2.2 and fig. 2.3, illustrating how these ratings varied for the two groups. For the group reporting better integration, the median ratings in all these measures were zero, indicating a balanced perception between HA and CI. Nevertheless, the answers varied largely, with some extreme ratings towards the CI being higher in pitch, louder, more reliable for localization, and better in sound quality. For participants in the group reporting worse integration, median ratings increased from zero towards the CI, indicating stronger interaural differences in these four aspects.

The relation between B1 and B2 to B5 was investigated for all participants together using Kendall rank correlations, corrected for multiple comparisons. The absolute value of the four difference measures was used, since it takes imbalance towards either device equally into account. The analysis showed that the rating for binaural-integration correlated significantly with the absolute ratings of pitch (B2; $p = 2.70 * 10^{-4}$), loudness (B3; p = $1.87 * 10^{-3}$), the preference for localizing sounds (B4; $p = 3.63 * 10^{-3}$), and sound quality (B5; $p = 3.30 * 10^{-4}$). Fig. 2.4 includes the scatter plots for B1 vs. B2 to B5, illustrating both the large variation and the correlations (cf. tab. 2.5).

The bottom panels in fig. 2.2 and fig. 2.3 show the self-assessed SSQ5 performance ratings in everyday listening-tasks in order: Speech in noise (C1), localization (C2), distance perception (C3), sound clarity (C4), and listening effort (C5). Performance in each group and for each question varied over a wide range. The median ratings for speech in noise, localization, sound clarity and listening effort were slightly better for the group that reported better binaural integration, while the rating for distance perception was not. But the differences in these performance-measures between the groups tending more towards integrated or separated sound percepts were not very pronounced. For speech reception in noise (C1), assessed in the SSQ5 with simply one talker in the background, most listeners rated their abilities only as moderate with exceptions towards either extreme. The ratings were slightly better for the group, which reported better binaurally-integrated percepts. The same can be concluded for sound localization (C2), even in the simple case of discriminating sounds from the left or right. Overall, the participants rated their abilities to judge the distance of sounds (C3) rather low-to-moderate, but a few participants reported both better integration and good ability to judge distances. Sound clarity (C4) also varied widely across listeners, with most reporting it to be moderate-to-good. Again, the participants who reported better integration also reported slightly better clarity. Most participants also described their listening effort (C5) to be rather high in general, with ratings showing a slightly lower effort for the group that reported better integration.

Relations between these performance-ratings (C1 to C5) and binaural integration (B1) were investigated using Kendall rank correlations (cf. tab. 2.6). The correlations for speech in noise (C1; $p = 0.0718$), localization (C2; $p = 0.695$), distance perception (C3; $p = 0.653$), sound clarity (C4; $p = 0.153$), and listening effort (C5; $p = 0.0478$) were all not significant, when accounting for the multiple comparisons, albeit the correlations for speech in noise and listening effort had relatively low p-values. Fig. 2.5 presents the scatter plots for these measures over the integration rating (B1).

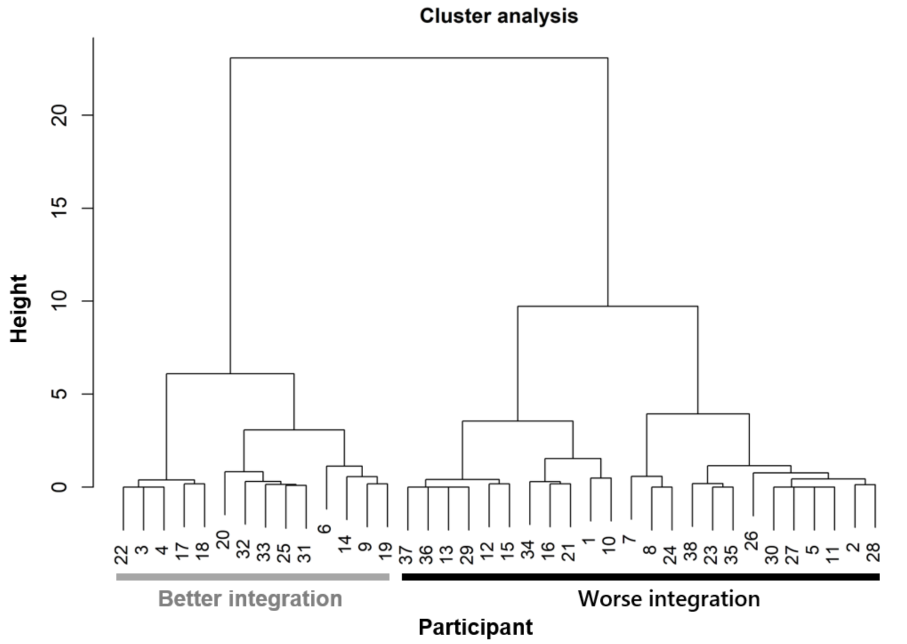


Figure 2.1: Cluster analysis results based on the perceived difference across ears (B1). Numbers at the bottom indicate individual participants. The group of participants to the left reported better binaural integration, while the others reported worse binaural integration (cf. fig. 2.2 - 2.3, top panel)

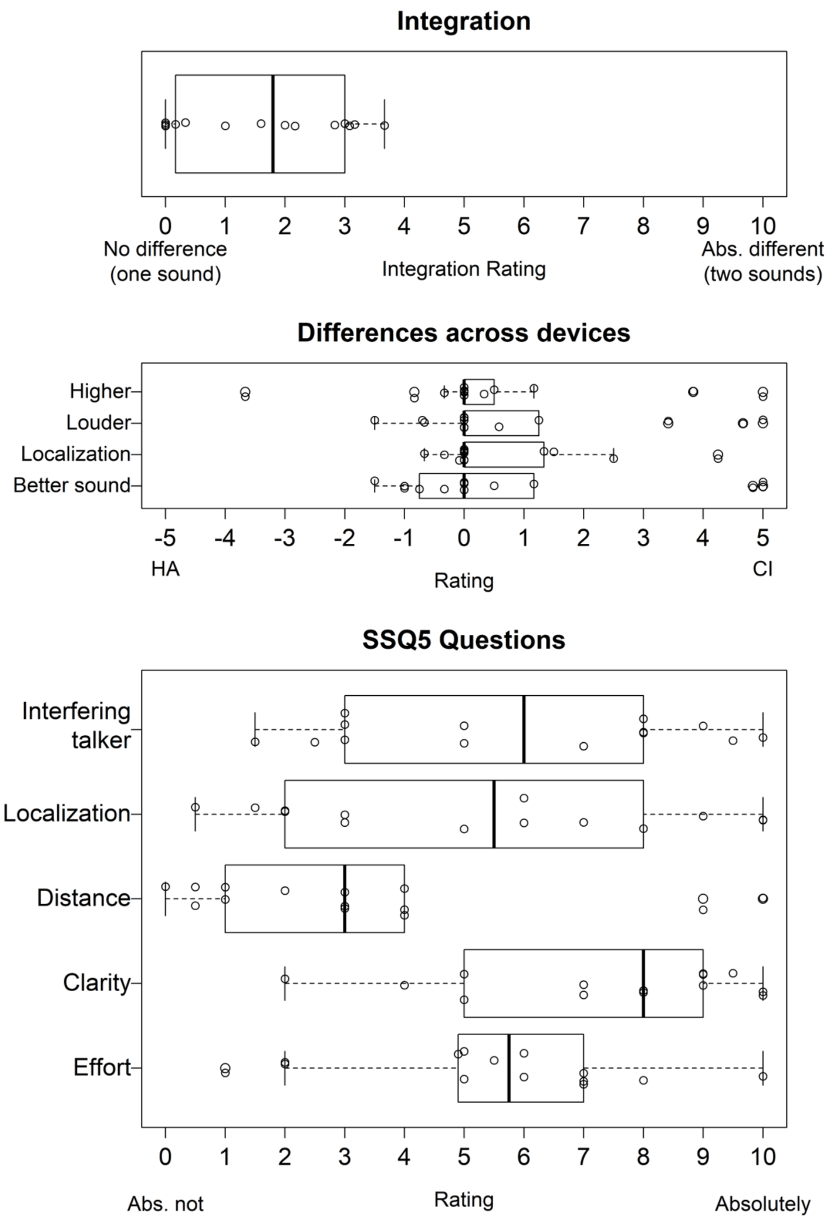


Figure 2.2: Results for participants who reported better binaural integration. Top panel: Binaural integration rating (B1). Middle panel: Ratings for differences in pitch (B2), loudness (B3), preferences for localization (B4), and sound quality (B5) across devices. Bottom panel: Listening performance ratings (C1-C5). Circles indicate individual ratings.

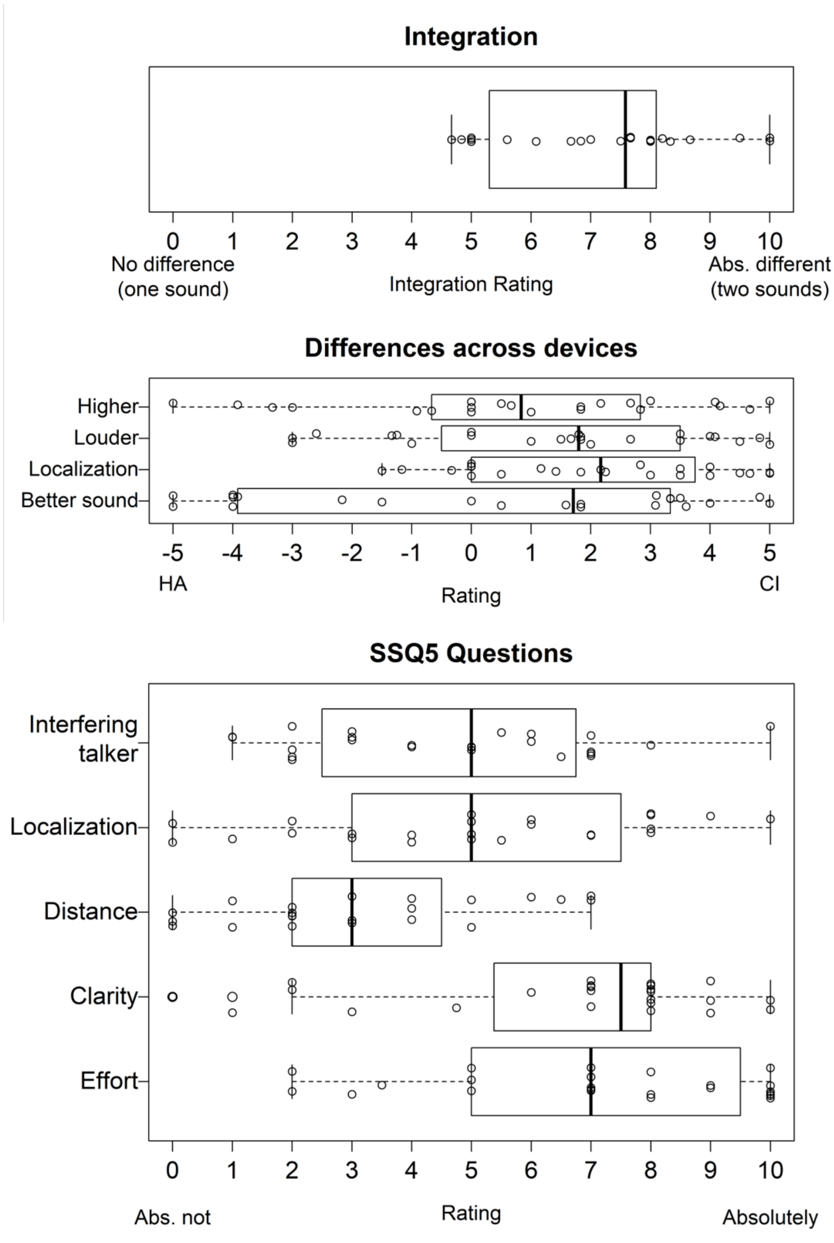


Figure 2.3: Results for participants who reported worse binaural integration. Top panel: Binaural integration rating (B1). Middle panel: Ratings for differences in pitch (B2), loudness (B3), preferences for localization (B4), and sound quality (B5) across devices. Bottom panel: Listening performance ratings (C1-C5). Circles indicate individual ratings.

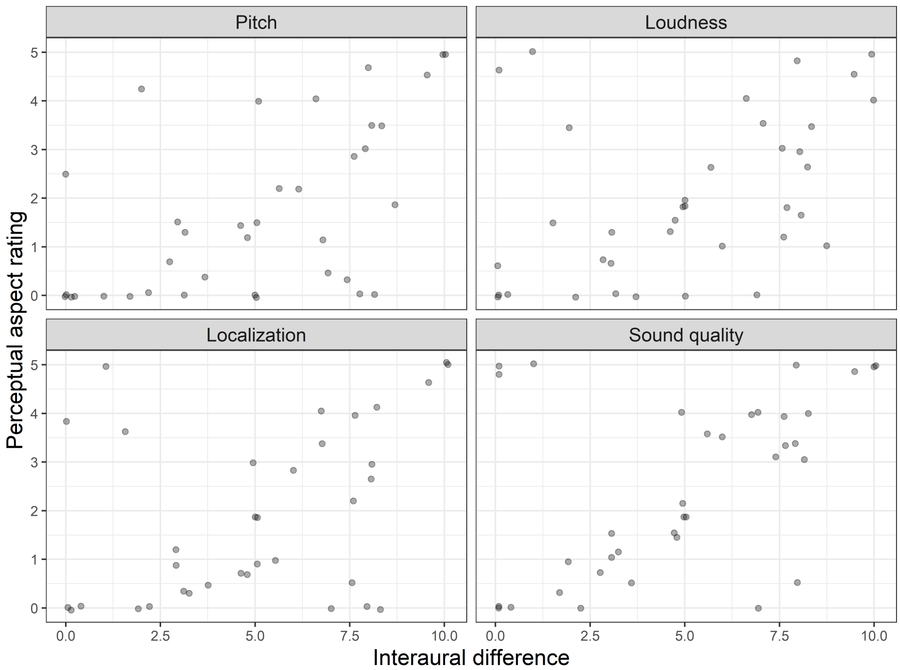


Figure 2.4: Scatter plots of the absolute ratings for the perceptual aspects, by row: pitch (B2), loudness (B3), localization-preference (B4), and sound quality (B5), plotted over the binaural integration rating (B1).

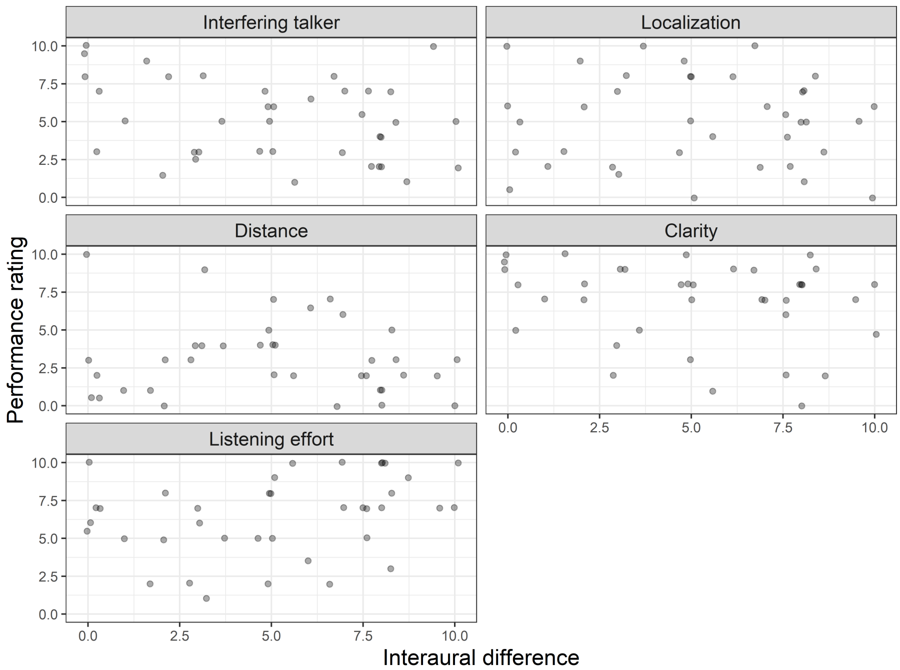


Figure 2.5: Scatter plots of the listening-performance ratings of section C, by row: interfering talker (C1), sound localization (C2), distance perception (C3), sound clarity (C4), and listening effort (C5), plotted over the binaural integration rating (B1).

Table 2.1: Additional information on the bimodal participants' gender, etiology, age in years (a), the ear aided by the CI, their grade of hearing loss and the site they were recruited at (DTU: Technical Univerisy of Denmark; MHH: Medical University of Hannover)

ID	Gender	Etiology	Age / a	CI Ear	Hearing loss	Site
1	F	Hereditary	35.2	R	moderate	DTU
2	F	Hereditary	58.6	R	severe	DTU
3	F	Meniere	67.2	L	severe	DTU
4	F	Meningitis	35.2	R	severe	DTU
5	M	Blood clot	58.6	R	moderate	DTU
6	F	Meniere	67.2	R	moderate	DTU
7	M	Hereditary	35.2	R	severe	DTU
8	F	Middle ear infection	58.6	R	severe	DTU
9	M	Presbycusis	67.2	L	Unknown	DTU
10	F	Unknown	35.2	L	Unknown	DTU
11	M	Hereditary	58.6	R	Unknown	DTU
12	M	Sudden hearing loss	67.2	L	severe	MHH
13	F	Meningitis	35.2	L	profound	MHH
14	M	Unknown	58.6	L	moderate	MHH
15	F	Unknown	67.2	L	profound	MHH
16	F	Sudden hearing loss	35.2	L	profound	MHH
17	F	Sudden hearing loudness	58.6	R	moderate	MHH
18	F	Unknown	67.2	R	profound	MHH
19	M	Hereditary	35.2	L	moderate	MHH
20	F	Unknown	58.6	R	profound	DTU
21	F	Sudden hearing loss	67.2	R	moderate	MHH
22	M	Noise	35.2	L	high-freq	MHH
23	M	Unknown	58.6	L	high-freq	MHH
24	F	Noise	67.2	R	profound	MHH
25	M	Otosclerosis	35.2	R	profound	MHH
26	M	Unknown	58.6	R	Unknown	MHH
27	M	Hereditary	67.2	R	moderate	MHH
28	F	Presbycusis	35.2	R	Unknown	MHH
29	F	Presbycusis	58.6	L	Unknown	DTU
30	F	Hereditary	67.2	R	Unknown	DTU
31	F	Presbycusis	58.6	L	moderate	MHH
32	M	Apoplexy	67.2	L	moderate	MHH
33	M	Hereditary	39.8	L	moderate	MHH
34	F	Sudden hearing loss	58.6	R	moderate	MHH
35	F	Sudden hearing loss	67.2	R	moderate	MHH
36	F	Presbycusis	39.8	L	profound	MHH
37	F	Presbycusis	58.6	L	moderate	MHH
38	F	Meniere	67.2	R	severe	MHH

Table 2.2: Additional information on the participants' hearing devices

	HA		CI	
ID	Brand	Type	Brand	Type
1	Phonak	Naida S V UP	Advanced Bionics	Naida SP Q70
2	Phonak	Solana micro P	Cochlear	Unknown
3	Phonak	Naida S V SP	Advanced Bionics	Unknown
4	Phonak	Naida S V SP	Med-EL	Flex 24, Opus 2
5	Phonak	Naida S V UP	Cochlear	Nucleus, CP910
6	Oticon	Chili SP5	Cochlear	Nucleus, CP800
7	Oticon	Alta Pro	Cochlear	Nucleus, CP910
8	Phonak	Naida Q70	Cochlear	CP810
9	Phonak	Naida S V SP	Advanced Bionics	Naida SP Q70
10	Phonak	Naida S V SP	Advanced Bionics	Naida
11	Oticon	Dynamo 10	Oticon Medical	Neuro One
12	Phonak	Naida Q70	Advanced Bionics	Naida Q90
13	Siemens	Unknown	Cochlear	CI532, CP920
14	Siemens	Pure	Cochlear	CI532, CP920
15	Phonak	Audeo	Med-El	Mi1200 Synchrony
16	Siemens	Unknown	Cochlear	CI532, CP910
17	Phonak	Naida Link	Advanced Bionics	Naida Q70
18	Oticon	Hit	Advanced Bionics	Unknown
19	Phonak	Naida Link	Advanced Bionics	Naida Q70
20	Phonak	Unknown	Cochlear	CI532, CP910
21	Siemens	Unknown	Med-El	Unknown
22	Widex	ME-G	Med-El	Concerto Flex 28
23	Phonak	Milo micro	Cochlear	CI24RE, CP910
24	Phonak	Unknown	Med-El	Unknown
25	Phonak	Naida SP Q90	Advanced Bionics	Naida Q90
26	Phonak	Unknown	Oticon Medical	Neuro One
27	Phonak	Naida SP Q70	Advanced Bionics	Naida Q70
28	Phonak	Unknown	Med-El	Unknown
29	GN Resound	Unknown	Cochlear	CP 910
30	Oticon	Dynamo SP10	Oticon Medical	Neuro One
31	GN Resound	Unknown	Cochlear	CI24RE (CA), CP910
32	Phonak	Naida SV SP	Cochlear	CI512, CP910
33	GN Resound	Unknown	Cochlear	CI522, CP910
34	Unknown	Unknown	Cochlear	CI522, CP910
35	Unknown	Unknown	Cochlear	CI522, CP910
36	Unknown	Unknown	Cochlear	CI24RE (CA), CP910
37	Unknown	Unknown	Cochlear	CP910

Table 2.3: Additional information on the participants on durations for their hearing loss and periods of deafness, as well as experience, given in years (a)

ID	Time with hearing loss / a HA	CI	Time with deafness / a CI-aided ear / a	Experience with / a HA	CI
1	35.2	32.6	0.0	31.2	2.6
2	29.6	29.3	0.0	24.6	0.5
3	22.2	26.2	0.0	22.2	2.0
4	22.3	20.1	0.0	22.3	2.3
5	25.9	23.5	15.5	25.9	2.4
6	27.8	22.0	0.0	23.8	5.8
7	25.3	23.0	0.0	25.3	2.3
8	39.4	37.6	0.0	39.4	3.8
9	63.6	62.9	0.0	15.6	0.7
10	60.8	59.8	0.0	2.1	1.1
11	19.8	29.6	0.0	19.8	0.5
12	7.4	0.6	0.0	6.8	6.8
13	52.5	51.9	5.0	47.5	0.6
14	20.8	15.3	1.6	20.5	0.5
15	5.0	4.5	0.9	1.3	0.5
16	8.4	7.3	0.0	11.4	1.2
17	36.0	30.0	0.0	10.0	6.0
18	30.8	30.3	1.0	30.8	0.5
19	46.5	45.8	0.0	6.5	0.7
20	30.5	28.3	0.0	30.4	2.2
21	40.8	37.2	0.0	14.8	3.7
22	16.3	12.9	0.0	16.3	3.4
23	8.8	29.8	0.0	8.7	1.0
24	25.3	0.1	0.0	25.3	25.3
25	22.3	0.1	0.0	22.3	22.3
26	21.3	32.3	30.0	21.2	1.0
27	57.4	53.9	0.0	1.0	3.5
28	31.8	26.4	0.0	6.8	5.3
29	21.1	19.0	0.0	21.0	2.1
30	48.9	48.1	0.0	48.9	0.8
31	40.8	0.0	2.3	40.8	6.8
32	46.7	23.0	0.0	22.7	4.7
33	56.7	54.1	0.0	12.7	2.6
34	12.7	10.3	0.0	10.3	2.3
35	12.7	10.3	0.0	10.3	2.3
36	38.3	25.7	0.0	32.7	12.6
37	21.4	19.0	0.0	21.2	2.4
38	61.5	53.0	0.0	49.7	8.5

Table 2.4: Kendall rank correlation results between the integration rating (B1) and various demographic factors, devices, and hearing loss from section A

Aspect	τ	Z	p
Age	-0.112	-0.982	0.326
Age at implantation	-0.104	-0.980	0.365
Time with hearing loss on HA side	-0.0029	-0.0252	0.980
Time with hearing loss on CI side	0.0866	0.756	0.4497
Time of experience with HA	0.0692	0.605	0.545
Time of experience with CI	0.0435	0.378	0.705
Duration without aid on HA side	-0.0439	-0.320	0.749
Duration without aid on CI side	-0.0240	-0.187	0.852
Year (first) HA was obtained	-0.0717	-0.618	0.536
Year of implantation	-0.0620	-0.513	0.608
HA manufacturer	-0.0915	-0.684	0.494
CI manufacturer	0.299	2.34	0.0194
Degree of hearing-loss (categorical)	0.117	0.820	0.412
Etiology	0.0296	0.193	0.847

Table 2.5: Kendall rank correlation results between the integration rating (B1) and absolute ratings for difference in pitch (B2), loudness, localization preference and preference for sound quality

Aspect	τ	Z	p
(B2) Pitch	0.436	3.64	$2.70*10^{-4}$
(B3) Loudness	0.368	3.11	$1.87*10^{-4}$
(B4) Localization preference	0.356	2.91	$3.63*10^{-3}$
(B5) Pref. for sound quality	0.432	3.59	$3.30*10^{-4}$

Table 2.6: Kendall rank correlation results between the integration rating (B1) and listening-performance ratings from section C

Question	τ	Z	p
(C1) Interfering talker	-0.212	-1.80	0.0718
(C2) Localization	-0.0462	-0.393	0.695
(C3) Distance perception	-0.0544	-0.450	0.653
(C4) Clarity	-0.172	-1.43	0.153
(C5) Listening effort	0.237	1.98	0.0478

2.5 Discussion

Assessing if bimodal CI patients experience sounds from one sound source as integrated into one sound object and whether this reflects in their self-reported percepts and listening performance was the inspiration for this study. The survey gave the participants the option to report their perception originating from the same source in a non-binary way (question B1) and most participants made use of this and answered in between the extremes of one and two sound percepts. This indicates a less than straightforward devision between either full integration of all sounds or the opposite. The subjective nature of the survey may have prevented that all participants interpreted the question in the same manner, so that their answers may also reflect other perceptual aspects than solely binaural integration. However, if their answers reflect such other perceptual differences, this again demonstrates that they did not (fully) integrate the representations of one sound source into one uniform sound object or stream, like normal-hearing listeners would, as they necessitate separated percepts.

Only three out of thirty-eight bimodal CI participants reported to perceive sounds originating from one sound source as integrated into one uniform sound object with no difference at all. Twenty-five participants reported to perceive sound from a single source with varying degrees of differences across ears and two patients reported even the extreme of two entirely separate sound percepts. This could be interpreted such, that binaural integration is not just either present or absent in most bimodal CI patients, but that the processes of integration might be affected by an adaptation to the bimodal stimulation these patients experience daily, comparable to findings for adaptive pitch percepts in Reiss et al. (2014b) and changes in binaural processing due to interaural asymmetries in Polonenko et al. (2019). Therefore, the present results suggest most bimodally-aided patients integrate sounds differently compared to normal-hearing listeners or not at all.

The participants were asked to answer the questions B1 to B5 in six everyday listening situations. Since every participant encounters these situations in a different form in their life, their answers may not necessarily be directly comparable across individuals, but nevertheless should give an impression as to how their percepts change with associated kind of acoustic background noise. Their ratings regarding binaural integration (B1) changed not significantly over the six situations, whether in quiet (S1) or with various background noises (S2 – S6).

This indicates that the ability to integrate may not depend on the listening environment, but (at least foremost) other factors. The participants were familiar with these listening situations from their life and instructed to not answer in case they felt uncertain about a described situation. However, they were not notified of the listening situations before the survey was conducted and, hence, did not have time to pay special attention to them. Such preparation could perhaps have resulted in more varied answers over the conditions.

It is possible for the large number of participants with integration ratings in the intermediate region, that they experience some form of binaural integration. Perhaps, this comes with greater effort and over a wider range than expected for normal-hearing listeners, just as Reiss et al. (2014a) reported from a psychoacoustic experiment. Many natural stimuli, such as the speech sounds addressed specifically in question B1, extend over many seconds and evoke auditory streaming in normal-hearing listeners (Shinn-Cunningham et al., 2017). It is possible, that the binaural integration of such streamed stimuli differs from that of short dichotic stimuli, offering an explanation for why most patients in the current study rated the binaural integration intermediate to low instead of high, as could be expected from the wider binaural fusion observed for bimodal CI patients in Reiss et al. (2014a) and similar results obtained for bilateral CI patients in Reiss et al. (2017). Another explanation could be that the participants with intermediate ratings only integrate over a limited spectral range covered by both devices. HAs can cover a lower spectral region than CIs and the residual hearing of many bimodal CI patients may in turn restrict stimulation from higher frequency sounds to the CI alone. Sounds in the overlapping frequency region could be integrated. The intermediate integration ratings of the participants could, thus, represent a spectral average. Some participants indicated this along with their answers, describing to perceive lower frequency sounds ("bass") only from the HA and higher frequency sounds only from the CI. Investigating binaural integration in bimodal CI patients further with listening experiments that require them to build streamed percepts out of binaural stimulation may add to understanding the potential differences in binaural integration.

Studies have shown relations between the history of hearing loss (Polonenko et al., 2019), the frequency range of residual hearing (Sheffield et al., 2016), and binaural processing or bimodal benefit. Despite that, it was not possible to predict whether participants integrate based on the various demographic factors, devices, hearing loss, experience with the devices, or etiology, that were assessed in section A of the survey, as none of these factors showed a significant correlation to question B1 after correction for multiple comparisons. Unfortunately, it was not possible to identify the exact HA model used by many participants, so that no analysis of the effect of HA model could be conducted.

Our first hypothesis was that imbalances in percepts across ears could lead to a lack of binaural integration. The results appear to confirm this: With lack of binaural integration (B1), bimodal patients also reported differences in pitch (B2) and loudness (B3) across ears, higher reliability of one of their devices, mostly the CI, for localizing sounds (B4), and, likewise, a preference for that device in sound quality (B5). In turn, a balanced perception of pitch and loudness (B2 and B3) made it more likely that the patients reported better integrated percepts (B1) and made them less inclined to judge CI or HA as more reliable for sound localization (B4) or better in sound quality (B5).

The significant correlation of both pitch (B2) and loudness (B3) to the rating for binaural integration (B1) suggests that both percepts should be balanced across ears as an optimal basis for binaural integration, as suggested in Gordon et al. (2012) and Gordon et al. (2017). This could for example include a customization of the frequency-to-electrode map of the CI according to pitch-matches in the non-implanted ear and loudness-balancing of stimuli in the resulting filterbands across ears. This could also reflect in better overall outcome-measures (Kan et al., 2013; Steel et al., 2015). With this, the results also agree with the findings of Suneel et al. (2017), who demonstrated that interaural mismatches in percepts lead to a lack of binaural fusion and reflect in a reduced localization performance.

If the survey had been conducted among bilateral HA patients, the similarities in the sound's representation should have led to a balanced pitch perception across ears. The loudness percepts would depend on the residual hearing of the patient, whether the same make and model of device was used across ears, since different models can feature different automated gain control systems and output levels, and the fitting of the devices' parameters (Kates, 2008). With balanced and similar representations of sounds across ears it is to be expected, that bilateral HA patients would binaurally integrated percepts.

For bilateral CI patients, despite the same or similar devices would be used across ears, the results could show mismatches in pitch percepts, since clinical fitting procedures typically do not include pitch matching of electrodes, but use

a standard frequency-to-electrode mapping, which does not consider, e.g., the insertion depth of the electrode array and resulting differences in percepts (Gordon et al., 2017; Oxenham, 2018). The loudness percepts would be balanced for these listeners if a bilateral loudness balancing was conducted with the device fitting. The possible differences in pitch percepts across ears could then lead to a lack of binaural integration for this group of patients (Gordon et al., 2012; Gordon et al., 2017; Moore and Gockel, 2002; Suneel et al., 2017).
Likewise, significant correlations to the ratings for sound localization (B4) and sound quality (B5) indicate that both are affected when patients do not integrate the percepts from both ears, resulting in a preference towards one of the devices, usually the one that was also perceived as louder. For most of the participants, the preferred device in these cases was the CI, as visible in the distribution of answers (cf. fig. 2.3, middle panel). Since localization in the healthy auditory system is based on binaural cues, the reported preference for one device with a reported lack of normal binaural integration could be a result of the fact that the CI was often perceived as louder. Hence the interaural loudness difference would make it more likely that patients localize sounds towards the the side with the device perceived louder. In line with previous results linking interaural mismatches to a lack of binaural fusion, changes in binaural processing, and decreases in sound localization performance (Easwar et al., 2018; Gordon et al., 2012; Kan et al., 2013; Polonenko et al., 2019; Suneel et al., 2017), these results may also indicate changes in binaural processing, strengthening the indicated lack of binaural integration. For patients reporting a lack of loudness balance across devices even after optimization of the devices' fittings, implantation of a second CI could perhaps lead to more balanced percepts and offer a better basis for binaural integration.
Despite these trends towards better ratings for speech reception in noise (C1), sound localization (C2), sound clarity (C4), and listening effort (C5) for participants who reported better binaural integration of sounds in B1, no significant correlations between these ratings were found. Hence, our second hypothesis regarding a connection between binaural integration and better performance in everyday listening-tasks could not be validated. The lower p-values for speech reception in noise and listening effort suggest that a much larger group of participants could lead to significant results. This was, however, not foreseeable in the study design, since it was difficult to estimate effect sizes for this explorative study. One might speculate that the abilities to understand speech in noise,

localize sounds, as well as sound clarity and listening effort depend not only on a binaurally-integrated percept of sounds from the same source, but also the ability to utilize binaural time-difference and loudness-difference cues, spectral differences and fine-structure information. The ability to utilize these cues can be hindered by various aspects of the devices' processing. Differences in compression systems and lack of synchronized automated-gain-control systems can severely distort binaural cues, leading to reduced performance in complex listening environments (Easwar et al., 2018; Fitzgerald et al., 2015; Litovsky and Gordon, 2016). Therefore, patients who reported integrated percepts might not necessarily also show better performance with their clinical devices, unless further efforts are undertaken to represent binaural cues accurately. This could explain, why the participants' answers varied largely, and no significant correlations between binaural integration (B1) and the performance-measures (C1 to C5) were found.

A pilot study with normal-hearing listeners could perhaps have helped assess potential shortcomings or misunderstandings. The fact that patients were interviewed at two different sites and in two different languages may further have contributed to some variation, although the translations have been conducted by native speakers and the personal interviews may have helped to ensure that participants understood the questions in the same way. An alternative approach would have been to design the study around an online questionnaire, which could allow to reach a greater number of listeners, albeit increasing the risks for misunderstandings, lack of responses, and, perhaps, a bias regarding demographic factors.

The addition of listening performance data would provide a means to verify the subjective performance ratings of the patients in section C and could perhaps have established stronger links between binaural integration and listening performance. However, this would have considerably impacted the duration of the assessment.

2.6 Conclusion

The results from this survey suggest that only a minority of bimodal CI patients integrate sounds binaurally like normal-hearing listeners do, when it comes to everyday sounds such as speech, independent of the listening environment.

Instead, most of the participants reported to perceive sounds from the same source not as a single, uniform sound object, but as something more complex. This could be a sign that they are aware of a form of abnormal binaural integration, such as binaural fusion over a wide pitch range, or that they only integrate in a limited spectral range, covered by both HA and CI. The integration rating could not be predicted by experience with the devices, device manufacturer, hearing loss, or etiology. Participants who did not report binaural integration reported pronounced differences in pitch and loudness, tended to rely more on one of the devices, mostly the CI, for localizing sounds, and preferred that device for sound quality. Patients who reported better integrated percepts also reported a better balance for pitch and loudness across devices, were less inclined to rate one device as more reliable for sound localization, or better in sound quality. As a takeaway, aligning the interaural percepts for bimodal CI patients could aid binaural hearing though better integration of the sounds from HA and CI and may lead to better speech reception in noise, better localization abilities, and reduce listening effort. Steps to better aligned percepts could be to customize the frequency-to-electrode map of the CI according to pitch-matches in the non-implanted ear with subsequent loudness-balancing of stimuli in the resulting filter-bands across ears.

3

A modification of the scale illusion into a detection task for assessment of binaural streaming[a]

Abstract

Binaural streaming by frequency-proximity was investigated without subjective listener-feedback by modifying the scale illusion of Deutsch (1975) into a detection task. Nineteen normal-hearing listeners had to detect one deviation within a repeating melody-stream, while simultaneously presented with a randomized distractor stream. Every second note in each stream was presented to the opposite ear, requiring binaural streaming to detect the deviant. Listeners performed well in this test but adding interaural delay or timbre-difference let the listeners group by location instead. This confirms the grouping by frequency-proximity. The method could be used to investigate binaural streaming in hearing-impaired patients, where interaural percepts might differ.

3.1 Foreword

As outlined in the general introduction to this thesis, auditory scene analysis stands at the core of how the human brain evaluates the acoustic environment, using perceptual similarities to form sounds objects from the manifold sound components through grouping and streaming as described in Bregman (1990) and Shinn-Cunningham et al. (2017). A main goal of these Ph.D. studies was to investigate if interaural differences in percepts, caused by the use of CIs for bilaterally or bimodally aided patients, lead to changes in binaural streaming.

[a] This chapter is based on Janssen et al. (2019b).

But studying this without relying on the patients' subjective descriptions required the development of a suitable experiment. This chapter describes the new method developed to assess binaural streaming and its test and verification with normal-hearing listeners.

3.2 Introduction

In 1975, Diana Deutsch aimed to investigate streaming of dichotically presented melodies and described an auditory percept called the scale illusion (Deutsch, 1975). The illusion was composed of two melody-streams at different frequency ranges, one high and one low, out of the components of one musical scale. Both streams ascend and descend in frequency over eight notes. The streams were presented in such a way that every second note from each of the two streams was played to the other ear, in a pattern symmetrical to the middle of the repeating sequence of eight notes. So, when one component of the higher melody was in one ear, a component of the lower melody was in the other ear, and vice versa. The intriguing effect of this stimulation was, that listeners did not group the stimuli by ear of input, but instead by frequency range. Most right-handed listeners perceived the higher stream coming from the right and the lower stream coming from the left, whereas only about half of the left-handed listeners did report this percept ("both streams"). A minority of the right-handed listeners and the other half of the left-handed listeners reported to perceive only the higher stream with four tones going up and down, and little or nothing of the lower stream ("single stream"). Deutsch argued that this illusion outlined binaural streaming in which the Gestalt principle of pitch-proximity overrides the lateralization cues. The illusion is only possible through binaural processing of the stimuli from both ears.
The participants in the original study had to report verbally what they perceived. Describing the complex percept with two melodies precisely might be difficult for listeners without musical training and, likewise, the researcher might find it challenging to interpret the listener's descriptions correctly. This study aims to validate a method in which the ability to perceive the scale illusion will be assessed through a detection task. Once validated this method could be performed by listeners with asymmetrical hearing loss and fitted with different hearing devices such as one hearing aid and one cochlear implant, to evaluate

their ability to integrate signals across ears and devices.

3.3 Participants

Nineteen normal-hearing (NH) listeners participated in the experiment (aged 24 to 31; means 26.1 years, hearing-loss of less than 20 dB). Seven of the listeners were female, eleven musically trained, and three left-handed. They were all recruited at the Technical University of Denmark and provided written informed consent. All experiments were approved by the Science-Ethics Committee for the Capital Region of Denmark (reference H-16036391).

3.4 Method

3.4.1 Stimuli

Participants were presented with sequences of notes that formed two melodies. Two types of stimuli were used to generate these melodies: pure tones, and complex tones designed to vary on two distinct dimensions of timbre, impulsiveness and brightness, as well as loudness. The interval between notes was 250 ms, as in the study by Deutsch (1975). All the pure-tone stimuli were loudness-balanced for each frequency and presented in a double-walled sound-isolated listening booth via HDA 200 headphones (Sennheiser GmbH & Co. KG, Wedemark, Germany), driven by a Scarlett 2i2 sound card (Focusrite Audio Engineering Ltd., High Wycombe, United Kingdom).

The pure tones had a duration of 250 ms with 10 ms half Hann-window ramps on the onset and offset. Their frequencies ranged from 262 to 523 Hz and were set according to the twelve-tone equal-temperament tuning system. The complex tones were five-tone-harmonic complexes with a duration of 200 ms including a 50 ms half Hann-window onset and offset. Their fundamental frequencies (F0) were set to equal those of the pure-tones, and to this F0, the first four harmonics were added. All complex tones' components were presented 6 dB below the level of the related pure tone.

All the pure-tone stimuli were loudness-balanced using an adjustment-procedure based on a reference sound: a pure tone at 508 Hz in the right ear, a frequency different from that of the pure-tone stimuli to be balanced. In the beginning, participants could adjust the reference sound to their preferred, most-comfortable loudness level. After that, they were presented with the reference sound followed by one of the pure tones in each trial. Their task was to adjust the loudness of the latter so that it matched that of the reference via an on-screen slider. The slider provided an adjustment range from 40 to 90 dB SPL. Participants could repeat the stimuli and readjust the loudness until they were satisfied. This balancing was performed twice for all frequencies used in the experiment in random order. Overall, the participants chose levels that ranged from 56.3 to 85.3 dB SPL with a mean of 67.7 dB SPL. The headphones were calibrated using a Brüel & Kjær (B&K; Nærum, Denmark) 2636 sound-level meter, with IEC-60318-1 ear simulator B&K 4153 and reference calibrator B&K 4230 and equalized by their frequency response.

3.4.2 Procedure

To test whether participants stream by frequency range instead of ear of input, this study used a detection task paradigm, based on the scale illusion experiment by Deutsch (1975). As in the original study, two melodies were presented concurrently. These will be referred to as the target stream and the distractor stream. The target stream was designed to be followed by the listener. It was chosen to be identical to one of the eight-note melodies in the scale illusion, either of the higher or lower frequency range. The distractor stream consists of eight notes with frequencies chosen randomly from the other frequency range, i.e., if the target stream was of the higher frequency range, the distractor stream was of the lower frequency range, and vice versa (see fig. 3.1). As in the experiment by Deutsch (1975), two notes were always presented simultaneously, one from each stream, except in a control condition with added delay (see Conditions below). If a note from the target stream was on the right, a note from the distractor stream was on the left, and vice versa. Consecutive notes for both streams were presented to the opposite ear, in a pattern that reversed in order after every four notes. Thus, the arrangement formed a pattern symmetric to the middle of the repeating sequence of eight notes (e.g., for one stream in one presentation of the eight notes: left-right-left-right-right-left-right-left).

If the listeners group by frequency-proximity, they should be able to follow the target-stream in one ear, while the distractor stream is either perceived

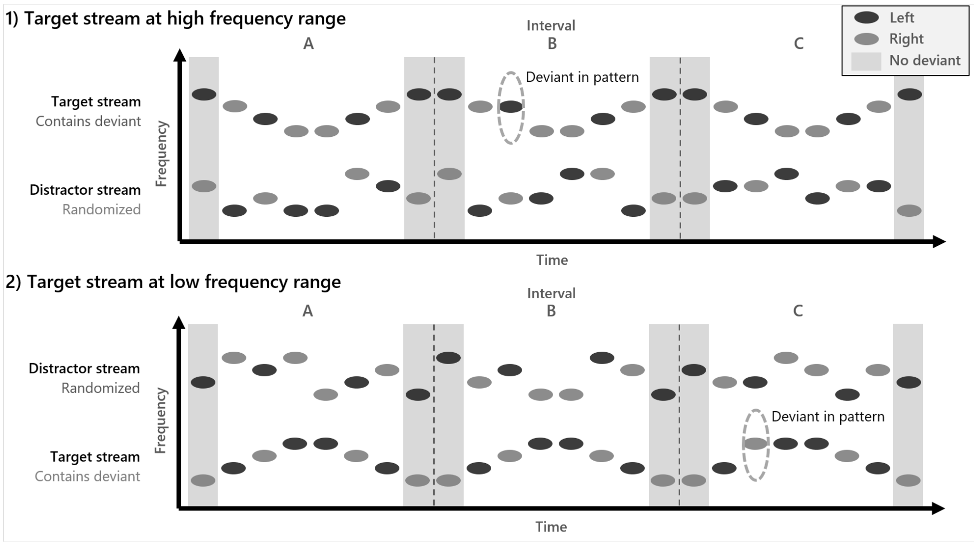


Figure 3.1: Schematic stimuli representation, based on Deutsch (1975). A random deviant occurs in the target stream, either at the higher (top) or lower (bottom) frequency range. The distractor stream consists of notes picked randomly from the other frequency range. The black and grey color of the ovals indicates the ear of presentation in the binaural test condition.

on the other side or not at all, as described by Deutsch (1975). Conversely, if the listeners were grouping based on lateralization cues, they would instead perceive random melodies on both sides, since the target notes and the random distractors from each side will be grouped. To test whether the listener could follow the target stream, a deviant note was introduced into it. The deviant consisted of shifting the note's respective frequency to one adjacent in the regular pattern. A listener who groups the binaural input by frequency-proximity was expected to detect this deviant easily, whereas somebody who groups by ear of input should experience difficulties to reliably detect the deviant note due to the unpredictable input in both ears.

At the beginning of each trial, the participant was presented with the target-stream alone twice, indicating which melody to listen for. After that, target and distractor streams were presented simultaneously six times. Of these six repetitions, the first three were intended to allow for a build-up of streaming, since it has been reported that a segregated percept arises gradually over several seconds (Bregman, 1990). One deviant note was introduced randomly in one of the last three repetitions, as depicted exemplarily in interval B at the top of fig. 3.1. The interval being presented was indicated on a graphical user interface by lighting up the respective part of the stimulus on the screen. It was possible

for the listener to repeat the stimuli.
The deviant could occur at any of the inner six of the eight notes within a presentation of the melody, but not at the first or last position. Its occurrence was counterbalanced with respect to the side of the presentation, interval, and frequency range of the target stream. Using a three-alternative, forced-choice paradigm prevented a detection-bias, since the participants knew that a deviant was always present in one of the three intervals as described in Wickens (2002), therefore allowing for an "objective" assessment of their grouping behavior.
Before conducting the test, participants underwent training. This training was only different from the experiment, in that the participants were given feedback on whether they answered correctly.

3.4.3 Conditions

The experiment featured four conditions, which are described below. Each of the four conditions consisted of twenty-four trials and used the pure-tone stimuli, except for a binaural control with changed timbre. The order of all trials from all conditions was randomized per participant to prevent it from affecting results.
In the binaural test condition, both the target and the distractor stream were presented in the same fashion as the original scale illusion, i.e., every second note from each stream was presented to the other ear. This condition, thus, required grouping by frequency-proximity to perceive (i.e., segregate) the target stream and detect the deviant.
In the monaural test condition, the target and the distractor streams were presented to the same ear. In this configuration, the listeners should easily group the target notes into one stream and the distractor notes in another one and, as a consequence, accurately detect the deviant. This condition was designed to verify that the listener could detect the deviant once the target notes were grouped into one stream.
As the ability to detect the deviant is used as a proxy to evaluate whether the listener can experience the scale illusion, it is necessary to add a control condition in which grouping the target notes across ears will be prevented by enhancing the lateralization cues to induce grouping by ear-of-input. In this configuration, the listener should perceive two random melodies in each ear, and should not be able to detect the deviant.

Two binaural control conditions were tested, in which the lateralization cues were enhanced through two different methods. In the first control condition, the binaural delay control condition, a delay of 125 ms (half a note's duration and larger than the range for natural interaural time-difference cues) was added to the sounds in one ear. As demonstrated by Deutsch (1979), introducing an asynchrony increases the tendency to treat each ear-input as emanating from a different source and will, therefore, induce a grouping by ear-of-input. Only eight of the nineteen participants were tested in this condition.
In the second binaural control condition, a significant timbre and loudness difference was introduced between the notes presented in each ear. Pure tones were presented to one ear as in the other conditions, whereas complex tones (cf. Stimuli) were presented to the other. Adding a salient perceptual difference across ears should lead the listener to group the notes by ear-of-input and result in a low detection performance. This condition will also help to quantify the possible performance when attending to just one ear.

3.5 Results

The results are presented in fig. 3.2. The detection performance is plotted as percent of correct answers for the four conditions: binaural and monaural test conditions, followed by the two binaural control conditions. For each condition, there are two data points: the average performance for the target-stream in the low frequency range ("low"), and the average performance for the target-stream in the high frequency range ("high"). The average performance for both the binaural and monaural test conditions lies at about 80 % correct, and the performance in the high frequency range is slightly higher than in the low frequency range. The average performance in the binaural control condition with delay is much lower at about 54 % correct, while that in the control condition with altered timbre lies at around 50 % correct. In both controls, again the performance in the high frequency range is slightly higher. The trend for high performance in both test conditions and lower performance in both controls was also found for each listener individually. A two-way repeated analysis of variance was performed in the software JMP® (SAS, Cary, NC, USA) with the "rationalized" arcsine transform (rau) applied (Studebaker, 1985), detection scores as the dependent variable, and the conditions, as well as the target stream's frequency range as factors. Both the condition ($F(3, 43.9) = 37.0$, $p < 0.0001$), as well as the frequency range ($F(1, 19.6) = 5.76$, $p = 0.0264$) were significant factors, while their interaction was non-significant ($F(3, 44.29) = 1.11$, $p = 0.355$). The estimated degrees-of-freedom are decimals because the number of participants tested in the delay condition is lower than in the other conditions. Posthoc analysis using the Bonferroni correction and again rau-transformed data, showed that results for both the binaural and monaural test conditions were not significantly different from each other ($p = 0.118$). The results in the binaural test condition were significantly different from the binaural controls with delay ($p = 0.0138$) and timbre ($p = 0.0012$), while results in these two binaural controls were not significantly different from each other ($p = 0.174$).

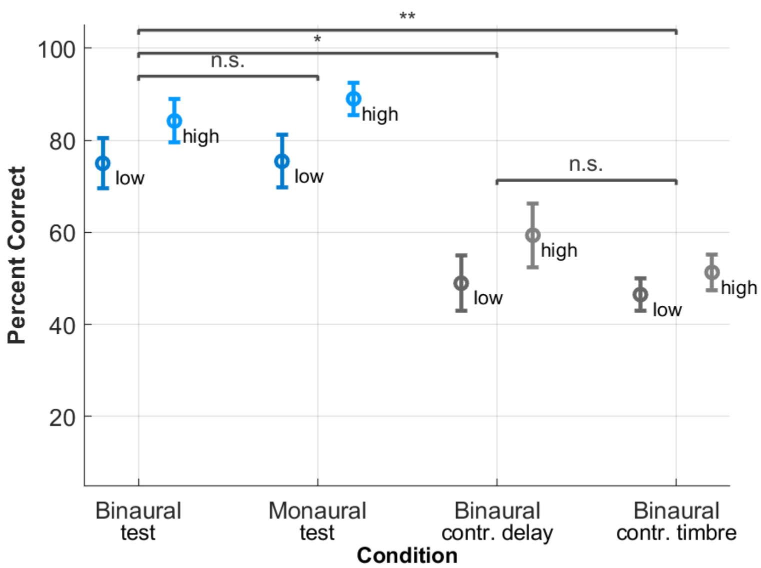


Figure 3.2: Results from 19 NH-listeners in percent of correct answers for the four conditions. Only eight of them were tested in the binaural delay control condition. In each condition, the average performance is plotted for the target stream in the high ("high"), and low ("low") frequency ranges with the standard error and indication of significant differences.

3.6 Discussion

The main aim of this study was to develop and validate a method in which the ability to perceive the scale illusion was assessed through a detection task. Three binaural and one monaural conditions were tested. We assumed that if the listeners can experience the scale illusion, they should be able to detect a deviant in the binaural test condition, but not in the two binaural control conditions. A monaural test condition was added to evaluate the average detection score that can be obtained when the target notes were easily grouped. The average score obtained shows a moderately high and similar score for both the binaural and monaural test conditions (see fig. 3.2). Furthermore, both control conditions yielded a score significantly lower that the binaural test condition. Taken together, these results indicate that this method can be used to demonstrate the ability to experience the scale illusion without the need to describe the melody pattern as in the original study (Deutsch, 1975).

Recently, the octave illusion (Deutsch, 1974) was used in Mehta et al. (2016) and Mehta et al. (2017) to assess binaural streaming in NH listeners with a detection task and either loudness or modulation cues. The octave illusion was based on only two tones, one high and one low and separated in frequency by an octave (400 and 800 Hz). These tones were presented in a dichotic sequence, so that one ear received a pattern high-low-high-low, while the other received the inverse pattern, low-high-low-high, simultaneously. As with the scale illusion, listeners could group these stimuli either based on pitch or based on location. The most common percept was a high tone lateralized to one ear alternating with a low tone lateralized to the other. Mehta et al. (2016) have modified the original method into a detection task, in which the participants had to detect a variation of loudness in a target stream. Their results demonstrated that the illusion is subject to the same constraints as auditory stream segregation and how the listener's attention influenced the obtained percepts, i.e., their lateralization. Further results in Mehta et al. (2017) revealed that the octave illusion arises from a "misattribution of time across perceptual streams, rather than a misattribution of location within a stream," providing a better understanding of the mechanisms involved in binaural streaming.

In the original experiment (Deutsch, 1975), the participants report hearing either one or two melodies. It is impossible to tell just from the performance in the binaural test condition, whether listeners perceived one stream or both high and low streams simultaneously, but with their high detection scores for the target in either frequency range, they must at least be able to attend to either stream selectively. However, the frequency range in which the deviant note is embedded had a significant impact on performance. Consistent across conditions, the detection performance in the higher frequency range is slightly higher, even though the stimuli were loudness-balanced at the different frequencies. Therefore, participants were able to focus selectively on either the low or high stream, but the better performance in the higher frequency range suggests that focusing on the higher stream was easier. This could explain the description of the single stream percept in Deutsch (1975), where listeners reported to perceive only the higher stream and nothing or little of the lower one. If it is easier to focus on the high stream and listeners voluntarily are not given further instructions, they should then be more likely to focus purely on the higher stream. The reported single stream percept is thus not at odds with the current results. Furthermore, these results match one of the conclusions from Deutsch (1985), where musically trained listeners had to identify dichotically presented sequences of tones correctly. There, listeners also performed consistently higher in transcribing higher tones compared to lower tones.

The effect of timbre on the streaming behavior could have implications for hearing-impaired patients, depending on the aids they use for listening. If these aids do affect the timbre of sounds differently across ears, binaural streaming could be affected negatively. This could apply to patients with CIs, providing electric instead of acoustic stimulation, especially for patients with a unilateral CI and acoustic hearing on the other side and this task could be adapted to assess if these changes in interaural percepts lead to changes in binaural streaming behavior for CI patients. Also, patients with highly asymmetric hearing loss, using entirely different hearing-aids across ears, could be affected. If the binaural object-formation was malfunctioning, performance in tasks relying on the evaluation of binaural cues, such as speech reception in noise and sound localization, could be reduced severely.

3.7 Conclusion

This study proposes a modification of the scale illusion experiment by Deutsch (1975) into a detection task to assess listeners' binaural streaming behavior without relying upon subjective reports of percepts. As in the original study, the NH listeners grouped the binaural stimuli by frequency-proximity instead of lateralization, for which they had to integrate stimuli from both ears into a common stream. Participants were able to focus on either the low or high stream and detect a deviant note, but it appears to be easier to focus on the higher stream. Two control conditions were added in which lateralization cues were enhanced. Detection scores dropped significantly, indicating that the task was facilitated by the ability to group stimuli by frequency-proximity. The new task could be suited to assess binaural streaming in the hearing-impaired population and explore possible changes and adaptations in the mechanisms.

4

Binaural streaming in bilateral cochlear implant patients

Abstract A detection task based on the scale illusion allows to assess if listeners sequentially integrate a dichotically presented stream (Deutsch, 1975; Janssen et al., 2019b). This study extends the method's applicability to bilateral cochlear implant (CI) patients and assesses if they can group dichotically presented stimuli into a single stream by pitch-proximity while ignoring spatial cues. Ten bilateral CI patients had to focus on a melody as a target stream and detect a note deviating from a repeating pattern in one of three intervals. Simultaneously, they were presented with a randomized distractor stream at a different pitch range. Since every second note in each stream was presented to the opposite ear, grouping of binaural stimuli by pitch-proximity, while ignoring the lateralization cues, was required to follow the target stream, and reliably detect the deviant. Eight of ten participants were able to segregate two concurrent streams by pitch-proximity monaurally. Five of these, and the CI listeners as a group, likewise demonstrated their ability to group the components of the target stream by pitch-proximity binaurally. Destroying the interaural correspondence in pitch, forced listeners to evaluate each ear separately instead.

4.1 Foreword

The previous chapter describes the new method developed to assess binaural streaming. In this chapter, this method is adapted to assess the capability for binaural streaming in a group of bilateral CI patients, where both ears are stimulated electrically by CIs. While this does not automatically mean that the percepts in both ears are the same, it may provide an opportunity to achieve a greater interaural similarity in percepts than possible for bimodal CI patients. Therefore, it marks an important step on the way to understand binaural streaming for bimodal CI patients.

4.2 Introduction

Auditory streaming is at the center of how the auditory system handles the complex acoustic scenes and, thus, plays an important role in everyday communication. It allows to segregate or to group subset of sounds into specific streams based on perceptual properties such as pitch, timbre, or spatial cues. As hearing impairment and hearing devices can affect the perception of these cues, the ability of hearing-impaired listeners to segregate auditory streams are often negatively impacted (Oxenham, 2008; Rose and Moore, 2005). In this study, we address the question whether people aided with two cochlear implants (CI) can perform streaming to segregate two melodies on the base of pitch and lateralization cues from binaural stimulation.

The formation of auditory streams is composed of two distinctive processes: sequential and simultaneous grouping (Bregman, 1990). Sequential grouping connects auditory objects over time (for example a melody composed of a sequence of notes). Simultaneous grouping combines simultaneously presented sounds into the same auditory object (e.g., the harmonics of a musical sound being grouped into the percept of a single note). Moore and Gockel (2002) argued that stream segregation was directly related to the degree of perceptual difference between sounds. Therefore, the listener could use any perceptual cue to segregate two auditory objects. In rare occasions, two such cues can be in conflict and even create an auditory illusion. Two examples of this are the octave illusion (Deutsch, 1974) and the scale illusion (Deutsch, 1975). The octave illusion has been used to assess binaural streaming for normal hearing listeners by Mehta et al. (2016, 2017), who have transformed it into a detection task using either loudness or modulation cues. However, the octave illusion relies upon rather precise pitch matches across ears to occur, which could make its application difficult with CI patients, who both face limitations regarding pitch perception and for whom precise pitch matches across ears may be impossible (Oxenham, 2018).

In the scale illusion two dichotic sequences of notes could be grouped either by pitch-proximity or by lateralization cues. Specifically, the illusion was formed by two melody streams at different pitch ranges (see fig. 4.1.), dividing the components of one musical scale into a higher melody and a lower melody, both of which ascend and descend in pitch over eight notes. Every second note from each of the two melodies was played to the opposite ear, so that when one component of the higher melody was in one ear, a component of the lower melody was in the other ear, and vice versa. Listeners could group these stimuli either by pitch range or by ear-of-input. Interestingly, most listeners did group only by pitch range, combining notes presented from the left with notes presented to the right, and vice versa. Furthermore, their percepts varied with handedness. The majority of right-handed listeners and about half of the left-handed listeners reported a higher pitch stream coming from the right, accompanied by a lower pitch stream from the left ("both streams"). The other left-handed listeners and a minority of the right-handed listeners reported to perceive only the higher pitch stream and little to nothing of the lower pitch stream ("single stream"). This illusion is only possible if the listener segregates simultaneously presented notes based on pitch or the lateralization cue (ear-of-input) and then groups sequentially presented notes into two lateralized streams based on pitch cue. Deutsch argued that the Gestalt-principle of pitch-proximity overrode the lateralization cues. Therefore, the ability to perceive this illusion can be used to assess the listener's ability to build streams from binaural stimulation.

In case that the pitch percept is weakened, altered or, especially, if the symmetry of the pitch percept between ears is affected, so that the pitch of a sound will depend on the ear of presentation, the illusion is unlikely to occur, so that listeners will then only group by ear-of-input. Such changes may be introduced to listeners by a hearing loss or the hearing devices used, particularly, if aided with CIs. CIs are used for patients with severe to profound hearing-impairment, where hearing-aids cannot provide sufficient benefit (Flynn et al., 1998). They provide stimulation by electric pulses emitted through electrode arrays, which are inserted into the cochlea. In the case of bilateral CI users, the two CI processors are not operating synchronized, which can affect the similarity in time and interaural loudness cues (Gordon et al., 2017). Moreover, different parameters such as the depth of electrode-array insertion, the nerve-fiber survival-state, the distance from the modiolus or the spiral ganglia, duration of deafness, and the fitting of the device parameters can influence the percept in both ears (Oxenham, 2018; Zeng et al., 2008). Therefore, stimulation at the same electrode number in both ears can be perceived differently. Interaural asymmetries in percepts could lead to changes in the binaural processing in these patients (Gordon et al., 2017; Polonenko et al., 2019). Since binaural alignments of percepts are not part of the standard clinical procedures, it is unclear if CI patients can use the pitch cues across ears to form stream percepts like normal-hearing (NH) listeners as required for the scale illusion to occur.

In order to assess whether bilateral CI patients can perform the binaural streaming necessary to perceive the scale illusion, some modifications of Deutsch's procedure were necessary. First, many CI listeners show some limited ability to identify a pitch contour due to the weak salience of pitch cues with the CI (Galvin et al., 2007, 2009a, 2009b; Zhu et al., 2011). Therefore, their description of their percept might be unreliable, so that a detection task appears to be more appropriate. A detection task based on the scale illusion was developed and validated in Janssen et al. (2019b), one of the streams, as grouped by pitch-proximity, was composed of randomized notes, while the other consisted of a repeating target melody. The listeners had to detect a note deviating from the target melody-pattern (cf. fig. 4.2) and the ability to detect this deviant was largely improved if the listeners were able to group the two streams by pitch-proximity. The high detection performance of the normal hearing listeners demonstrated, that they were able to perform the task and to group the two streams by pitch-proximity.

Second, to perform this task the CI listeners will need to segregate two simultaneous stimuli by pitch and/or localization cues and group sequential stimuli by proximity in pitch. To the authors' knowledge only few studies have investigated their abilities to segregate simultaneous stimuli by pitch cues (Carlyon et al., 2007; Cooper and Roberts, 2010; Deeks and Carlyon, 2004). Altogether, these studies concluded that CI listeners were unlikely to utilize the temporal pitch cues provided via pulse rate changes to segregate or group streams, albeit Deeks and Carlyon (2004) also found that NH listeners in a CI simulation were able to follow a voice stream presented together with a masker stream at a slightly lower pulse rate (77 pps vs. 100 pps), but not when the masker was presented at the higher and the target speech at the lower pulse rate. The latter results suggest that the pulse rate could be utilized as a grouping cue, perhaps with better results at larger pulse rate differences.

The ability of CI patients to group or segregate sequential stimuli based on temporal pitch cues via the pulse rate (Duran et al., 2012) and place pitch cues via the electrode position (Chatterjee et al., 2006; Cooper and Roberts, 2007) have been studied and shown that both cues can be used for streaming. However, along with the saturation of the temporal pitch percept beyond pulse rates of 300 pps (Zeng, 2002), also the effects as a cue for stream segregation diminish along with the perceptual differences (Duran et al., 2012). Regarding the electrode place difference, it was found that for some patients a difference of about 0.75 mm along the electrode array, i.e., one electrode for Cochlear devices used (Zeng et al., 2008), may be sufficient for sequential stream segregation, while others can require 3 mm or more to reliably segregate sequential stimuli, due to the reduced perceptual space of the electrode place pitch coding (Chatterjee et al., 2006).

Moreover, Kan et al. (2013) and Reiss et al. (2017) found that most bilateral CI patients report to perceive only a single sound when presented with two equally loud sounds with different pitches and send to opposite ears. This means those patients could binaurally fuse the short dichotic stimuli, i.e., integrate them into a single common percept, over a wide pitch range, spanning up to 12 mm along the electrode array, i.e., encompassing the entire array. This could present a potential hindrance to perceiving the scale illusion, if the simultaneously presented components of the two melodies were instead integrated into a single stream. Such integration would lead to chance-level performance, as all binaurally fused pitches in the paradigm would have a random pitch, since one of the two streams was randomized.

Consequently, it may be advisable to provide a larger separation in pitch across target and distractor streams. The results of Luo et al. (2012) suggest that CI patients can effectively integrate (consistent) place and temporal pitch cues, so that combined electrode place as well as pulse rate cues could allow more CI listeners to overcome the issue of broad pitch fusion and, on top of this, segregate the two streams of the scale illusion.

Third, for the illusion to be effective, it is required that sequential notes played on opposite ears will be perceived as having a small difference in pitch. In NH listeners, it can be assumed that 200 Hz tone played on one ear will be perceived as fairly similar to a 210 Hz tone played on the other ear. But, as discussed earlier, various physiological reasons, such as varying insertion depth or distance to the modulus, can lead to the situation that a pulse train presented on, e.g., electrode 15 is perceived as much higher, or much lower, than a pulse train

presented on electrode 14 in the opposite ear. Thus, to provide the best-possible foundation for allowing the listeners to perceive the illusion, in this study stimuli should be delivered using direct bilateral stimulation of the implants, through electrodes perceptually matched across ears.
For this study we hypothesized that by implementing those three modifications, bilateral CI listeners might experience the scale illusion. Finding out if CI patients can form streams by pitch-proximity from binaural stimulation like NH listeners could aid optimization of the implants' processing, stimulation scheme, fitting parameters, and potentially lead to better speech reception in noise, as well as reduced listening effort. Moreover, the paradigm allows to assess if CI listeners can form separate streams of simultaneously presented stimuli when both streams are presented monaurally.

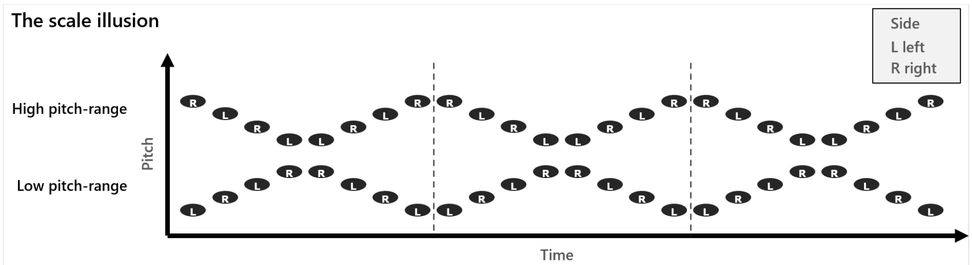


Figure 4.1: Schematic representation of the stimuli forming the scale illusion (Deutsch, 1975). There are two melody streams, one at a higher and one at a lower pitch range. Both are formed by components presented sequentially to the opposite ear, i.e., left-right-left-right-right-left-right-left, presented so that when a note from the higher stream is in the left ear, a note from the lower stream is in the right ear, and vice versa. This is indicated by the black (left) and grey (right) of the ovals indicates the ear of presentation in the binaural test condition.

4.3 Participants

Ten bilateral cochlear implant (CI) patients participated in this study (aged 48 to 82; average 65.2 years), who became deaf post-lingually. One additional participant was tested but failed an inclusion criterium (cf. 4.4.3). Five of the listeners were female, none musically trained, all right-handed, and all CIs were produced by Cochlear (Macquarie University, Australia). They were either recruited at the Technical University of Denmark or at the German Hearing Center of the Medical University Hannover, Germany, and provided written informed consent. All experiments were approved by the Science-Ethics Committee for the Capital Region of Denmark (reference H-16036391) and the ethical committee of the Medical University Hannover (reference 7885 BO S 2018). More information on the participants can be found in tab. 4.1.

4.4 Method

Stimuli were monopolar 200 ms pulse trains on single electrodes, with stimulation rates of either 70 pps or 600 pps. The pulses were biphasic with a pulse phase of 25 µs and a pulse gap of 8 µs. Stimuli were presented directly through a RF Gen XS research stimulation platform (Cochlear Ltd., Macquarie University, NSW, Australia), allowing for synchronous bilateral stimulation.
All stimuli were loudness-balanced to the most-comfortable presentation level in a procedure alike to the clinical fitting procedure, for electrodes in the range of electrode no. 22 at the apical end of the array to no. 7 towards the basal end. Patients had to indicate the loudness on a seven-step scale (1: inaudible; 2: barely audible; 3: soft; 4: comfortable, but a little soft; 5: most-comfortable; 6: comfortable, but a little loud; 7: too loud, in the participant's native language), allowing for a fine indication of the desired presentation level. The level was adjusted by the experimenter according to listener feedback; starting in small steps of six current units (CU) from no current and using one CU steps around the most-comfortable level. One CU is equivalent to about 0.168 dB for the Cochlear devices used. The stimuli were loudness-balanced in random order and the results verified by a final comparison of all stimuli. If necessary, additional fine adjustments were made in this step.

4.4.1 Stimuli

4.4.2 Interaural alignment and electrode selection

The electrodes used for stimulation were interaurally matched in percepts, i.e., not just pitch, with the same stimuli parameters as used later in the experiment, i.e., loudness-balanced pulse trains of 200 ms duration at pulse rates of 70 pps or 600 pps. The matching's aim was to find eight electrodes with different place pitch per side, four for a lower pitch range and four for a higher pitch range,

Table 4.1: Demographic information on the bilateral CI participants with times given in years

ID	Etiology	Age	Dur. CI use / a		Dur. hearing loss prior to CI / a		Dur. deafness prior to CI / a		Implant	
			Left	Right	Left	Right	Left	Right	Left	Right
1	Unknown	67	3	2	15	16	0	0	CI522	CI522
2	Noise induced	72	10	12	14	12	0	0	CI24R (CA)	CI24R (CA)
3	Acute hearing loss	82	16	7	17	26	0	0	CI24R (CS)	CI512
4	Acute hearing loss	61	5	7	16	14	0	5	CI422	CI422
5	Acoustic shock, middle ear infection	66	9	15	39	27	6	0	CI24 RE (CA)	CI24R (CA)
6	Middle ear infection, acute hearing loss	70	9	10	60	60	0	0	CI24 RE (CA)	CI24 RE (CA)
7	Acute hearing loss	69	1	3	43	37	0	4	CI532	CI512
8	Anemia	68	6	7	22	21	0	0	CI422	CI512
9	Middle ear infection	66	9	17	44	36	0	0	CI24 RE (CA)	CI24R (CS)
10	Unknown	76	15	9	15	15	0	0	CI24R (CS)	CI24 RE (CA)

with best-matching correspondents across ears. In between the lower four and the higher four electrodes, a gap of up to four electrodes was left, if the difference ratings allowed. This enhanced the place-pitch differences between the two pitch ranges. So, for an ideal case with electrodes interaurally matching in percepts by electrode number, the selection would be electrodes 22 to 19 from the apical end of the array for the lower pitch range and electrodes 14 to 11 from the middle region of the array for the higher pitch range on both sides.

In each trial of the matching-procedure, participants were presented with combinations of stimuli from one electrode on the left and one electrode on the right side, which they could repeat as often as they desired. Then, they had to rate how different the two stimuli sounded using an on-screen slider with a scale from "no difference" (rating zero) to "absolutely different" (rating one). Explanations and labeling of the user interface were given in the participants' native language. For each interaural combination of electrodes, the rating was repeated three times and the results were averaged. Electrodes with the smallest difference were considered as pairs.

The electrode combinations for testing were limited by the experimenter, which greatly reduced the number of necessary comparisons and combinations from the selection of electrodes were assessed in random order. The side presented first was also chosen randomly in each trial. Electrodes were evaluated separately for the lower pitch range with a pulse rate of 70 pps, and for the higher pitch range with a pulse rate of 600 pps. In a first step, three electrodes from each side were evaluated with one electrode in between, no. 22, 20 and 18. If no clear closest-match was found from these electrodes, the evaluation was repeated with a different selection of electrodes, shifted according to the direction of decreasing difference-ratings. Once a closest match was found, the matching electrode pair was used as a basis to select six consecutive electrodes around the pair, including the electrodes adjacent to the ones used in the initial comparisons. This selection was then evaluated and optimized using the same procedure until four interaural electrode pairs were found. If the matches allowed, consecutive electrodes were used on each side, but in case of ties for electrodes with a very close average difference ratings (< 0.1), one of them was excluded to ensure that patients could perceive the differences between the electrodes. The process was then repeated for the higher pitch range, starting with electrodes 14, 12 and 10. Finally, the assessment was repeated for the selection in the lower pitch range, to ensure that the participants, now with

Table 4.2: Electrodes selected for the bilateral CI participants per side and pitch range

	Electrodes per side and pitch range (low/high)			
	Left		Right	
ID	Low	High	Low	High
1	22 21 20 19	14 13 12 11	22 21 20 19	14 13 12 11
2	22 21 20 19	13 12 11 10	20 19 18 17	13 12 11 10
3	22 21 20 19	13 12 11 10	22 21 20 19	14 13 12 11
4	22 21 20 19	11 10 09 08	22 21 20 19	11 10 09 08
5	21 20 19 18	13 12 11 10	21 20 19 18	13 12 11 10
6	22 21 20 19	14 12 10 09	22 21 20 19	14 12 11 09
7	22 21 20 19	14 13 12 10	20 19 18 17	14 13 12 10
8	22 21 19 18	12 11 10 09	21 20 19 18	13 12 11 09
9	21 20 19 18	14 13 12 11	14 13 12 11	07 06 05 04
10	22 20 19 17	14 12 10 08	22 20 19 17	14 12 10 08

greater experience in the matching task, still paired the same electrodes. The electrode pairs found for the participants are listed in tab. 4.2.

4.4.3 Procedure

The test procedure used the paradigm developed in Janssen et al. (2019b), a detection task based upon the auditory scale illusion described in Deutsch (1975). The method was based on two melodies presented concurrently. Participants were asked to focus on a target melody stream. This melody stream was built of a repeating pattern of eight notes going up and down in pitch like the melodies in the original scale illusion experiment (cf. fig. 4.1). The target melody stream was either of a lower or a higher pitch range, while a distractor melody stream was presented simultaneously with notes picked randomly from the other pitch range. If the target stream was of the higher pitch range, the distractor stream was of the lower pitch range, and vice versa (cf. fig. 4.2). As in the experiment of Deutsch (1975), notes from these two streams were presented at the same time, so that if a note from the target stream was on the left, a note from the distractor stream was on the right, and vice versa. Consecutive notes for both streams were presented to the opposite ear, in a pattern that reversed in order after every four notes. Thus, the arrangement formed a pattern symmetric to the middle of the repeating sequence of eight notes (e.g., for one stream in one presentation of the eight notes: left-right-left-right-right-left-right-left).

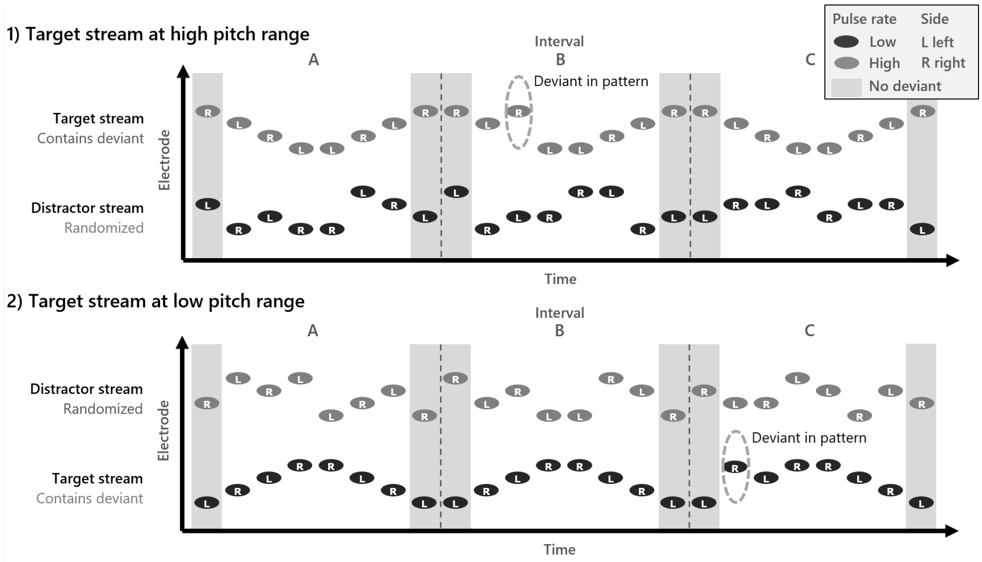


Figure 4.2: Schematic representation of the experiment's stimuli, adapted from Janssen et al. (2019). A random deviant occurs in the target stream, either at the higher (top) or lower (bottom) pitch range. The distractor stream consists of notes picked randomly from the other pitch range. The black (low) and grey (high) of the ovals indicates the pulse rate.

If the bilateral CI listeners behaved like the normal-hearing listeners in Deutsch (1975) and Janssen et al. (2019b), they should group the stimuli by pitch-proximity, and be able to follow the target stream on one side, with the concurrent distractor stream either perceived on the other side ("both streams" percept) or not at all ("single stream" percept). These percepts require the integration of sequential dichotic stimuli into a common stream. To test whether the listener could follow the target stream, a deviant note was introduced into it. The deviant consisted of shifting a note's respective place pitch, i.e., electrode, to the second-next electrode within the respective pitch range's four electrodes (cf. deviant indicated in interval B in the top drawing of fig. 4.2). Listeners who grouped by pitch-proximity, were expected to show a high detection-performance. If the CI listeners were grouping the sequential dichotic stimuli by ear-of-input, they would have to evaluate both ears independently. Due to the random nature of the distractor stream's components, every other stimulus in both ears was random, making the detection of a deviant more difficult.
To indicate which melody was the target stream, it was twice presented alone in the beginning of the stimulation sequence during each trial. Target and distractor streams were presented together six times after that. The first three of the six repetitions served the purpose to allow for a build-up of streaming, since

the formation of segregated percepts have been reported to arise gradually over several seconds (Bregman, 1990; Paredes-Gallardo et al., 2018b, 2018c). The last three repetitions form the three intervals of which one contains the deviant note, such as indicated for interval B at the top of fig. 4.2. The presented interval was indicated on a graphical user interface by lighting up the respective part of the stimulus on the screen. Listeners were allowed to repeat the stimuli. The deviant note was placed randomly at any of the inner six of the eight notes within one of the three last repetitions of the target stream, but never the first or last position. The placement was counterbalanced regarding the side of presentation, interval, and pitch range of the target stream. Since participants had to choose the deviant interval out of three options, the task was not prone to a detection-bias (Wickens, 2002), allowing for an "objective" assessment of the listeners' grouping behavior.
Participants were trained in the task before conducting the test. The training differed from the experiment in that only the binaural and monaural test-conditions were used (cf. 4.4.4) and that the participants were provided with feedback on whether they answered correctly. In total, the experiment could take four to six hours including breaks and was divided into two sessions. In the first session, loudness balancing and electrode pairing were conducted, followed by a test to determine if participants could perform the task when there was only the target melody presented. This test was analog to the main experiment, but only presented the target stream monaurally in twelve trials and with the same counterbalancing. Participants had to score at least 75 % correct in this task and listeners unable to reach this performance were excluded from further testing. This affected one listener. In the second session, the training and final testing were performed. These two sessions were performed on separate days to reduce fatigue.

4.4.4 Conditions

Four conditions were used in the experiment, as described below. Each condition was composed of twenty-four trials. The order of trials was randomized per participant across all conditions to prevent order effects on the results.
A binaural test-condition presented both the target and the distractor stream in the same way as the original scale illusion, i.e., every second note from each stream was presented to the other ear (fig. 4.3). Hence, this condition required

both segregation of the simultaneously presented dichotic stimuli, as well as grouping of sequential stimuli, by pitch-proximity (due to rate and place cues) to perceive, i.e., segregate, the target stream and detect the deviant.

A monaural test condition presented both streams to the same ear instead and,

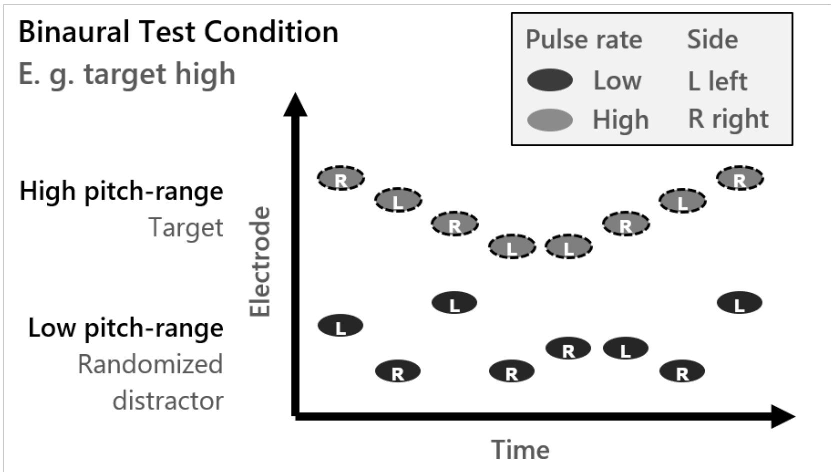


Figure 4.3: Schematic stimuli representation for the binaural test condition on the example of the target stream (marked with dashed lines around the notes) at the higher pitch range with the randomized distractor stream at the lower pitch range. The black (low) and grey (high) of the ovals indicates the pulse rate.

hence, did not require binaural processes (fig. 4.4). Listeners should always segregate the simultaneous stimulation based on the difference in place and rate pitch and group sequential stimuli by pitch-proximity in this condition. They should, therefore, be able to easily detect the deviant. This allowed for an assessment of the best-possible performance in the task without binaural processes, and whether listeners are able to form separate stream-percepts out of simultaneous stimulation at all. Therefore, it allows to assess the contribution of the binaural processing when compared to the binaural test condition.

In a previous study, we have shown that when normal listeners were presented with a control condition in which the perceptual difference across the ears was enhanced, they could not detect the deviant within the target stream reliably (Janssen et al., 2019b), as they grouped the two sequences by ear-of-input instead of pitch-proximity. Consequently, instead of perceiving the target melody in one ear and the distractor in the other, they would perceive two random streams in both ears. In the current experiment the distractor and the target melody were presented with two different rates, so the CI listeners could the-

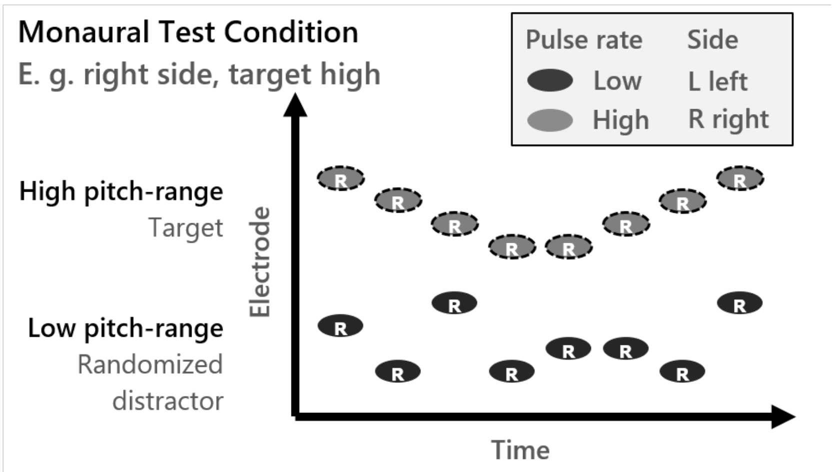


Figure 4.4: Schematic stimuli representation for the monaural test condition on the example of the target stream (marked with dashed lines around the notes) at the higher pitch range with the randomized distractor stream at the lower pitch range, both presented to the same ear, e.g., the right. The black (low) and grey (high) of the ovals indicates the pulse rate.

oretically have been able to detect the deviant by attending to one ear and segregating the target from the distractor components. In that case, they would perceive two streams, one composed of a two notes melody and one of random notes. If the deviant occurs in the attended ear, it should be easily detectable, but this would be complicated by the randomization of the deviants' position regarding the ear-of-input. Therefore, in order to assess if CI listeners used that cue, it is important to introduce a control condition that will promote this strategy. If CI listeners experienced the illusion as described in Deutsch (1975), their score should be higher than in any condition promoting monaural listening.

Two such binaural control conditions presented the target and distractor streams binaurally in the same pattern as the binaural test condition, but with the intend to force listeners to group by ear-of-input instead of using pitch cues. To achieve this, the interaural correspondence in pitch cues was destroyed.

A binaural electrode place control condition interchanged the electrodes used to represent the place pitches between the two pitch ranges but maintained the pulse rates for high and low pitch ranges (fig. 4.5). This destroyed the interaural correspondence in place pitch but preserved the rate-pitch cue, leading to a mismatch in overall pitch-percepts across ears. Thus, participants were expected to group the stimuli by ear-of-input in this case, leading to a reduced detection-performance. This condition allowed to quantify the possible detection performance when grouping by ear-of-input, with elements from both target- and the randomized distractor stream present.

A binaural pulse rate control condition interchanged the pulse rates used for

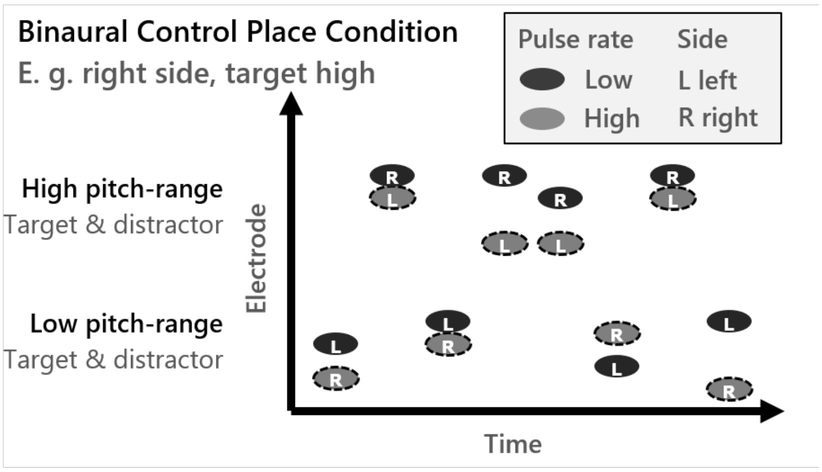


Figure 4.5: Schematic stimuli representation for the binaural control condition with mismatched electrode place pitch cues across ears, e.g., with inverted electrode place cues for the right ear. In this example, the target stream (marked with dashed lines around the notes) is at the higher (pulse rate) pitch range with the randomized distractor stream at the lower (pulse rate) pitch range. The black (low) and grey (high) of the ovals indicates the pulse rate in the binaural test condition.

the lower and high pitch range in one ear compared to the other, while keeping the electrode selection the same. This destroyed the interaural correspondence in rate pitch but preserved the correspondence in place pitch (fig. 4.6), again leading to a mismatch in overall pitch percepts across ears. This allowed to evaluate the effect of rate-pitch mismatches on the streaming behavior of the participants and to compare it to that of place-pitch mismatches. As with the place-pitch mismatches, participants were expected to group the stimuli by ear-of-input in this condition, resulting in a reduced detection performance.

4.5 Results

One participant failed to meet the inclusion criterium to reach a detection performance of at least 75 % correct when performing the task with the target-melody alone. For the other ten participants, the individual performance in the four conditions is depicted in fig. 4.7 with individual plots for each listener. The

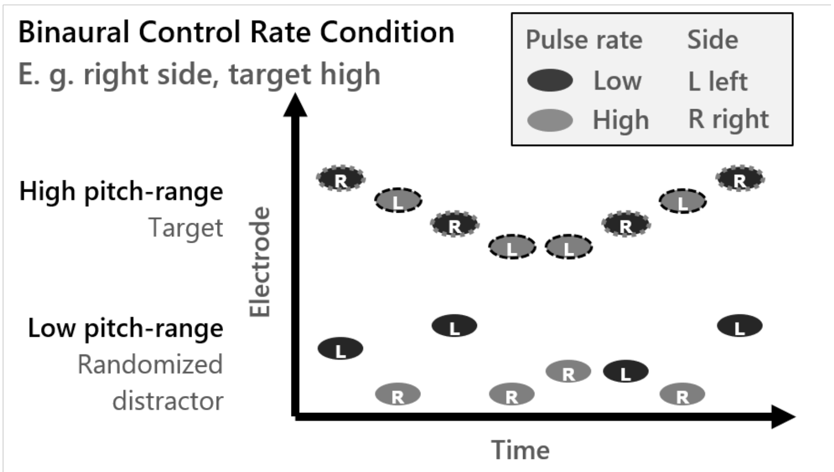


Figure 4.6: Schematic stimuli representation for the binaural control condition with mismatched pulse rate pitch cues across both streams for one ear. In this example, the target stream (marked with dashed lines around the notes) is at the higher (electrode place) pitch range with the randomized distractor stream at the lower (electrode place) pitch range. The black (low) and grey (high) of the ovals indicates the pulse rate, which is interchanged for the right ear in this example.

participant ID is plotted in bold type at the top right of every individual plot. A dotted line in the figures indicates 75.0 % correct, the performance criterion at which the individual performance in a pitch range becomes significantly different from chance. This level was determined according to the binomial distribution for twelve trials per pitch range and using the Bonferroni-correction for eight comparisons at the significance level of five percent (nine or more correct trials out of twelve trials results in $p \leq 0.00385$). The performance in the monaural test condition fell at or above this criterion for eight out of ten participants (no. 1, 3, 4, 6, 7, 8, 9, and 10) in at least one pitch range, while all performance in the control conditions fell below that limit. Of these eight listeners, five (no. 1, 4, 7, 8, and 10) also showed such high performance in the binaural test condition. Two participants (no. 2 and 5) did not show performance significantly different from chance in any condition.

The average results in the four conditions are depicted in fig. 4.8, plotted as the detection performance in percent of correct answers for the binaural and monaural test conditions, as well as the binaural control conditions with mismatches in pulse rates and electrodes across ears. Each condition has two data points, indicating the average performance in the lower pitch-range ("low") and the higher pitch-range ("high"). The binaural test condition shows on average a slightly lower performance than the monaural test condition, about

48 % vs. 66 % for the lower pitch-range and about 58 % vs. 75 % for the higher pitch-range. Average performance in the binaural control conditions is lower at around 40 % for the condition with misaligned pulse rates and between 36 % and 50 % for the misaligned electrodes. All conditions but the binaural control condition with mismatched pulse rates show higher performance in the higher pitch range.

A two-way repeated analysis of variance was performed with the "rationalized" arcsine transform (rau) applied to the detection scores (Studebaker, 1985), which were the dependent variable, and the conditions, as well as the target stream's pitch range as factors. The condition was a significant factor ($F(3, 27) = 19.14$, $p < 0.0001$), but not the pitch range ($F(1, 27) = 3.5$, $p = 0.0944$). However, the interaction between condition and pitch range was significant ($F(3, 27) = 3.27$, $p = 0.0363$).

Posthoc analysis was carried out using paired T-tests on the rau-transformed performance data and considering the Bonferroni correction (criterion: $p = 0.01$). The scores in the binaural and monaural test conditions were not significantly different from another ($p = 0.0850$). The performance difference between in the binaural test condition and the binaural controls with mismatched pulse rates ($p < 0.001$) and mismatched electrodes were significant ($p < 0.001$). Taking into account that the interaction of condition and pitch range was significant, also the differences across higher and lower pitch ranges were assessed for the binaural ($p = 0.118$) and monaural test conditions ($p = 0.0820$), neither difference was significant.

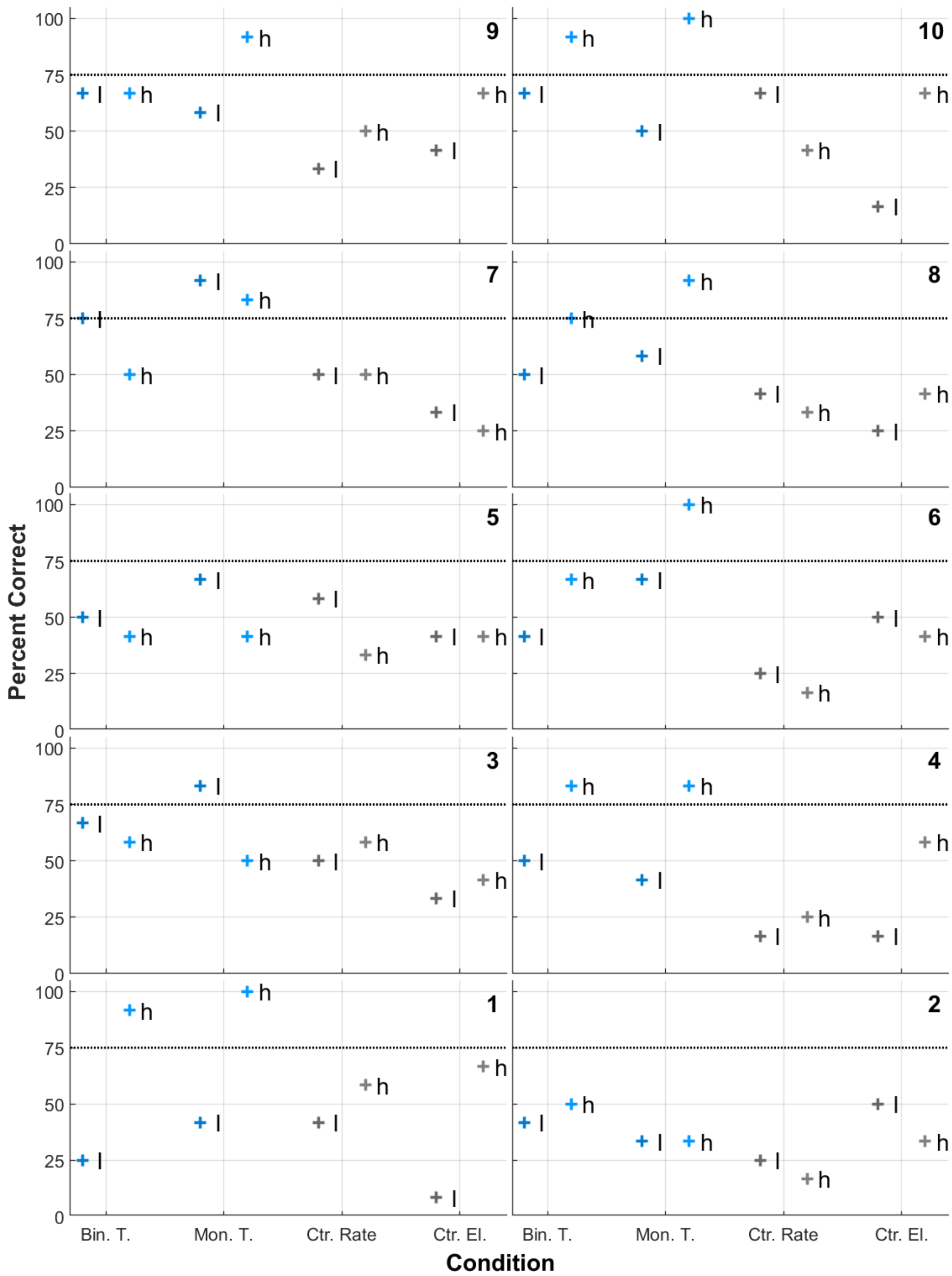


Figure 4.7: Individual results from the ten bilateral CI listeners in percent of correct answers for the four conditions (denoted “Bin. T.”: binaural test condition, “Mon. T.”: monaural test condition, “Ctr. Rate”: binaural control condition with mismatched pulse rate, and “Ctr. El.”: binaural control condition with mismatched electrodes) with one plot per participant. A bold number at the top right of every plot indicates the participant ID. In each condition, the performance is plotted for the target stream in the lower (“l”), and higher (“h”) pitch ranges.

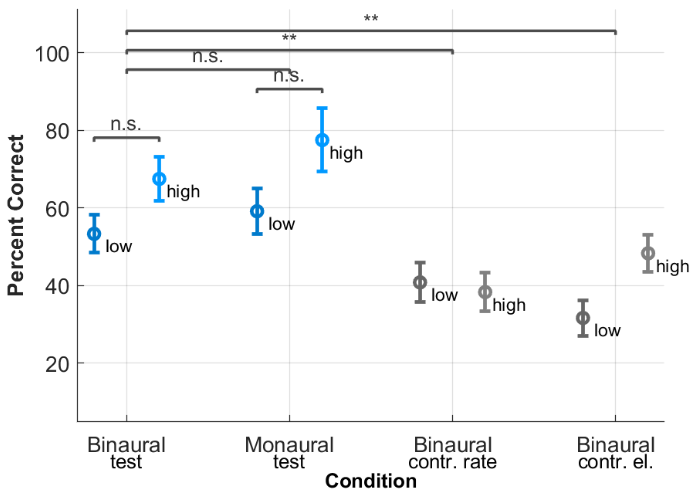


Figure 4.8: Results from ten bilateral CI listeners in percent of correct answers for the four conditions. In each condition, the average performance is plotted for the target stream in the lower ("low"), and higher ("high") pitch ranges with the standard error and indication of significant differences.

4.6 Discussion

In this study, the bilateral CI listeners had the task to focus on a target stream and point out the interval in which a deviant note occurred. In one of the four conditions, i.e., the binaural test condition, this required them to use binaural processing to a) segregate simultaneous dichotic stimuli based on pitch (and possibly also the lateralization cue provided by ear-of-input) and b) to group sequentially presented stimuli by proximity in pitch across ears, while ignoring lateralization cues. The high performance in the binaural test condition of the bilateral CI listeners as a group indicates, that they were able to perform this binaural streaming with grouping behavior similar to the normal-hearing listeners in Janssen et al. (2019b) and the original study on the scale illusion in Deutsch (1975). Notably, this was achieved at rather large pitch differences through combined temporal pitch and place pitch cues, as provided via pulse rate and electrode place.

When the listeners had to perform the same task monaurally, their average performance was even a bit higher compared to the binaural test condition, but the difference between the two conditions was not significant. This demonstrates that the listeners were also able to segregate concurrent stimuli into separate streams based on proximity in pitch when no lateralization cues were present and no binaural processes were required. This suggests, that while the lateralization cues were present in the arrangement of the binaural test condition and could have aided segregation of the simultaneous dichotic elements of the stimuli, they were not required to perform the streaming necessary to perceive the scale illusion.

The significantly lower performance in both binaural controls is a result of the mismatches in interaural pitch percepts, introduced either via changes in pulse rates across streams or via changes in electrode place. These changes affected the proximity in pitch-percepts required to perform the grouping of sequential stimuli into the target and distractor steams, and, thus, prevented listeners from perceiving the scale illusion. Therefore, they could only evaluate the stimuli grouping by ear-of-input and, due to the random nature of the distractor stream components, achieve only a limited detection performance, that was not significantly different from chance. A similar reduction in performance was obtained with normal-hearing listeners when mismatches in timbre percepts across ears were introduced (Janssen et al., 2019b). Since the difference between both control conditions was non-significant, we can conclude that destroying the correspondence in rate-pitch cues was just as detrimental for the binaural streaming as destroying the correspondence in place-pitch cues, despite the participants' clinical processors do not utilize the pulse rate to encode information.

At the individual level, five out of the ten bilateral CI listeners reached a performance level significantly different from chance in the binaural and monaural test conditions, showing that they were grouping by pitch-proximity both monaurally, as well as binaurally. Three other participants showed a lower performance in the binaural test conditions instead but still performed significantly above chance in the monaural test condition. This leads to the conclusion, that they were able to segregate the simultaneous stimuli and form streams from sequential stimuli based on pitch-proximity only monaurally, but not when binaural processes were required to integrate the stimulation across ears. Another two listeners did not score significantly above chance in any condition apart from the initial test with just the target stream, which indicates that even monaurally, they were unable to (reliably) form separate simultaneous streams from the stimuli like the other participants.

Since the participants overall showed a slightly lower performance in the binaural test condition compared to the monaural one, one might speculate about changes in the binaural grouping-processes, which made the task more difficult when binaural processing is required. Nevertheless, this difference was non-significant, so that the data does not allow for such a conclusion. A future study with a larger number of participants could clarify this, potentially combining the task with an assessment of listening effort, e.g., through pupil-dilation measures.

For both binaural and monaural test conditions, the results showed a trend towards higher scores in the higher pitch range, but this difference was non-significant. Still, the task may have been easier to perform for the target in the higher pitch range. This fits in with results from musically-trained normal hearing listeners, who had to identify dichotic sequences of tones and performed consistently higher for the stimuli higher in pitch (Deutsch, 1985), as well as with the results of Deeks and Carlyon (2004), who investigated concurrent streaming with a target speech stream and a masker stream, presented to normal-hearing listeners via a CI simulation at different pulse rates. Their participants could follow the target stream when it was presented at the higher rate compared to the masker, but not when the masker was presented at the higher rate. This suggests that the effect making it easier to follow a stream at a higher pulse rate or pitch may not be limited to the specific stimuli used here and could become a significant factor for larger groups of participants. An optimization of the experiment design could account for this trend, by always presenting the target stream at a higher pitch range. This could still be tested with electrodes over the full range of the arrays, by always using a high pulse rate for the target stream and a low one for the distractor.

As mentioned in the introduction, it may be of interest to compare the present results to findings for binaural fusion, since the present task relies on listeners segregating simultaneous, dichotic stimuli instead of integrating them into a common percept. Kan et al. (2013) used short electric pulse trains delivered synchronously to interaurally pitch-matched electrodes and deliberately mismatched interaural electrode-combinations to test the effect of interaural pitch-mismatches on binaural fusion and sound localization for short dichotic stimuli. They found that bilateral CI users reported to perceive one sound with electrode place-mismatches up to about 3 mm along the electrode array, corresponding to four electrodes for the Cochlear devices used. But increasing interaural pitch-mismatches resulted in lateral sound-image shifts. At higher mismatches, participants were more likely to report percepts of two separate sounds. Interaural mismatches in perception also increased the just-noticeable differences (JND) for interaural time-difference (ITD) and loudness-difference (ILD) cues, limiting the listeners' abilities to accurately detect the location of sounds. Reiss et al. (2017) also investigated binaural fusion in bilateral CI patients with short, dichotic stimuli. Listeners were presented with loudness-balanced, biphasic pulse-train stimuli from various interaural electrode-combinations. They had to answer whether they perceived (1.) one and the same single sound in both ears, (2.) one sound in the left ear only, (3.) one sound in the right ear only, or two sounds with either (4.) the left ear higher in pitch, or (5.) the right ear higher in pitch. Stimuli were interpreted as fused in case the participants answered same (1.) or if they were reported as lateralized to either side (2., 3.). Results for the bilateral CI listeners showed abnormally-wide fusion ranges instead, averaging at 6.1 ± 3.9 mm distance along the cochlea, with minimum range of about 1.7 mm up to a maximum of 10.3 mm or even 12.0 mm electrode distance. This corresponds to stimuli being fused binaurally over a range of two to fourteen electrodes with respect to the reference electrode in the other ear for Cochlear

CIs, an even larger range than found by Kan et al. (2013). The difference between the two studies may partly be explained by the interpretation of the data. Reiss et al. (2017) interpreted lateralized stimuli as binaurally integrated, while Kan et al. (2013) assessed binaural fusion and lateralization separately. Steel et al. (2015) suggested that there may be two separate psychophysical ranges for binaural fusion and lateralization, based on the results in pediatric bilateral CI patients. Hence, while the extend of the pitch range over which bilateral CI patients fuse such short dichotic stimuli may be unclear, it appears to be larger than for NH listeners, which fuse over up to 0.3 octaves (Brink et al., 1976). This could, thus, also affect their abilities for binaural streaming, both due to limited differentiation in pitch percepts, but also through possible (mal-) adaptations of the binaural processes involved.

Nevertheless, the bilateral CI listeners in the present study were able to segregate the simultaneously presented bilateral stimuli into different streams and focus their attention on one of them, while the electrodes generally were between about 3.75 mm and 9 mm apart along the arrays. This leaves enough separation by place pitch for monaural sequential stream segregation (Chatterjee et al., 2006; Cooper and Roberts, 2007) and to generally prevent binaural fusion according to Kan et al. (2013), but would be too small a separation in electrode place according to Reiss et al. (2017). However, the two pitch ranges in this study were further separated by rate pitch. Both cues combined appear to have separated the streams' components sufficiently, so that listeners did not fuse the dichotic stimuli in the paradigm. The place and temporal pitch cues may have been integrated into a common pitch percept, as Luo et al. (2012) suggested, widening the perceptual separation across streams. If listeners had fused the stimuli into one common sound percept, they would have perceived (most of) the stimulation in the current experiment as only one melody of random pitches out of the combination of target and randomized distractor, resulting in chance level performance.

In a pilot study for this experiment the same pulse rate, 900 pps, was used for both higher and lower pitch ranges and the two ranges were presented with a minimum separation of about 2.25 mm along the electrode arrays for the two pitch ranges. With these parameters, three out of five bilateral CI listeners performed high in the monaural test condition and two of those also in the binaural test condition, indicating that the two listeners streamed binaurally, while the other one could only perform the streaming monaurally. The remaining two listeners always performed at chance level. Thus, for some listeners the two streams may only be segregated with a pitch difference enhanced by use of different pulse rates, while for others the additional context provided by the stimuli continuing over time may already lead their auditory system to group over a smaller pitch range. The limits of this binaural grouping range could be evaluated in further studies, e.g., by performing the task while systematically varying rate-pitch and place-pitch differences between the two streams. Also, a tracking scheme could be implemented to determine the necessary separation for streaming. This could benefit the strategies for encoding multiple, competing streams for CIs and might also indicate if binaural percept alignments should be integrated into clinical practice. Furthermore, this could help optimize the patients' speech reception and reduce their listening effort. This may also apply to patients aided bimodally, i.e., with a hearing-aid contralateral to the CI.

The present results are of interest also on the background of several studies about concurrent stream segregation in CI listeners which concluded that CI patients would likely not utilize temporal pitch differences, i.e., pulse rate, to segregate streams (Carlyon, Long, Deeks, and McKay, 2007; Deeks and Carlyon, 2004). While this may be true for small differences, such as the 77 pps vs. 100 pps used by Carlyon et al. (2007), already the results in Deeks and Carlyon (2004) show effects of pulse rates, with the stream presented at a higher pulse rate easier to follow, as discussed earlier. The results of the present study in both the binaural and monaural test conditions demonstrate that the CI listeners were able to segregate two concurrent streams based on pitch cues resulting from a combination of electrode place and pulse rate, whether presented to the same ear or across ears. This combined effect of both cues becomes more apparent considering the pilot testing for this study without differences in pulse rate across streams, as well as the results of the studies regarding stream segregation and binaural fusion discussed above (Chatterjee et al., 2006; Cooper and Roberts, 2007; Kan et al., 2013; Reiss et al., 2017), all of which suggest that a higher perceptual separation is necessary for segregated percepts. Moreover, the binaural control with mismatched pulse rates demonstrates the significant effect of the rate cues on binaural streaming.

This study was conducted using an interaural electrode-pairing and for most participants, the best matches were not those used in the clinical mapping. The clinical mapping for all participants just assigned the same frequency-bands to electrodes by number with the only exceptions being deactivated electrodes. The results from the electrode pairing suggest instead, that this clinical mapping might be suboptimal, as often the electrodes with the same number are not perceived the same. For five out of ten participants, a shift by one or two electrode numbers, i.e., up to 1.5 mm along the electrode array, produced stimuli perceived more similar and in one case, a shift of seven electrodes was necessary, i.e., 5.25 mm along the electrode array. These shifts also often only affected one of the two pitch ranges and, hence, must be assessed over the entire electrode arrays for optimal interaural matchings. Hence, all but one of the participants fulfill the limit of 3 mm of interaural mismatch required to binaurally fuse sounds according to Kan et al. (2013) with their clinical standard mapping. But at least for sound localization, even such small mismatches already resulted in a distorted perception, Kan et al. (2013) reported. It might, therefore, be necessary to address binaural electrode-pairing clinically, at least for the sake of those CI patients with larger mismatches, even if they present a minority. This would have to be done over at least multiple points along the electrode arrays, to cover eventual changes in matchings over the extend of the arrays. One could investigate the effect of such matchings on the patients' acclimatization to the CI stimulation after implantation and how that reflects in outcome measures such as speech reception in noise and listening effort. Speech reception in noise, as a more complex measure, may require that participants grow accustomed to the interaurally matched pitch-encoding, making such studies more complex and time-consuming.

4.7 Conclusion

This study assessed if bilateral CI patients can perform streaming based on a combination of electrode place and pulse rate pitch cues, both monaurally and when they needed to combine percepts from binaural stimulation. The paradigm used to test this was a detection task based on a modification of Diana Deutsch's scale illusion experiment. The participants demonstrated their ability to monaurally segregate two concurrently presented streams based on pitch. Moreover, they were able to perform both segregation of simultaneous dichotic stimuli and grouping of sequential stimuli across ears by pitch-proximity when the two streams were presented binaurally, so that binaural processing was required to combine stimuli from both ears into one stream. On an individual level, two listeners were unable to perform the necessary streaming monaurally. Of the remaining eight listeners, five were able to perform the streaming both monaurally and binaurally, while the remaining three were only able to stream monaurally but not when binaural processing was required. The ability to perform this binaural streaming was taken away by binaural mismatches in pitch percepts introduced either via changes in pulse rate or electrode place, which destroyed the interaural correspondence in pitch percepts. Without those interaurally matching pitch percepts, listeners were unable to group sequential stimuli by proximity in pitch, forcing them to group by ear-of-input instead. Consequently, also in clinical practice an alignment of binaural percepts may be required to let patients build streams from binaural stimulation. This may be an important topic for further studies, which could also investigate the effect on outcomes for speech reception and music perception. Most current clinical CI stimulation strategies do not utilize the pulse rate to convey additional information. The present results suggest the participants were utilizing the pulse rate differences as a streaming cue, so that it could be worth to investigate its further adoption for clinical CI stimulation schemes, may it be to enhance perceptual differences across electrodes or as a cue for stream segregation.

5

Binaural streaming in bimodal cochlear implant patients

Abstract

A previous study confirmed the ability of bilateral cochlear implant (CI) patients to form streamed percepts from binaurally presented stimuli based on pitch cues while ignoring lateralization cues (chapter 4). It used a detection task paradigm based on the scale illusion (Janssen et al., 2019b). This raised the question whether bimodal CI patients, aided with a CI in one ear and a contralateral hearing-aid (HA), likewise can stream binaurally. For the HA-aided ear, complex tones were perceptually matched to the CI's stimulation on single electrodes, using the methodology described in (Lazard et al., 2012). Seven bimodal CI patients were asked to concentrate on a target stream and had to point out the one interval out of three in which a note deviated from a repeating pattern, while a distractor stream of randomized notes was presented simultaneously at a different pitch range. For both streams, every second note was presented to the opposite ear, so that binaural processing was required to group sequential stimuli by pitch-proximity, as necessary to follow the target stream and reliably detect the deviant. Five of seven bimodal listeners were only able to perform the task monaurally and none binaurally, despite rating the similarity of matched stimuli across ears high. This suggests that they did not form streams by pitch-proximity from binaural stimulation, as normal-hearing and bilateral CI listeners would.

5.1 Foreword

The study described in this chapter is again focussed on bimodal CI patients. After the bilateral CI patients' abilitiy for binaural streaming has successfully been demonstrated in the previous chapter, here the same method is adapted once more to assess whether bimodal CI patients, with their special combination of acoustically and electrically stimulated hearing, are also capable of binaural streaming.

5.2 Introduction

In everyday listening, such as when listening to another speaker, the human auditory system combines stimulation received by the two ears into common percepts through streaming. Auditory streams are formed by grouping (resp. segregation) of sounds based on perceptual properties such as pitch, timbre, spatial, or timing cues. However, hearing impairment and the devices used to aid patients with it can affect the perception of these cues, which often negatively affects the patients' abilities to segregate auditory streams (Oxenham, 2008; Rose and Moore, 2005). The question whether the devices affect the patients' streaming might be especially important for patients aided with a cochlear implant (CI) and a contralateral hearing aid (HA), referred to as bimodal CI patients, which we investigate in this study.
Auditory streams are formed by the mechanisms of sequential and simultaneous grouping, which are based on similarity in perceptual aspects (Bregman, 1990). Listeners can utilize a difference in any perceptual cue to segregate auditory objects (Moore and Gockel, 2002). As a result, simultaneously presented sounds can be grouped into auditory objects (e.g., the harmonics of a musical sounds are grouped into the percept of a single note) and such auditory objects can be linked into a stream by sequential grouping (such as a melody composed of a sequence of notes). However, these cues can conflict with each other, which may lead to an auditory illusion. The scale illusion described in Deutsch (1975) presents a famous example of such, presenting sounds that could be grouped either by pitch-proximity or by lateralization cue. This illusion can also be utilized to assess binaural streaming, as demonstrated for bilateral CI patients in 4. In the illusion, two melody streams were formed out of dichotic stimulation

based on grouping by pitch-proximity, one at a lower and one at a higher pitch-range (cf. fig. 4.1 in chapter 4). Both melodies ascended and descended in pitch over eight notes. These two streams were presented such that every other note per stream was delivered to the other ear. As one component of the higher stream was in one ear, a component of the lower stream was in the other ear, and vice versa. These stimuli could have been grouped either by pitch range or by ear-of-input by the listeners, but most listeners did group only by pitch range, combining notes presented from the left with notes presented to the right, and vice versa. The listeners' handedness appeared to influence their exact percepts: A higher-pitched stream coming from the right, accompanied by a lower-pitched stream from the left was reported by the majority of right-handed listeners and about half of the left-handed listeners ("both streams"). A higher-pitched stream and little to nothing of the lower-pitched stream was reported by a minority of the right-handed listeners and the other left-handed listeners ("single stream"). This illusion was only possible if the listener segregated simultaneously presented notes based on pitch or the lateralization cue (ear-of-input) and then grouped sequentially presented notes into two lateralized streams based on proximity in pitch. Hence, Deutsch argued that the Gestalt-principle of pitch-proximity overrode the lateralization cues. In this study, we aim to utilize this illusion to assess the bimodal CI patients' ability to build streams from binaural stimulation, as we did with bilateral CI patients in the previous chapter (4).
To enable testing of the illusion with bimodal CI listeners required several changes. First, CI patients often have only limited abilities to identify musical pitch-contours (Galvin et al., 2007, 2009a, 2009b; Zhu et al., 2011), so that a detection task may offer a better assessment than letting them describe their percepts. Hence a new paradigm developed and verified in Janssen et al. (2019b) was used, in which one of the two melody streams of the scale illusion was randomly chosen as a target stream into which a deviant note was embedded. Participants had the task to detect this deviant in a three-alternative, forced-choice paradigm. Parallel to the target stream, they were presented with notes of random pitch picked from the other pitch range of the scale illusion (fig. 5.2). As characteristic for the scale illusion, these two streams were presented such that every other note per stream was presented to the other ear. As one component of the higher stream was in one ear, a component of the lower stream was in the other ear, and vice versa. Hence, this arrangement required listeners

to a) segregate concurrent dichotic stimuli by pitch and b) group sequential stimuli by pitch-proximity, in order to perceive and follow the target stream. If listeners were able to follow this target stream, detecting the deviant note would be relatively easy.

Second, to allow the bimodal CI listeners to both segregate the simultaneous dichotic stimuli by pitch cues and group sequential stimuli by pitch-proximity, the pitch cues had to provide sufficiently large differences to allow for streaming to occur. A number of studies have shown limitations in CI patients' stream segregation abilities for stimuli with smaller differences in electrode place and pulse rate for streaming of concurrent stimuli (Carlyon et al., 2007; Cooper and Roberts, 2010; Deeks and Carlyon, 2004), as well as sequential stimuli (Chatterjee et al., 2006; Cooper and Roberts, 2007; Duran et al., 2012; Paredes-Gallardo et al., 2018a, 2018b). In addition to that, studies found that CI patients could also fuse simultaneous dichotic stimuli into one common pitch percept over a wide range in electrode-place pitch (Kan et al., 2013; Reiss et al., 2014a, 2017; Steel et al., 2015). Therefore, the two pitch ranges of the scale illusion stimuli were encoded using both electrode place pitch cues and different pulse rates across the two streams on the CI side. These place and temporal pitch cues could be integrated into a common pitch percept, as results from Luo et al. (2012) and Lamping et al. (2017) suggest, effectively enhancing the differences in pitch percepts.

Third, regarding the acoustical stimulation to the ear opposite to the CI, a pure-tone stimulation by the hearing aid may lead to a very different neural excitation pattern compared to that of a pulse train stimulation on a single electrode. So, to allow the bimodal CI listeners to group by perceptual similarity, a matching of the acoustic stimulation to the electric stimulation would need to be conducted. The matching procedure used in Lazard et al. (2012) was adapted for this experiment to match acoustic tone complexes presented via headphones to the stimuli from single CI electrodes. For these, the fundamental frequency was set equal to the matched electrode's pulse rate, since both represented temporal pitch cues. The participants were asked to adjust the center frequency, bandwidth and inharmonicity of the tone-complexes. Fourth, the limited residual hearing of the bimodal CI patients, often mild-to-moderate up to 500 Hz and changing to severe-to-profound above that, must be considered (Ching et al., 2007; Gifford et al., 2007). This restricts the choice of electrodes useable for interaural alignment to the ones located more apically.

We hypothesize that these modifications should allow bimodal CI patients to segregate concurrent streams and perceive the scale illusion, as with the bilateral CI patients in our previous study (chapter 4). However, the daily exposure to the different sound representations across ears in the bimodal configuration could have let to adaptations in the patients' perception and/or binaural processing (Gordon et al., 2017; Polonenko et al., 2019; Reiss et al., 2014b), so that they may not form the streamed percepts that lead to the scale illusion. Various forms of adaptations bimodal CI patients' pitch percepts over time have been reported in Reiss et al. (2014b), leading either to (1.) a reduction in interaural mismatch by adapting pitch percepts, (2.) converging of pitch percepts to a common low pitch percept across multiple electrodes, or (3.) static, non-adapting pitch percepts over time. This large individual variation in adaptation to bimodal stimulation also indicates a possible wide spread in streaming abilities.
Determining if bimodal CI patients form streamed percepts like bilateral CI patients and normal hearing listeners could be valuable information to guide optimization of the implants' and hearing aids' processing, stimulation scheme, and fitting parameters for the specific case of bimodal stimulation. It might help explain variability in outcome measures and could be a key to improve speech reception in noise and reduce listening effort.

5.3 Participants

This study encompassed seven bimodal CI patients (aged 40 to 79; average: 65.5 years), who became deaf post-lingually, including two male and two musically trained listeners. All participants were right-handed and all CIs were produced by Cochlear (Macquarie University, Australia). They were recruited at the German Hearing Center of the Medical University Hannover, Germany, and provided written informed consent. All experiments were approved by the ethical committee of the Medical University Hannover (reference 7885 BO S 2018). All participants were aided with CIs produced by Cochlear Ltd. (Macquarie University, NSW, Australia). More information on the participants can be found in tab. 5.1 and their average audiogram in fig. 5.1.

Table 5.1: Demographic information on the bimodal CI participants and their hearing aids (HA) with age and durations (Dur.) given in years (a)

	Etiology	Age	Dur. device use / a		Dur. hearing loss prior to aid / a		Dur. deafness prior to aid / a		Implant	
ID			CI	HA	CI	HA	CI	HA	Type	Ear
1	Middle ear infection	77	7.5	17.9	9.0	9.0	9.0	0	CI512	Left
2	Meningitis, progressive hearing loss	52	8.5	49.7	2.3	2.3	0	0	CI512	Right
3	Progressive hearing loss	63	12.8	32.9	6.0	6.0	0	0	CI24RE(CA)	Left
4	Progressive hearing loss	40	8.0	25.8	0	14.6	0	0	CI422	Right
5	Meniere's disease, progressive hearing loss	76	1.5	21.9	0	0	15.4	0	CI522	Right
6	Acute and progressive hearing loss	67	2.5	10.3	1.5	1.5	0	0	CI522	Right
7	Hereditary, progressive hearing loss	79	2.7	12.8	56.9	44.8	0	0	CI522	Right

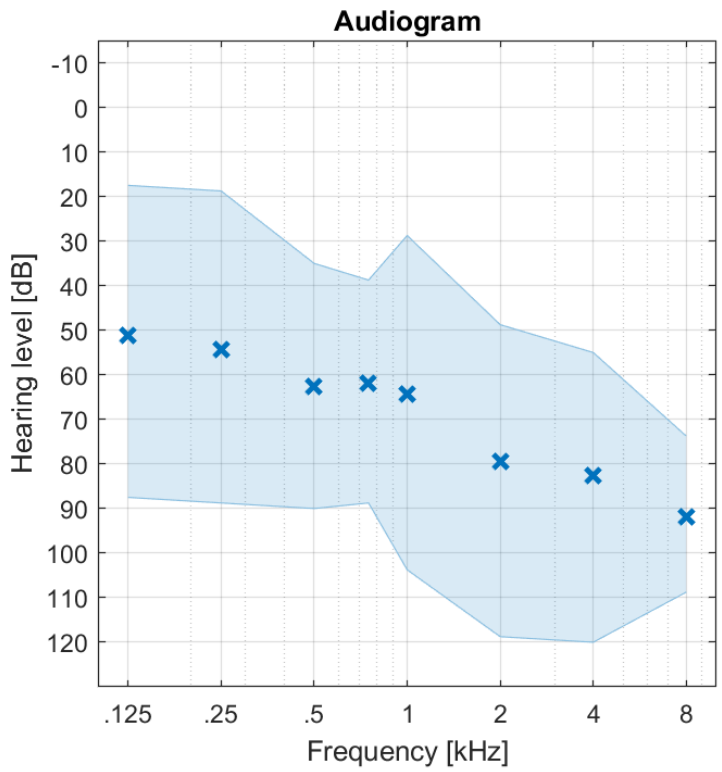


Figure 5.1: Average audiogram for the acoustically aided ear of the participants. The shaded area shows the maximum and minimum thresholds.

5.4 Method

5.4.1 Stimuli

For the ear with the CI, stimuli were 200 ms monopolar biphasic pulse-trains on eight single electrodes from the apical end of the CI, no. 22 to 15, with stimulation rates of either 70 pps or 250 pps, a pulse phase of 25 µs, and a pulse gap of 8 µs. The stimuli were delivered directly through a RF Gen XS research stimulation-platform (Cochlear Ltd., Macquarie University, NSW, Australia). For the acoustically aided ear, 200 ms-long bandpass-filtered acoustic tone-complexes with n = 50 harmonics were delivered via HDA 200 headphones (Sennheiser Electronik GmbH & Co. KG, Wedemark, Germany), driven by a Scarlett 2i2 soundcard (Focusrite Audio Engineering Ltd., High Wycombe, United Kingdom). A frequency-dependent amplification was applied according to the participants' audiogram and the half-gain rule. A fourth-order Butterworth filter was used for the filtering. The headphones were calibrated using a Brüel & Kjær

(B&K; Nærum, Denmark) 2636 sound-level meter, with IEC-60318-1 ear simulator B&K 4153 and reference calibrator B&K 4230 and equalized by their frequency response. The hearing threshold and most-comfortable listening level for each electrode were determined as in the clinical fitting, by the experimenter carefully increasing the current level from zero, while the participant could indicate the loudness on a seven-step scale (1: inaudible; 2: barely audible; 3: soft; 4: comfortable, but a little soft; 5: most-comfortable; 6: comfortable, but a little loud; 7: too-loud; in the participant's native language). Around the threshold and most-comfortable level, the loudness was adjusted in one current unit steps, equivalent to 0.168 dB. The sound processor was then programmed with the most-comfortable level found and threshold set to the same level, to ensure that stimulation occurs at the desired level. The loudness of the acoustic stimuli was adjusted in the same manner, using 1 dB steps around the most-comfortable level.

The acoustic stimuli were matched in percept to the single electrode stimulation using a modification of the procedure described in Lazard et al. (2012). The matching was conducted using the MAX/MSP 7 software (Cycling '74, Walnut, CA, USA), allowing for real-time modification of the stimuli. An acoustic equivalent was determined for each of the eight electrodes separately and in random order. During the matching, the electric stimulation was provided via a Freedom sound-processor (Cochlear Ltd., Macquarie University, NSW, Australia) connected via direct audio input, programmed so that only a single electrode was stimulated at most-comfortable listening level and a pulse rate of 250 pps, the lowest pulse rate selectable in the clinical programming software Custom Sound 5 (Cochlear Ltd., Macquarie University, NSW, Australia). Hence, the interaural matching could only be conducted at the higher of the two pulse rates used in the experiment, under the assumption that the pulse rate does not significantly affect the matching in other parameters, namely, the center frequency f_c of the filter, its bandwidth Δf, and inharmonicity iof the tone-complexes. The inharmonicity i determines the frequency-spacing of the complexes' components $f_n = f_0 * n^i$. The bandwidth was controlled via adjustments of the filter's Q-factor $Q = f_0/\Delta f$. The fundamental frequency f_0 of the tone complexes was fixed to either 70 Hz for the more apical four electrodes or 250 Hz for the other four electrodes, to match the pulse rate used during the later experiment. Therefore, the temporal pitch percept should match both acoustically and electrically, where temporal pitch differences are generally only perceived up to a pulse rate of about 300 pps (Vandali et al., 2013; Zeng, 2002). In the matching procedure, electric and acoustic stimuli were constantly presented in turn, while participants could vary the parameters of the acoustic tone-complex using a graphical interface Bamboo Fun pen and touch (Wacom Co. Ltd., Kazo, Saitama, Japan). This allowed them to control two parameters at once by moving a stylus on a two-dimensional surface. Hence, the adjustment for each electrode's equivalent was conducted in two steps. The participants were encouraged to explore the entire mapping space during the adjustment before deciding on the optimum point.

First, the abscissa was mapped to control the center frequency (ranging from 40 to 22050 Hz on a logarithmic scale) and the ordinate to control the Q-factor (ranging from 0.15 to 300 on a logarithmic scale), while the inharmonicity was kept at one, i.e., producing a harmonic sound. In case the center frequency of the filter fell below the fundamental frequency, the filter acted as a low-pass and, likewise, if the center frequency fell above the highest harmonic ($50 * f_0$), the filter acted as a high-pass. This limited the number of audible harmonics. The participants were instructed to report eventual changes in loudness during the adjustments, in which case the loudness was re-matched. In the second step, the parameters adjusted in the first step were kept constant and the abscissa was mapped to control the inharmonicity instead, leaving the ordinate without function. Once the participants were satisfied with the equivalence of the two stimuli, they were asked to rate this equivalence on a scale from zero (for no similarity at all) to ten (for a perfect match). The adjustment was repeated at least once for all stimuli and the result with the better rating was used in the further experiment.

Next, the loudness-balancing was repeated for the stimuli to be used in the further experiment, i.e., with the parameters determined in the matching, as generated by the hard- and software for the experiment, which was programmed in Matlab R2015b (The Mathworks Inc., Natick, MA, USA). The triggering feature of the research platform was used to ensure simultaneous binaural stimulation.

5.4.2 Procedure

The patients' ability to stream binaurally by pitch-proximity was investigated using the procedure described in Janssen et al. (2019b). Building upon the auditory scale illusion described by Deutsch (1975), two melodies were presented concurrently using the eight binaurally aligned stimuli. The participants' task was to concentrate on a target melody stream, which consisted of eight notes repeated in a pattern going up and down in pitch, similar to the melodies in the original scale illusion experiment. This target melody stream was presented either at a lower or a higher pitch-range, together with a concurrent distractor melody stream with notes picked randomly from the other pitch range. Hence, when the target stream was presented at the lower pitch-range, the distractor stream was presented at the higher pitch-range, and vice versa (cf. fig. 5.2). Moreover, the components of these two streams were presented simultaneously, so that if a component from the target stream was on the right, a component of the distractor stream was on the left, and, again, vice versa. Consecutive components for both streams were presented contralaterally, in a pattern that reversed in order after every four notes. So, the arrangement formed a pattern symmetric to the middle of the repeating sequence of eight notes (e.g., for one stream in one presentation of the eight notes: left-right-left-right-right-left-right-left).

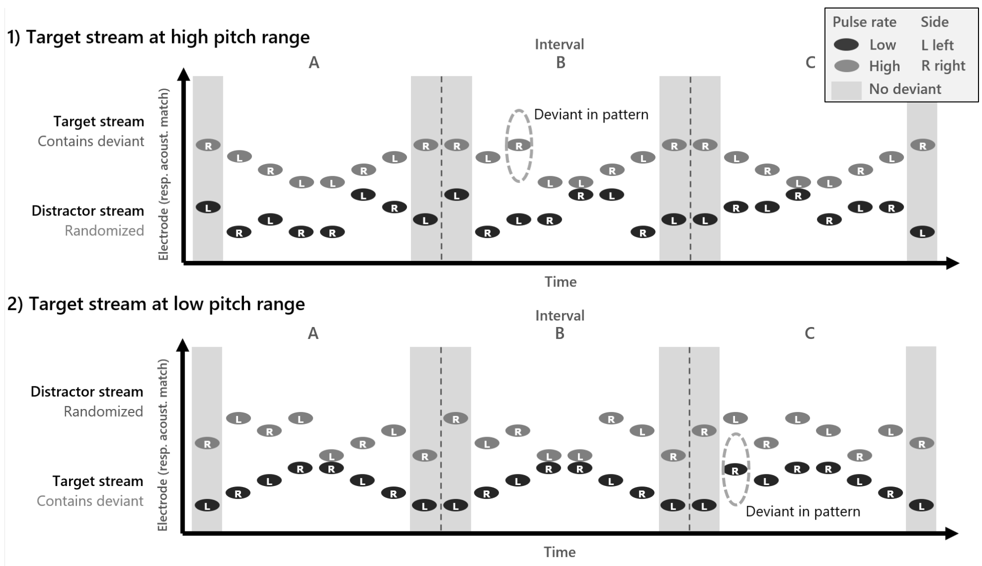


Figure 5.2: Schematic stimuli representation. A random deviant occurs in the target stream, either at the higher (top) or lower (bottom) pitch range. The distractor stream consists of notes picked randomly from the other pitch range. The black (low) and grey (high) of the ovals indicates the pulse rate, resp. fundamental frequency for the tone complexes with which the acoustically aided ear was stimulated. These tone complexes were matched in percept to the corresponding electrode in the other ear.

In case the bimodal CI listeners showed the same streaming behavior as the bilateral CI listeners in chapter 4, they should, likewise, be able to concentrate on and track the target stream. To perceive it, they must segregate the simultaneous dichotic components of the stimuli and group sequential components by proximity in pitch, ignoring the localization cues from the ear-of-input. The test to determine whether the listener could follow the target stream consisted of a deviant note in the target stream, falling out of the regular pattern (but within the regular pitches of the target stream). The deviant was introduced by shifting a note's respective pitch, i.e., electrode, or its acoustic equivalent, by two positions within the respective pitch range's four pitches (as indicated in interval C of the bottom drawing in fig. 5.2). The detection task cannot reveal if they perceive the concurrent distractor stream on the other side along with the target or not ("both streams" or "single stream" percept, as described in Deutsch, 1975). However, either percept requires that listeners perform the pitch-based grouping from binaural stimulation. In case listeners grouped by pitch-proximity, they would show a high detection performance. If listeners were grouping by ear-of-input instead of pitch-based, evaluating both ears individually, the random nature of the distractor stream's components, i.e., every second note, would make the detection of a deviant more difficult.

The target stream was indicated at the beginning of each trial by two exclusive presentations without the distractor stream. Following, both target and distractor streams were presented concurrently six times. During the first three repetitions, a build-up of streamed percept could occur, as the formation of segregated percepts has been described to form gradually over several seconds (Bregman, 1990; Paredes-Gallardo et al., 2018b, 2018c). Three intervals for the detection task are formed by the last three repetitions, of which one contains the deviant note, as indicated for interval C at the bottom of fig. 5.2. During playback, the current interval was lit up on a graphical user interface on the screen. Listeners could repeat the stimuli when desired.

The positioning of the deviant note occurred randomly at any of the inner six of the eight notes within one of the three last repetitions of the target stream, but never the first or last of the eight notes. The deviant was placed counterbalanced with regard to the side of presentation, interval, and pitch range of the target stream. As a three-alternative, forced-choice task, the task was free of a detection-bias (Wickens, 2002).

A training was conducted before the experiment, which consisted only of the binaural and monaural test conditions (see Conditions below). During the training, the participants were given feedback on whether they answered correctly. The overall duration of the experiment amounted to 4 to 6 hours, which included breaks. The experiment was, therefore, split into two separate sessions to reduce fatigue. Loudness balancing and interaural matching were conducted in the first session, which concluded in a basic test to determine if participants were able to perform the task with only the target melody presented in the paradigm. Here, twelve trials presented the listener with the target stream alone and in the same counterbalanced for as during the experiment. Participants who were unable to reach a performance level of at least 75 % correct in this simple task were excluded from further testing, since in that case no clear conclusions regarding binaural streaming could have been drawn. The training and final testing were performed in the second session.

5.4.3 Conditions

The experiment featured four conditions with twenty-four trials each, which are described below. All conditions together were presented in individually randomized order to prevent the order from affecting the results.
Both target and distractor stream were presented in the same way as the original scale illusion in the binaural test condition, i.e., every second note from each stream was presented to the other ear (fig. 5.3). Hence, perception of the target stream and the reliable detection of the deviant required listeners to a) segregate concurrent dichotic stimuli based on pitch and b) group sequential stimuli by proximity in pitch, while ignoring the lateralization cues from the ear of input.
Conversely, both streams were presented to the same ear in a monaural test condition (fig. 5.4). This allowed assessing the listeners' capability to segregate the two concurrent streams when no binaural processes were required and to evaluate the influence of the binaural processing when compared to the binaural test condition. In this condition, listeners were always expected to group the stimulation based on pitch-proximity and detect the deviant. The performance in this condition demonstrates the best-possible performance without binaural processes and show if the listeners were capable of forming separate streams from concurrent stimulation. Thus, listeners unable to reach a performance significantly different from chance in this condition, and therefore unable to even segregate the streams monaurally, were excluded from the evaluation regarding binaural streaming. Based on the bimodal distribution for twelve

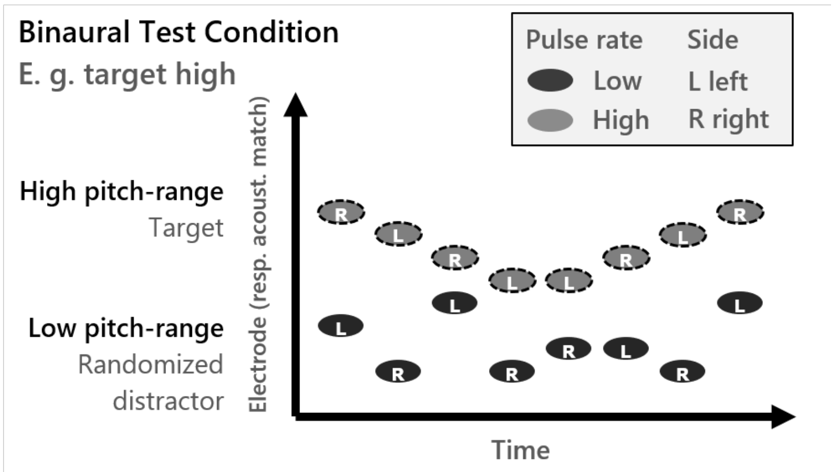


Figure 5.3: Schematic stimuli representation for the binaural test condition on the example of the target stream (marked with dashed lines around the notes) at the higher pitch range with the randomized distractor stream at the lower pitch range. The black (low) and grey (high) of the ovals indicates the pulse rate. Stimuli on the ear stimulated acoustically were acoustic tone complexes, matched in percept to stimulation by the corresponding electrode in the opposite ear.

trials per pitch range and using the Bonferroni correction for two comparisons at a significance level under five percent, the required performance-criterion amounts to 66.67 % correct answers ($p \leq 0.0188$).

In addition, two binaural control conditions were designed to force listeners to perform the task without grouping of stimuli across ears by pitch-proximity, but by ear-of-input instead. In these, the target and distractor streams were again presented binaurally as in the binaural test condition. This allowed to contrast the performance in the test conditions and to quantify the possible detection performance in case that no binaural grouping was possible. But even without such binaural grouping, a theoretical optimal detector would yield a result of 66.67 % correct answers, compared to 100 % correct in the test conditions. In both control conditions, the interaural correspondence in percepts was destroyed so that no interaural grouping of stimuli by similarity could occur, leaving the listeners to evaluate the stimuli by ear-of-input. This was realized by modifying the stimuli on the CI-side, either by exchanging the pulse rate across streams, maintaining the electrode's places (fig. 5.5), or by inverting the order of electrodes while maintaining the pulse rates (fig. 5.6). The listeners, evaluating the stimuli by ear-of-input in both control conditions, would encounter stimuli where every second note came from the randomized

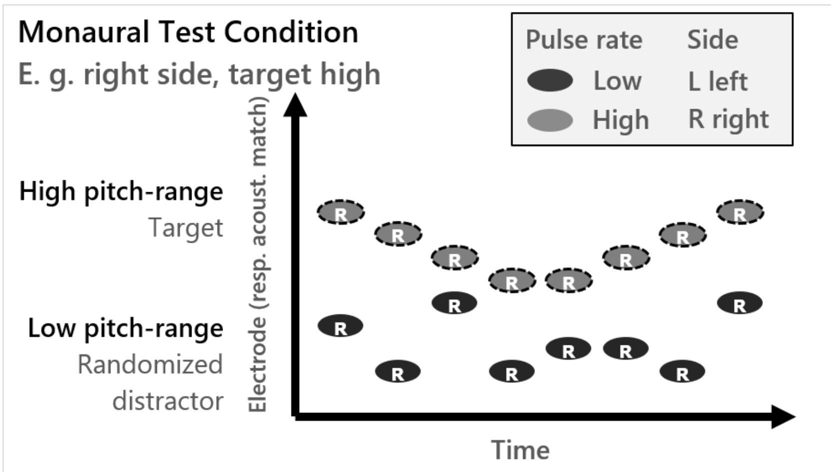


Figure 5.4: Schematic stimuli representation for the monaural test condition on the example of the target stream (marked with dashed lines around the notes) at the higher pitch range with the randomized distractor stream at the lower pitch range, both presented to the same ear, e.g., the right. The black (low) and grey (high) of the ovals indicates the pulse rate. Stimuli on the ear stimulated acoustically were acoustic tone complexes, matched in percept to stimulation by the corresponding electrode in the opposite ear.

distractor stream, making the detection of the deviant more difficult. As in the results of the bilateral CI patients (chapter 4), a low performance was expected in the control conditions.

5.5 Results

5.5.1 Results of the binaural matching

The results of the binaural matching to stimulation by the eight individual CI electrodes are depicted in fig. 5.7 for the filter center frequency, fig. 5.8 for the Q-factor of the filter, and fig. 5.9 for the inharmonicity. The parameter values are plotted both as the average across participants in black circles with bars for the standard error, as well as in shaded grey plusses with dotted lines for the individual participants.

Differences in the matched parameters were observed especially between the four more apical electrodes with a fundamental frequency of 70 Hz (denoted 1 to 4 here, referring to electrodes 22 to 19 as named by Cochlear) compared to the four following electrodes with a fundamental frequency of 250 Hz (denoted 5 through 8, referring to electrodes 18 to 15 as named by Cochlear). For the

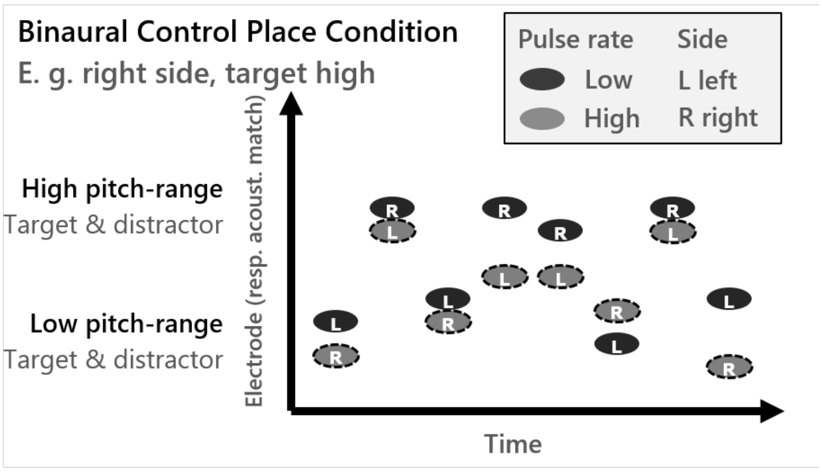


Figure 5.5: Schematic stimuli representation for the binaural control condition with mismatched pulse-rate pitch cues across both streams for one ear. In this example, the target stream (marked with dashed lines around the notes) is at the higher (electrode place) pitch range with the randomized distractor stream at the lower (electrode place) pitch range. The black (low) and grey (high) of the ovals indicates the pulse rate, which is interchanged for the right ear in this example. Stimuli on the ear stimulated acoustically were acoustic tone complexes, matched in percept to stimulation by the corresponding electrode in the opposite ear.

more apical electrodes, the average filter center frequency does only increase slightly with electrode place, matched to values within about 200 Hz and 600 Hz, whereas an increase from this range is visible for all but one participant for the other four electrodes.

Possible relations between the electrode number (from 1 to 8 with increasing distance from the apex) as the independent variable and the matching parameters as the dependent variables were explored using Pearson correlations. Overall, the filter center frequency shows a highly significant correlation to the electrode ($r = 0.606$, $p < 0.0001$). The filter's Q-factor does not change substantially over the electrodes and its correlation to the electrode is non-significant ($r = -0.0841$, $p = 0.509$). Lastly, for the four more apical electrodes the inharmonicity shows more individual variation between about 1.0 and 2.4 with a rather constant average value of about 1.5. For the other four electrodes, the inharmonicity drops to both lower values and lower individual variation between about 0.7 and 1.6 with an average of about 1.2 (i.e., more harmonic). The inharmonicity was also significantly correlated to the electrode ($r = -0.433$, $p < 0.0001$).

The participants rated how satisfied they were with the resulting correspondence of acoustic and electric stimulation on a scale from zero for no correspondence at all to ten for perfectly matching sounds. The resulting ratings can be found in

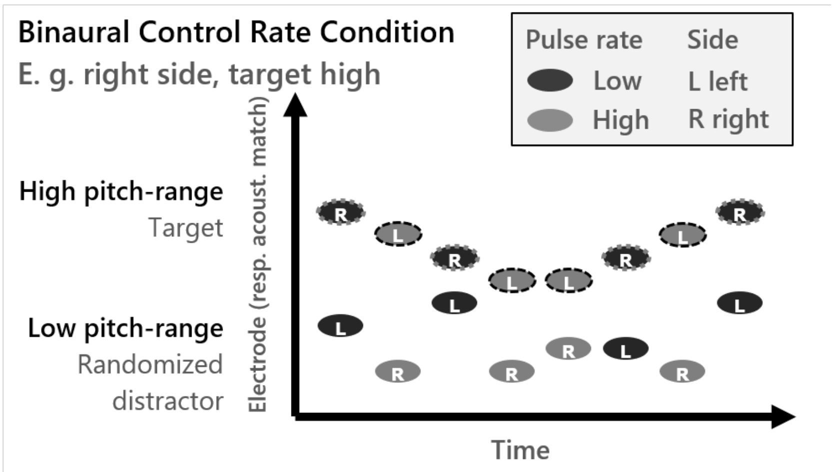


Figure 5.6: Schematic stimuli representation for the binaural control condition with mismatched electrode-place pitch cues across ears, e.g., with inverted electrode place cues for the right ear. In this example, the target stream (marked with dashed lines around the notes) is at the higher (pulse rate) pitch range with the randomized distractor stream at the lower (pulse rate) pitch range. The black (low) and grey (high) of the ovals indicates the pulse rate. Stimuli on the ear stimulated acoustically were acoustic tone complexes, matched in percept to stimulation by the respective electrode on the opposite ear.

tab. 5.2 and indicate that the participants were rather satisfied with the resulting acoustic stimuli, with an average rating of 8.4 ± 0.80.

5.5.2 Results of the detection task

All participants were able to perform the task with just the target melody alone at a performance of 75 % correct or higher. Fig. 5.10 shows the performance achieved with the CI compared to the HA in the monaural test condition. The performance is plotted separately for the low (“low”) and the high pitch range (“high”) for the HA-aided side and the CI-aided side. For both pitch ranges, the CI shows slightly higher performance, with about 47 % vs. 45 % difference in the low pitch range and about 57 % vs. 67 % in the high pitch range across acoustic and electric stimulation, respectively.
For this monaural data comparing the listeners’ performance in the two kinds of stimulation, a two-way repeated analysis of variance was performed, using the “rationalized” arcsine transform [rau; Studebaker, 1985], applied to the detection scores as the dependent variable, and the kind of stimulation, as well as the target stream’s pitch range as factors. The results demonstrate that neither

Table 5.2: Ratings for the participants' satisfaction with the correspondence of electric and acoustic stimulation after the interaural matching on a scale from zero for no correspondence at all to ten for perfectly matching sounds. Along with this, the number of repetitions for the matching procedure (Rep.), as well as the average rating for the two pitch ranges and overall (tot.) are given.

	Correspondence satisfaction rating per electrode									Average			
ID	El.	1	2	3	4	5	6	7	8	Rep.	Low	High	Tot.
1		6.5	6.5	7.0	7.0	7.0	7.0	8.0	7.0	1-2	6.7	7.2	7.0
2		10.0	8.0	9.0	10.0	9.0	9.0	10.0	10.0	1-2	9.2	9.5	9.3
3		9.0	8.0	9.0	8.0	9.0	9.0	9.0	9.0	1-3	8.5	9.0	8.7
4		9.0	9.5	9.5	9.5	10.0	10.0	9.0	9.5	1-2	9.4	9.6	9.5
5		8.0	8.0	8.0	8.0	8.0	8.0	8.0	8.0	2	8.0	8.0	8.0
6		8.5	8.5	8.0	8.0	9.0	9.0	9.0	9.0	1-2	8.2	9.0	8.6
7		9.0	9.0	8.0	7.0	8.0	8.0	8.0	8.0	2	8.2	8.0	8.1

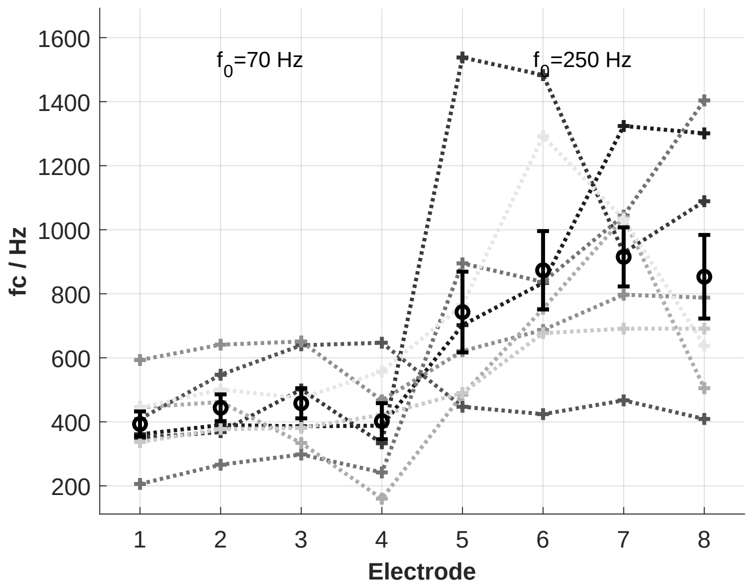


Figure 5.7: Results of the interaural matching for the filter center frequency f_c plotted over the electrodes from the most apical, denoted 1, to the least apical, denoted 8 (referring to el. 22 to 15 as named by Cochlear). For the four more apical electrodes a fundamental frequency of 70 Hz was used, while a fundamental frequency of 250 Hz was used for the other four electrodes. The average results are plotted in black circles with the standard error, while greyscale plusses with dotted lines denote the individual data.

the kind of stimulation ($F(2, 6) = 0.62$, $p = 0.461$), the pitch range ($F(1, 6) = 1.26$, $p = 0.304$), nor their interaction ($F(1, 6) = 3.56$, $p = 0.847$) were significant factors. Hence, the results were evaluated averaging over both devices.

Fig. 5.11 shows the detection performance for the participants individually, plotted in percent of correct answers for the binaural and monaural test conditions, as well as the binaural control conditions with mismatches in pulse rates and electrodes across ears. There are two data points for each condition, one indicating the average performance in the lower pitch range ("l") and the other in the higher pitch range ("h"). The bold type at the top right of every individual plot indicates the participant ID. Regarding the criterion for a performance significantly different from chance of 66.67 % in the monaural test condition, five of the seven participants (no. 1, 3, 5, 6, and 7) reached or surpassed this, but only when the target stream was in the high pitch range. The performance in other the binaural test condition and the binaural control conditions is lower for these five listeners as well as the remaining two participants (no. 2 and 4).

The performance on the group level for the listeners who scored significantly

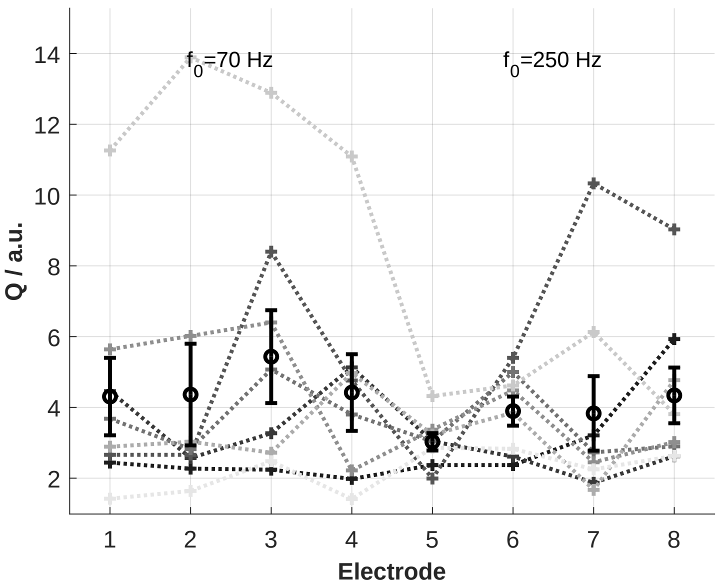


Figure 5.8: Results of the interaural matching for the filter Q-factor plotted over the electrodes from the most apical, denoted 1, to the least apical, denoted 8 (referring to el. 22 to 15 as named by Cochlear). For the four more apical electrodes a fundamental frequency of 70 Hz was used, while a fundamental frequency of 250 Hz was used for the other four electrodes. The average results are plotted in black circles with the standard error, while greyscale plusses with dotted lines denote the individual data.

above chance in the monaural test condition is given in fig. 5.12, plotted in percent of correct answers for the binaural and monaural test conditions, as well as the binaural control conditions with mismatches in pulse rates and electrodes across ears. As before, there are two data points for each condition, one indicating the average performance in the lower pitch range ("low") and the other in the higher pitch range ("high"). The monaural test condition shows the highest performance, with about 75 % on average in the high pitch range, but only about 58 % in the low pitch range. The latter result is only marginally better compared to the scores in the binaural test condition with around 42 % in both pitch ranges, as well as the binaural control conditions with mismatched pulse rate around 33 % and mismatched electrode place around 38 %.

A two-way repeated analysis of variance was performed for the data from these five listeners, again with the "rationalized" arcsine transform applied to the detection scores (rau), which were the dependent variable, and the conditions, as well as the target stream's pitch range as factors. The condition was a significant factor ($F(3, 12) = 7.64$, $p = 0.00410$), but not the pitch range ($F(1, 12) = 3.01$,

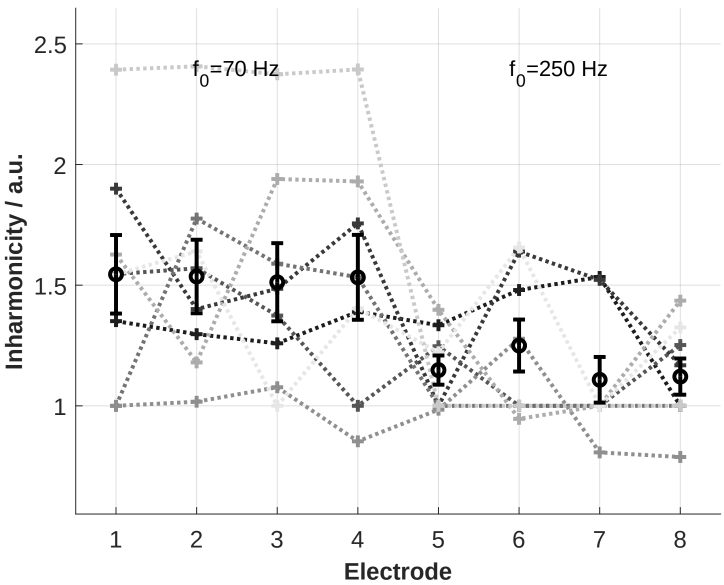


Figure 5.9: Results of the interaural matching for the inharmonicity plotted over the electrodes from the most apical, denoted 1, to the least apical, denoted 8 (referring to el. 22 to 15 as named by Cochlear). For the four more apical electrodes a fundamental frequency of 70 Hz was used, while a fundamental frequency of 250 Hz was used for the other four electrodes. The average results are plotted in black circles with the standard error, while greyscale plusses with dotted lines denote the individual data.

$p = 0.158$). The interaction between condition and pitch range was significant ($F(3, 12) = 4.61$, $p = 0.0228$). Following this, a post-hoc analysis was conducted using the rau-transformed performance data of the five listeners. The paired T-tests were evaluated using the false discovery rate control (FDRC) method, based on a significance criterium of $\alpha = 0.05$ (Benjamini and Hochberg, 1995). Although the p-value was rather low, the average scores in the binaural and monaural test conditions overall, i.e., including both pitch ranges, were not significantly different from each other when the correction was considered ($p = 0.0309$; FDRC-criterion: $p = 0.0167$). The differences between the binaural test condition and the binaural controls with mismatched pulse rate ($p = 0.206$) and mismatched electrodes ($p = 0.641$) were also non-significant. Taking into account the significant interaction of condition and pitch range and the noticeably higher performance in the high pitch range in the monaural test condition, also the differences across high and low pitch ranges were assessed. The differences across pitch ranges within the binaural ($p = 0.178$) and monaural test conditions ($p = 0.0128$), were both non-significant (FDRC-criterion: $p = 0.0125$).

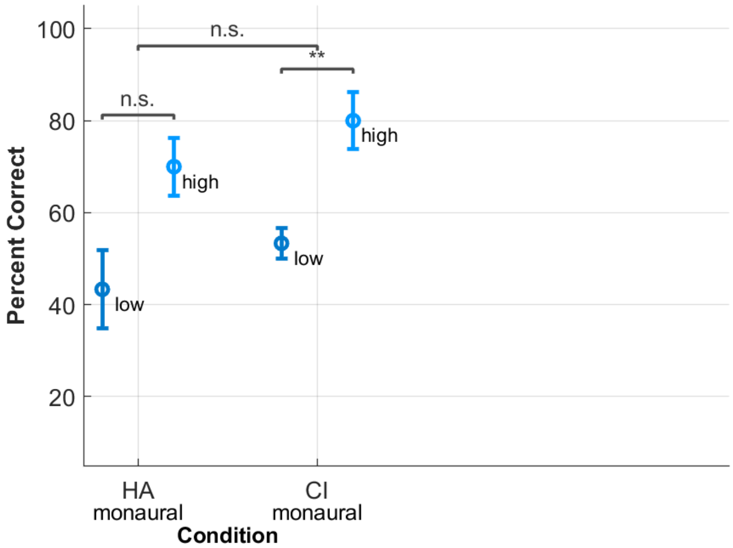


Figure 5.10: Averaged monaural results from the bimodal CI listeners in percent of correct answers separated by kind of stimulation (HA: acoustic; CI: electric). In each condition, the average performance is plotted for the target stream in the low ("low"), and high ("high") pitch ranges with the standard error and indication of significant differences.

Neither significant was the difference across binaural and monaural test condition for the low pitch range ($p = 0.225$), but the difference across these two conditions was significant for the high pitch range ($p = 0.0134$; FDRC-criterion: $p = 0.0167$).

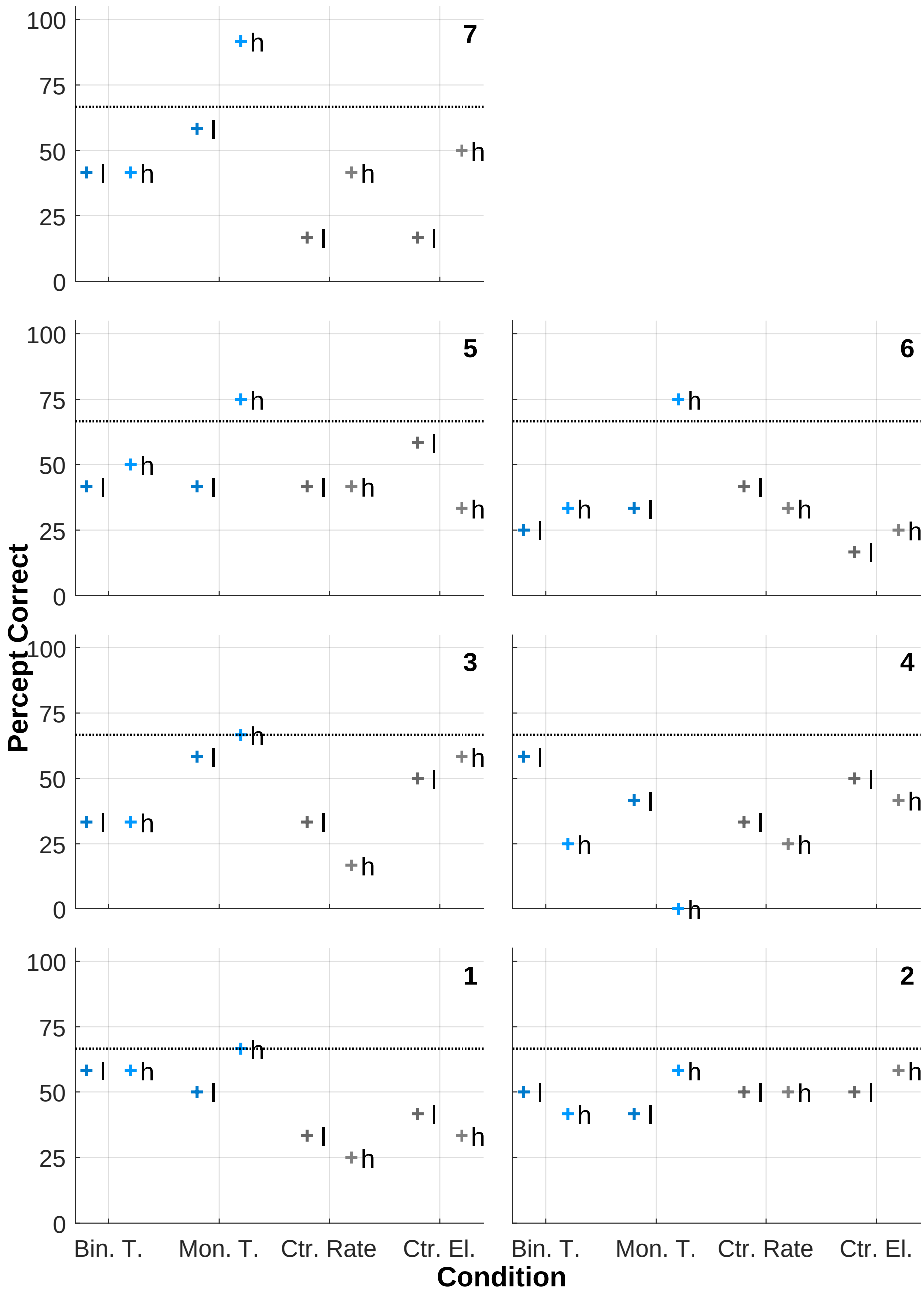


Figure 5.11: Individual results from the seven bimodal CI listeners in percent of correct answers for the four conditions (denoted “Bin. T.”: binaural test condition, “Mon. T.”: monaural test condition, “Ctr. Rate”: binaural control condition with mismatched pulse rate, and “Ctr. El.”: binaural control condition with mismatched electrodes) with one plot per participant. A bold number at the top right of every plot indicates the participant ID. In each condition, the performance is plotted for the target stream in the low (“l”), and high (“h”) pitch ranges (color online). The dotted line indicates a performance of 66.67 % correct, the criterion used to assess the individual performance in the monaural test condition.

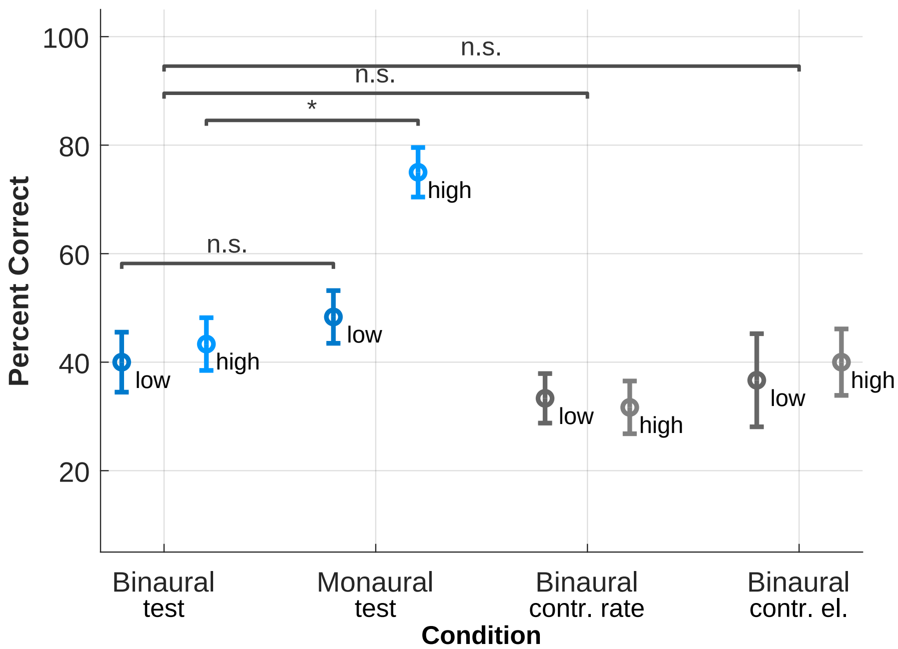


Figure 5.12: Averaged results from the five bimodal CI listeners who scored significantly above chance in the monaural test condition in percent of correct answers for the four conditions. In each condition, the average performance is plotted for the target stream in the low ("low"), and high ("high") pitch ranges with the standard error and indication of significant differences

5.6 Discussion

The task for the bimodal CI participants of this study was to focus on a target stream in the presence of the randomized distractor and determine the interval in which a deviant note occurred. The binaural test condition required binaural processing a) for segregation of simultaneous dichotic stimuli based on pitch, possibly aided by the lateralization cues from the ear-of-input, and b) grouping of sequentially presented stimuli by interaural pitch-proximity, while ignoring lateralization cues. While all bimodal CI listeners could perform the task with just the target melody alone, none of them scored significantly above chance in the binaural test condition and only five of the seven listeners were able to score significantly above chance level in the higher pitch-range of the monaural test condition, where both target and distractor were presented to the same ear. The binaural controls with mismatches in pulse rate and electrode place showed scores not significantly different from the binaural test condition, as visible both individually and for these five listeners as a group. This suggests that the bimodal listeners were unable to perform the binaural streaming required in the binaural test condition. However, the high performance in the higher pitch-range of the monaural test condition indicates that they were able to segregate concurrent streams based on the temporal and place pitch cues provided via pulse rate and electrode place resp. fundamental frequency for the acoustically stimulated ear. Since the listeners were unable to stream binaurally with interaurally aligned percepts in the binaural test condition, such streaming was also impossible in the binaural control conditions. The other two listeners did not score significantly above chance in any condition, which suggests that they cannot segregate the concurrent streams monaurally, foreclosing a study of their binaural streaming behavior.

The lower performance shown in the lower pitch-range, could at least in part explained by a generally higher difficulty to focus on the stream presented at a lower pulse rate resp. pitch, since the bilateral CI listeners in chapter 4 showed the same trend. Likewise, a study of concurrent streaming using CI simulations for normal-hearing listeners found that presenting one target stream at a higher rate and a concurrent masker at a lower rate let listeners follow the target, but interchanging the rates appeared to have prevented the listeners from following it (Deeks and Carlyon, 2004). Musically-trained normal-hearing listeners, who had to identify dichotically presented tonal sequences in Deutsch (1985), also performed consistently better in the higher of two melodies. There appears, thus, to be a general tendency that segregating a stream of higher pitch (or at a higher pulse-rate) is easier, which then likely factors into the lower performance observed in the lower pitch-range. The participants' on average stronger high-frequency hearing loss, even though compensated for via the half-gain rule, could have led to a lower performance in the high pitch range, but not a higher one. As for bilateral CI listeners, the experiment design could be optimized regarding this by always presenting the target stream at a higher pitch-range, making use of the pulse rate resp. fundamental frequency to ensure a higher pitch for the electrodes desired to present the target stream.

Remarkably, the performance differences in the monaural test condition were not significant across the type of stimulation, i.e., the participants scored slightly lower on the HA-aided side, but not significantly so. This indicates that the salience of the pitch cues in the acoustic stimuli were sufficient to allow for the melodies to be encoded, even though some of the matching parameters only changed moderately over the corresponding electrodes. Still, the trend towards a slightly lower acoustical performance could lead to a significant difference across devices with a larger number of participants. The lower salience of the cues could be explained by the reduced frequency selectivity in the impaired auditory system (Florentine et al., 1980).

Before subjected to the streaming tests, the bimodal listeners first had to match acoustic tone-complexes to the electric stimulation of eight individual electrodes from the apical side of their CI. Overall, the participants were rather satisfied with how well the resulting acoustic stimuli matched the sensation from the electric stimulation. The results for the matching parameters indicate that the filter's Q-factor was not varying significantly with the electrode location. Thus, it would perhaps only have to have been matched once per participant. This presents a way to optimize the interaural matching procedure. The other two parameters matched varied significantly over the electrodes, with the filter's center frequency increasing from around 400 Hz to around 850 Hz and the inharmonicity decreasing from very inharmonic towards more harmonic sounds with increasing distance from the most apically located electrode. The eight consecutive electrodes used in this experiment spanned about 6 mm along the extend of the electrode array.

The participants adjusted the matching parameters to more similar values for the sounds with a fundamental frequency of 70 Hz (paired to the more apical

four electrodes) and the sounds with a fundamental frequency of 250 Hz (paired to the other four electrodes). This could be explained by the different fundamental frequency used across the two ranges, which the participants could not adjust themselves. If this parameter had been accessible to the participants, they might have increased the fundamental frequency with increasing distance from the most apical electrode, along with the filter's center frequency. Although the participants were rather satisfied with the matching results, letting them also adjust the fundamental frequency could have led to an even better correspondence. Another noteworthy factor is that the stimulation via direct audio input to a CI processor, as used in the matching, did only allow to stimulate with certain pulse rates, of which 250 pps was chosen as the lowest possible. Hence, an exact correspondence between matched stimuli and the stimuli used during the experiment was only given for the higher pitch-range. This may have contributed to the significantly better performance of the listeners in that pitch range, visible in the monaural test condition. This may also reflect in the participants' satisfaction ratings, since those in the higher pitch-range were slightly better at 8.6 compared to 8.3 for the lower pitch-range and a paired T-test showed that the difference was close to significant ($p = 0.0533$).
The difference in streaming behavior compared to the bilateral CI listeners in chapter 4 could be explained based on the bimodal patients' daily experience with HAs and CIs, representing sounds in fundamentally different ways, i.e., without binaurally-matched percepts. As Reiss et al. (2014b) found, such daily "training" can affect the auditory system when it comes to pitch percepts. This could also lead to adaptations in the binaural processing of sounds, which may prevent the bimodal listeners from integrating the binaural stimuli into common streamed percepts like normal-hearing listeners and bilateral CI patients would (Gordon et al., 2017; Polonenko et al., 2019).
Further studies could attempt to find if stimuli inspired by the way the devices commonly represent sounds, i.e., based on the common frequency-mapping for the devices' signal processing, lead to binaurally streamed percepts instead, if the bimodal listeners do only form streams from binaural stimulation in a limited spectral range, or not at all. A survey of 38 bimodally-aided CI patients (chapter 2) suggests that only a minority perceiveed sounds from the same source as integrated into one uniform sound object, whereas the majority perceived notable differences in pitch and loudness in a sound's represenation that may lead to split percepts. Some patients even reported two entirely separate

percepts. These results thus fit in with the observed lack of binaural streaming in the bimodal participants, as binaural streaming equally would require them to form common percepts from binaural stimulation. Hence, it appears unlikely that bimodal CI patients generally can form such common percepts from binaural stimulation. This does not necessarily mean that patients must be unable to utilize cues from either device in combination, as studies have shown that bimodal patients can utilize fundamental frequency cues only available from the HA, together with the stimulation from the CI, which fundamentally enables them to understand speech, e.g., in Kong et al. (2005). However, the changes in binaural processing might hamper the efficiency of such combination and, therefore, lead to higher listening effort, limit the patients' abilities for sound localization and speech reception in noise, and factor into the variability commonly seen with CI listeners. The patient's ability to effectively utilize both ears together with the bimodal aids, including binaural streaming, may be an essential question to consider when discussing if a patient should receive a second CI.

In further research, one could assess if patients show better results regarding binaural streaming with stimuli oriented towards the perceptually different clinical acoustic stimulation and clarify how a lack of binaural streaming influences the patients' listening performance. Current results suggest that the prolonged experience of binaurally-unmatched bimodal stimulation changes the patients' binaural processing. Studies with bimodal listeners, carried out within the first six months after implantation, could track eventual adaptations that prevent perceptually-similar binaural stimuli from being integrated into a common streamed percept.

A possible explanation for the low performance in the binaural test condition could be an abnormal integration of simultaneous binaural stimuli as that reported for short dichotic stimuli in other studies (Kan et al., 2013, Reiss et al., 2014a, Steel et al., 2015, cf. chapters 1 and 4). Such wide integration could let listeners perceive the binaural stimuli in the current task as one common stream, instead of segregating them into two concurrent streams by pitch-proximity of the sequential stimuli. Such a single fused stream would consist of random pitches, allowing only for chance level performance. Compared to the monaural test condition, this would present a fundamental change of streaming behavior, and signal for the binaural auditory system that binaural stimulation could act as an inhibitor to streaming, rather than aiding listening. In such a case, bilateral

aids of the same kind, HA or CI, would be highly preferable over a bimodal CI configuration.
Despite no indication of binaural streaming was found, the present findings are encouraging regarding previous studies of stream segregation in CI listeners, as most listeners appear able to segregate streams by pitch-proximity monaurally. Studies of sequential monaural stream segregation by Chatterjee et al. (2006), as well as Cooper and Roberts (2007) had found that a larger separation in electrode place across streams was necessary to segregate sequential stimuli into separate streams, at least for most listeners. In the current study, the streams were presented with eight consecutive electrodes, which by itself may prevent many CI patients from segregating the two streams. However, with the use of different pulse rates across lower and higher pitch range, the perceptual separation appears to have been sufficiently high to allow for streaming by pitch-proximity, even of the concurrently presented stimuli. Other studies regarding concurrent streaming in CI listeners had concluded that the patients were unlikely to utilize temporal pitch cues, i.e., differences in pulse rate, to segregate streams (Carlyon et al., 2007; Deeks and Carlyon, 2004). While not directly obvious from the present results, the results in the binaural control with mismatched pulse rate across ears for the bilateral CI patients in chapter 4 clearly demonstrated the effect of the pulse rate on the listeners' stream segregation. Together with the other previous studies of streaming, this suggests that also the majority of the bimodal CI participants were able to utilize the enhanced perceptual separation provided by the temporal pitch cues on top of the place pitch cues to segregate the streams by pitch-proximity in the monaural test condition. Hence, using the pulse rate to enhance pitch cues and/or provide an additional cue for stream segregation could prove useful for future CI stimulation strategies.

5.7 Conclusion

This study was aimed at investigating whether bimodal CI patients were able to form streams from binaural stimuli so that they were required to integrate input to both ears to form common streams. The bimodal participants were unable to perform the detection task when the stimuli were presented in the characteristic binaural pattern that evoked the scale illusion in normal-hearing and bilateral CI listeners through interaural grouping of sequential stimuli by proximity in pitch and let to a high detection performance there. However, their higher monaural performance indicated, that they were able to segregate concurrent monaural stimuli into separate streams based on pitch. The use of pulse rate cues in addition to the electrode place cues likely aided this segregation. This presents an encouragement to utilize the pulse rate also in clinical stimulation strategies to enhance pitch cues and the patients' abilities to segregate streams. Also, the participants were rather satisfied with the correspondence of the acoustic stimuli matched to the electric stimulation of the CI and there were no significant differences in the monaural performance across electric and acoustic stimulation. Together, this suggests differences in the participants' binaural streaming behavior compared to normal-hearing and bilateral CI listeners. The reasons for the deviant binaural streaming behavior remain a matter of investigation. The prolonged exposure to bimodal stimulation, with the HA presenting the same sounds with quite different quality compared to the CI, may have let to adaptations in the perception and the binaural processing of the bimodal CI patients. Such adaptations may prevent the integration of similar binaural stimuli into a common stream percept or could have led listeners to integrate dichotic binaural stimuli into a common percept instead of two separate ones. Further studies could investigate the effects of the binaural streaming behavior on patients' speech reception in noise and listening effort. Should these be affected negatively, bilateral aids of the same kind, whether HAs or CIs, may be preferable to bimodal stimulation in the current state. However, one could also attempt to align the percepts better across the two devices to retain the patients' abilities for binaural streaming.

6 Overall discussion

6.1 Summary of main results

The works in this thesis were motivated by the question whether the two devices aiding bimodal CI patients, a CI combined with a HA on the other ear, work well together. From this broad question, a more specific one was formed: Auditory streaming is essential to how the human brain analyses the surrounding acoustic scene, and similarity in perceptual aspects, such as pitch, is commonly used to group auditory objects and streams. CIs, especially in combination with a contralateral HA, can lead to fundamentally different representations of the same sound source across ears. In turn, this may affect the binaural processes, which govern binaural streaming. Therefore, the studies presented in this thesis centered on the question if the bimodal patients can form steamed percepts, integrating binaural stimuli. This has been investigated via the self-assessment in a survey for bimodal CI patients, as well as via psychophysical experiments. Since CIs by themselves also change the percepts of sound by representing them in a profoundly different way, changes in those binaural processes could be contributed to either the influence of cochlear implantation or the stark differences of sound representations across ears in the bimodal combination. Therefore, listening experiments were conducted with both bilateral and bimodal CI listeners. The design of these new listening experiments also made possible an investigation of the monaural streaming abilities of the listeners, allowing to compare them to the binaural results and to verify if the streaming cues provided were sufficient to segregate concurrent stimuli.
In the following, the chapters' main findings are summarized and discussed, starting with a comparison of the monaural and binaural streaming abilities of the bilateral and bimodal CI participants, before giving future perspectives and a summary of the conclusions.

6.1.1 Monaural and binaural streaming with cochlear implants

The results for both bilateral (chapter 4) and bimodal CI patients (chapter 5) suggest that most of the listeners can segregate concurrent melodies monaurally based on pitch, at least when the two streams are separated well by both electrode place and pulse rate cues. This befits the argument brought forward by Moore and Gockel (2002), that every perceptual difference in the stimuli could be used as a segregation cue and, likewise, the results of Luo et al. (2012), which suggest that CI patients benefit from combination of congruent place and temporal pitch cues.

The place coding for pitch is typical for clinical CI stimulation strategies (Oxenham, 2018; Zeng et al., 2008), but previous studies on streaming and binaural fusion in CI listeners demonstrated that the low salience of place pitch cues alone often proves incapable of providing sufficient perceptual separation to allow listeners to segregate multiple sound objects. This has been demonstrated for short dichotic (Kan et al., 2013; Reiss et al., 2014a, 2017; Steel et al., 2015), sequential (Chatterjee et al., 2006; Cooper and Roberts, 2007; Duran et al., 2012), and concurrent stimuli (Carlyon et al., 2007; Cooper and Roberts, 2010; Deeks and Carlyon, 2004). Some of the named studies also utilized the pulse rate as a segregation cue. On its own and for smaller changes, the pulse rate appeared insufficient to let listeners segregate streams. Studies by Luo et al. (2012) and Lamping et al. (2017) also suggested that both place and temporal pitch cues are effectively integrated into a common percept, which explains how both together enhance the perceptual separation of the pitch-percepts. The results presented in chapters 4 and 5 suggest that both cues together sufficiently separated the two streams perceptually to allow for their segregation. This also forms a basis to assess the listeners performance binaurally, since if they had been unable to segregate the streams monaurally, they would likely also be unable to segregate them binaurally, so that no conclusion about their abilities for binaural streaming could have been made.

In binaural streaming, i.e., the formation of common streams from binaurally distributed stimuli, differences among bilateral and bimodal CI listeners emerged. The results of chapter 4, dealing with the bilateral CI patients, indicated that these listeners were able to stream binaurally, grouping stimuli across ears. Hence, binaural streaming should generally be possible with CIs, at least when using stimuli that are perceptually aligned across ears. This alignment may be vital to binaural streaming, since streaming relies upon the grouping of similar sound components. Bilateral CIs may allow for a better correspondence in percepts across ears compared to a bimodal solution, since both ears are stimulated in the same manner.

Whether CI patients can perform binaural streaming in practice may depend on how well aligned the sound representations in their two devices are, but also on the stimulation strategy used. To allow for streaming, monaurally or binaurally, the results suggest that concurrent sounds must be presented with sufficient perceptual separation, e.g., in pitch. The precise thresholds of that separation could be investigated in further studies. If concurrent sound sources are close in frequency, they may either fall into the same channel or into adjacent channels. Considering a stimulation strategy that only uses the electrode place to convey pitch information, for sounds (or their components) falling within the same channel, the patients would be presented with the temporal envelope of the sound mixture in the respective band. It is unlikely that CI patients can utilize these temporal envelope cues alone to segregate in this case (Oxenham, 2008; Quin and Oxenham, 2005). If concurrent sounds fall into adjacent channels, the low separation in electrode place for neighboring channels may likewise provide insufficient separation to segregate concurrent sound sources. When concurrent sounds are separated enough in frequency, so that they end up in channels mapped to electrodes separated in place by about 3 mm and more, it becomes more likely that (most) patients can use the resulting pitch difference to segregate them (Chatterjee et al., 2006; Paredes-Gallardo et al., 2018b). Real, broad-band sounds that extend over multiple bands can lead to partial spectral overlaps, which should provide limited cues for segregation based on parts of their non-overlapping spectrum. If a stimulation strategy also employs the pulse rate, this could aid segregation by providing an additional means of separation over the electrode place.

However, the stimulation strategies first have to present the concurrent sounds at all. If a strategy limits the number of electrodes stimulated per time frame (Bingabr et al., 2008; Fishman et al., 1997; Friesen et al., 2001), it depends on the channel selection algorithms, whether concurrent sound sources are presented. If, for example, the selection is made purely upon sound level, this could lead to the presentation of only the louder sound source or a mix of components from multiple sources.

The timing, as marked by common onset of sounds, provides another stream-

ing cue (Bregman, 1990). For this reason the stimuli in the experiments were delivered synchronously. If clinical CI processors (and HAs) are not operating synchronized, they could induce a delay, which may lead patients to segregate stimuli by ear-of-input, like in the delay control-condition employed in the verification of the experimental method in chapter 3.
The bimodal CI listeners in chapter 5 reported a rather high satisfaction with the correspondence between electric stimulation and the acoustic tone complexes they had matched to it. Despite that perceptual proximity of the interaurally corresponding stimuli, the bimodal CI patients appear to have been unable to group interaurally by pitch-proximity. This suggests, that changes in their binaural processing prevent the formation of streams from binaural stimulation rather than perceptual mismatches, although the latter cannot fully be ruled out. The changes in pitch percepts over time reported in Reiss et al. (2014b) support that the auditory systems of CI patients can adapt to the stimulation over time. But eventual changes in pitch or other percepts would already be included in the matched acoustic stimuli, so that a proximity in percepts across ears was likely given, as reflected by the participants' ratings. Consequently, the extended exposure to the different sound representations by bimodal stimulation may alter binaural processing, so that sound objects are no longer formed normally from binaural stimuli.
The design of the survey in chapter 2, dealing with the patient's subjective experience, was a first step to assess the listener's perspective. It was of an explorative nature and, owing to the nature of a questionnaire study, it cannot be guaranteed that all participants interpreted the questions in the same way or mapped their perception to the provided scales in the same way. This could account for some of the observed variability. Therefore, the interpretations should be taken with a bit of caution. Nevertheless, the results present a document of the patients' personal perspective on their perception with the devices and the balance in percepts between them. The results also agree with the findings of the listening experiment. In the survey among thirty-eight bimodal CI patients, only three reported to perceive the voice of a person as one single uniform sound object, while most participants perceived it with differences in pitch and loudness across ears, in some cases so extreme that they reported to perceive two entirely separate voices. The large number of intermediate ratings in between the extremes could be a sign that participants only integrated over a limited spectral range, covered by both HA and CI. The ratings were not changed

significantly by the listening situation. A lack of integration of the percepts from HA and CI did come with trends to worse speech reception in noise and higher listening effort as assessed through the SSQ5 questionnaire. This connection makes sense, as both could be explained as side-effects of a lack of integration: If the binaural percepts are not integrated into a common sound object, this doubles the number of sounds, leading to a more complex listening situation. However, these listening performance measures depend not only on binaural streaming, but can also be limited by how well the devices represent auditory cues and how well the listener's auditory system can decode them, which may explain why the correlations were not significant.
A lack of the capability for binaural streaming implies that patients may perceive their acoustic surroundings in a profoundly different way compared to normal-hearing. As some patients indicated in the bimodal survey, in the extreme that can mean every sound is perceived twice, separately for left and right ear. Consequently, there are twice as many background sounds and twice as many relevant sounds a patient may desire to focus on. This could make it harder or impossible to localize sounds using binaural cues, lead to worse speech reception in noise, and higher listening effort. It may further make listening to music more complex, since there would be twice as many melodies present. Because of this, it is a worthwhile goal to further investigate how to optimize binaural streaming for CI patients and how it links to listening performance.

6.2 Perspectives

It is encouraging to see that the bilateral CI listeners generally appear capable of forming streams from concurrent stimulation both monaurally as well as binaurally. More research could be conducted to understand the exact limits required for perceptual separation of streams, in order to devise new stimulation strategies that aid CI patients in forming streams from binaural stimulation and segregate concurrent sound sources. Stimulation strategies that utilize only place coding to provide patients with pitch cues appear to fall short in separating concurrent sound sources well enough if they fall into bands close in frequency. This reflects in the studies of streaming in CI listeners, as well as the patients limited abilities when it comes to speech reception in noise.
One step could be to explore the effect of binaural matchings for bilateral CI

patients, like the one employed in chapter 4, in clinical practice and whether the matched percepts, over time, contribute toward better outcomes for the patients. For bimodal CI patients, there is not such a rather straightforward way of optimization. With the results of chapter 5, it appears possible that the extended exposure to very different representations of sounds across ears has led to changes in binaural processing, so that common sound objects are no longer formed by proximity in perceptual aspects from binaural stimulation. These results are in agreement with previous studies of binaural integration and binaural processing (Gordon et al., 2012; Gordon et al., 2017; Polonenko et al., 2019). A logical conclusion is to align the percepts again across ears, but, as the results of chapter 5 demonstrate, this may not directly, if ever, lead their auditory system to group again by proximity in perceptual aspects, such as pitch, from binaural stimuli. Thus, more research is necessary to understand if these patients can learn to adopt such natural grouping behavior again.

The limitations of electric stimulation make it seem unlikely that CI-stimulation will approximate natural sound characteristics, even though newer techniques, such as current steering, could minimize the problem of current spread. The perhaps most straightforward way to align the percepts across ears and allow patients to stream binaurally would be bilateral cochlear implantation with appropriate perceptual alignment across ears. Ideally, one could measure the binaural interaction component of the auditory brainstem response to perfectly align the electrodes' interaural percepts physically in an objective way or select matching electrode pairs from a much larger, more closely spaced, number of electrodes. However, the cost for bilateral implantation is much higher than for the bimodal solution and, thus, its reimbursement may present a challenge. Moreover, patients may be reluctant to fully commit to CIs and give up the acoustic hearing they are used to. Another solution could be to change the output of the hearing aid, so that it more closely matches the CI's sound representaton. While such matching in the bimodal study (chapter 5) did not appear to allow the participants to form binaural streams immediately, providing them with better matched sound representations across ears in their daily life could preserve the processes that integrate sounds binaurally. For this, the devices' frequency-channel mapping could be aligned according to the patients' percept. Similar to the matching utilized in the bimodal study, the parameters of an acoustic tone complex could be paired to the stimulation by the CI's electrodes. Then, instead of the typical HA output per channel, the signals could be filtered with the matched tone-complex to increase the similarity to the sound's representation by the CI, while maintaining the HA's ability to present fundamental frequency and fine structure of the sounds. However, since at least some bimodal CI patients also benefit from the present stimulation by their HA, as in, e.g., Kong et al. (2005), fundamentally changing the HAs output is not an easy decision to make. However, it could be an alternative for bimodal patients not benefitting significantly from the HA and for whom bilateral implantation is not an option. Alternatively, bilateral implantation may provide the patients with a better chance to preserve their abilities for binaural streaming.

6.3 Conclusions

The findings presented throughout the chapters of this thesis suggest that:

- In general, CIs can allow for streaming of concurrent melodies both monaurally, as well as binaurally.
- Larger perceptual separation in the sound's representations than available in current clinical devices may be required to segregate concurrent streams and overcome abnormally-wide binaural fusion.
- A certain level of perceptual similarity across ears is required for binaural streaming in normal-hearing listeners, which extends to CI patients. Clinical devices may not always offer this similarity without undertaking further steps to align percepts interaurally.
- Bilateral CI appear to provide greater interaural similarity, at least after perceptual alignment, as required for successful streaming.
- Most bimodal CI patients may be unable to integrate binaurally distributed stimulation into a common stream, likely due to the differences in interaural representations of sounds, which these patients experience on a daily basis.
- The lack of binaural streaming and integration is also reflected in self-reported data and could degrade speech reception in noise and listening effort.

Overall, these results may provide a starting point for further research regarding the limits of binaural stream segregation in CI patients, which could lead to new technologies allowing the patients to better segregate concurrent streams in their daily life. For this, changes in implantation criteria, new devices with optimized stimulation strategies, and adoption of interaural alignment into clinical practice may be necessary.

Bibliography

Advanced Bionics (2016). *Optimizing Hearing for Listeners with a Cochlear Implant and Contralateral Hearing Aid.* URL: `https://advancedbionics.com/content/dam/advancedbionics/Documents/Regional/US/libraries/bimodal/Bimodal-Fitting-Formula-Whitepaper.pdf` (visited on 06/12/2019).

Anstis, S. M. and S. Saida (1985). “Adaptation to auditory streaming of frequency-modulated tones”. In: *J. Exp. Psychol. Hum. Percept. Perform.* 11.5, 257–271.

Aronoff, J. M., C. Shayman, A. Prasad, D. Suneel, and J. Stelmach (2015). “Unilateral spectral and temporal compression reduces binaural fusion for normal hearing listeners with cochlear implant simulations”. In: *Hearing Research* 320, pp. 24–29.

Aronoff, J., M. Padilla, J. Stelmach, and D. M. Landsberger (2016a). “Clinically Paired Electrodes Are Often Not Perceived as Pitch Matched”. In: *Trends Hear* 20, pp. 1–9.

Benjamini, Y. and Y. Hochberg (1995). “Controlling the false discovery rate: A practical and powerful approach to multiple testing”. In: *J. R. Statist. Soc. B* 57.1, pp. 289–300.

Bernstein, J. G. W., M. J. Goupell, G. I. Schuchman, A. L. Riviera, and D. S. Brungart (2016). “Having two ears facilitates the perceptual separation of concurrent talkers for bilateral and single-sided deaf cochlear implantees”. In: *Ear Hear.* 37, pp. 289–302.

Bingabr, M., B. Espinoza-varas, and P. C. Loizou (2008). “Simulating the effect of spread of excitation in cochlear implants”. In: *Hearing Research* 241, pp. 73–79.

Breebaart, J., S. v. d. Par, and A. Kohlrausch (2001). “Binaural processing model based on contralateral inhibition I . Model structure”. In: *J. Acoust. Soc. Am.* 110, pp. 1074–1088.

Bregman, A. S. (1978). “Auditory streaming is cumulative”. In: *J. Exp. Psychol. Hum. Percept. Perform.* 4.3, 380–387.

Bregman, A. S. and J. Campbell (1971). "Primary auditory stream segregation and perception of order in rapid sequences of tones". In: *J. Exp. Psychol.* 89.2, 244–249.

Bregman, A. (1990). *Auditory scene analysis. The perceptual organization of sound.* MIT Press, Boston.

Brink, G. van den, K. Sintnicolaas, and W. S. van Stam (1976). "Dichotic pitch fusion". In: *J. Acoust. Soc. Am.* 59, 1471–1476.

Carlyon, R. P., C. J. Long, J. M. Deeks, and C. M. McKay (2007). "Concurrent sound segregation in electric and acoustic hearing". In: *J. Ass. Res. Otolaryngol.* 8.1, 119–133.

Chang, Y.-P., R. B. Chang, C.-Y. Lina, and X. Luo (2016). "Mandarin Tone and Vowel Recognition in Cochlear Implant Users". In: *Ear Hear.*, 37, pp. 271–281.

Chatterjee, M., A. Sarampalis, and S. I. Oba (2006). "Auditory stream segregation with cochlear implants: A preliminary report". In: *Hear. Res.* 222, 100–107.

Cherry, C. E. (1953). "Some further experiments upon the recognition of speech, with one and with two Ears". In: *J. Acoust. Soc. Am.* 25.5, 975–979.

Ching, T. Y. C., E. van Wanrooy, and H. Dillon (2007). "Binaural-Bimodal Fitting or Bilateral Implantation for Managing Severe to Profound Deafness: A Review". In: *Trends in Amplification* 11.3, 161–192.

Ching, T. Y., M. Hill, H. Dillon, and E. van Wanrooy (2004). *Fitting and evaluating a hearing aid for recipients of a unilateral cochlear implant: the NAL approach. Part 1: Hearing aid prescription, adjustment, and evaluation.* URL: `http://www.hearingreview.com/2004/07/fitting-and-evaluating-a-hearing-aid-for-recipients-of-a-unilateral-cochlear-implant-the-nal-approach/` (visited on 06/12/2019).

Christiansen, S. K., M. L. Jepsen, and T. Dau (2014). "Effects of tonotopicity, adaptation, modulation tuning, and temporal coherence in 'primitive' auditory stream segregation". In: *J. Acoust. Soc. Am.* 135.1, 323–333.

Cooper, H. R. and B. Roberts (2007). "Auditory stream segregation of tone sequences in cochlear implant listeners". In: *Hear. Res.* 225, 11–24.

Cooper, H. R. and B. Roberts (2010). "Simultaneous grouping in cochlear implant listeners: Can abrupt changes in level be used to segregate components from a complex tone?" In: *J. Ass. Res. Otolaryngol.* 11.1, 89–100.

Cusack, R. and B. Roberts (2000). "Effects of differences in timbre on sequential grouping". In: *Percept. Psychophys.* 62.5, 1112–1120.

David, M., N. Grimault, and M. Lavandier (2015). "Sequential streaming, binaural cues and lateralization". In: *J. Acoust. Soc. Am.* 138.6, 3500–3512.

Deeks, J. M. and R. P. Carlyon (2004). "Simulations of cochlear implant hearing using filtered harmonic complexes: Implications for concurrent sound segregation". In: *J. Acoust. Soc. Am.* 115, pp. 1736–1748.

Deutsch, D. (1975). "Two-Channel Listening to Musical Scales". In: *J. Acoust. Soc. Am.* 57, pp. 1156–1160.

Deutsch, D. (1974). "An auditory illusion". In: *Nature* 251, pp. 307–309.

Deutsch, D. (1979). "Binaural integration of melodic patterns". In: *Percept. Psychophysics* 25.5, pp. 399–405.

Deutsch, D. (1985). "Dichotic Listening to Melodic Patterns and Its Relationship to Hemispheric Specialization of Function". In: *Mus. Percept.* 3.2, pp. 127–154.

Dobie, R. A. and S. J. Norton (1980). "Binaural Interaction in Human Auditory Evoked Potentials". In: *Electroencephalogr. Clin. Neurophysiol.* 49, 303–313.

Dorman, M. F. and R. H. Gifford (2010). "Combining acoustic and electric stimulation in the service of speech recognition". In: *Internat. J. Audiol.* 49.12, 912–919.

Dunn, C. C., R. S. Tyler, and S. A. Witt (2005). "Benefit of Wearing a Hearing Aid on the Unimplanted Ear in Adult Users of a Cochlear Implant". In: *J. Speech Lang. Hear. Res.* 48, 668–681.

Duran, S. I., L. M. Collins, and C. S. Throckmorton (2012). "Stream segregation on a single electrode as a function of pulse rate in cochlear implant listeners". In: *J. Acoust. Soc. Am.* 132.6, 3849–3857.

Easwar, V., J. Sanfilippo, B. C. Papsin, and K. Gordon (2018). "Impact of Consistency in Daily Device Use on Speech Perception Abilities in Children with Cochlear Implants: Datalogging Evidence". In: *J. Am. Acad. Audiol.* 29.9, pp. 835–846.

Erikson-Mangold, M. and S. G. Carlsson (1991). "Psychological and somatic distress in relation to perceived hearing disability, hearing handicap, and hearing measurements". In: *J. Psychosom. Res.* 35.6, pp. 729–740.

Fishman, K. E., R. V. Shannon, and W. H. Slattery (1997). "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor". In: *J. Speech Lang. Hear. Res.* 40.5, pp. 1201–1215.

Fitzgerald, M. B., A. Kan, and M. J. Goupell (2015). “Bilateral Loudness Balancing and Distorted Spatial Perception in Recipients of Bilateral Cochlear Implants”. In: *Ear Hear.* 36.5, e225–e236.

Florentine, M., B. Buus S. Scharf, and E. Zwicker (1980). “Frequency selectivity in normally-hearing and hearing-impaired observers”. In: *J. Speech Lang. Hear. Res.* 23.3, pp. 646–669.

Flynn, M. C., R. C. Dowell, and G. M. Clark (1998). “Aided speech recognition abilities of adults with a severe or severe-to-profound hearing loss”. In: *J. Speech Lang. Hear. Res.* 41.2, pp. 285–299.

Fredelake, S. and V. Hohmann (2012). “Factors affecting predicted speech intelligibility with cochlear implants in an auditory model for electrical stimulation”. In: *Hear. Res.* 287.1-2, pp. 76–90.

Friesen, L. M., R. V. Shannon, D. Baskent, and X. Wang (2001). “Speech recognition in noise as a function of spectral channels: Comparison of acoustic hearing and cochlear implants”. In: *J. Acoust. Soc. Am.* 110.2, pp. 1150–1163.

Galvin, J. J., Q.-J. Fu, and G. Nogaki (2007). “Melodic contour identification by cochlear implant listeners”. In: *Ear Hear.* 28, 302–319.

Galvin, J. J., Q.-J. Fu, and S. I. Oba (2009a). “Effect of a competing instrument on melodic contour identification by cochlear implant users”. In: *J. Acoust. Soc. Am. Expr. Lt.* 125.3, pp. 98–103.

Galvin, J. J., Q.-J. Fu, and R. V. Shannon (2009b). “Melodic contour identification and music perception by cochlear implant users”. In: *Ann. N. Y. Acad. Sci.* 1169, 518–533.

Gifford, R. H., M. F. Dorman, S. A. Mckarns, and A. J. Spahr (2007). “Combined Electric and Contralateral Acoustic Hearing: Word and Sentence Recognition With Bimodal Hearing”. In: *Journal of Speech Language and Hearing Research* 50.4, pp. 835–843.

Gordon, K. A., C. Salloum, G. S. Toor, R. van Hoesel, and B. C. Papsin (2012). “Binaural Interactions Develop in the Auditory Brainstem of Children Who Are Deaf: Effects of Place and Level of Bilateral Electrical Stimulation”. In: *Journal of Neuroscience* 32.12, pp. 4212–4223.

Gordon, K. A., S. L. Cushing, V. Easwar, M. J. Polonenko, and B. C. Papsin (2017). “Binaural integration: A challenge to overcome for children with hearing loss”. In: *Current Opinion in Otolaryngology & Head and Neck Surgery* 25.6, pp. 514–519.

Goupell, M. J., C. Stoelb, A. Kan, and R. Y. Litovsky (2013). "Effect of mismatched place-of-stimulation on the salience of binaural cues in conditions that simulate bilateral cochlear-implant listening". In: *J. Acoust. Soc. Am.* 133, pp. 2272–2287.

Guérit, F., S. Santurette, J. Chalupper, and T. Dau (2014). "Investigating Interaural Frequency-Place Mismatches via Bimodal Vowel Integration." In: *Trends Hear.* 18, pp. 1–10.

Hamzavi, J., S. M. Pok, and W. Gstoettner (2004). "Speech perception with a cochlear implant used in conjunction with a hearing aid in the opposite ear". In: *International Journal of Audiology* 43, pp. 61–65.

Hollander, M., D. A. Wolfe, and E. Chicken (1973). *Nonparametric Statistical Methods.* Hoboken, New Jersey: Jon Wiley & Sons.

Hong, R. S. and C. W. Turner (2006). "Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients". In: *J. Acoust. Soc. Am.* 120.1, 360–374.

Iverson, P. (1995). "Auditory stream segregation by musical timbre: effects of static and dynamic acoustic attributes". In: *J. Exp. Psychol. Hum. Percept. Perform.* 21.4, 751–763.

Janssen, N. A., L. Bramslow, S. Riis, and J. Marozeau (2019b). "A modification of the scale illusion into a detection-task for assessment of binaural streaming". In: *J. Acoust. Soc. Am.* 145.6, pp. 457–462.

Jürgens, T., V. Hohmann, A. Büchner, and W. Nogueira (2019). "The effects of electrical field spatial spread and some cognitive factors on speech-in-noise performance of individual cochlear implant users—A computer model study". In: *PLoS ONE* 13.4, pp. 1–20.

Kan, A., C. Stoelb, R. Y. Litovsky, and M. J. Goupell (2013). "Effect of mismatched place-of-stimulation on binaural fusion and lateralization in bilateral cochlear-implant users". In: *J. Acoust. Soc. Am.* 134, 2923–2936.

Kates, J. M. (2008). *Digital hearing aids.* San Diego: Plural Publishing.

Kong, Y.-Y., G. S. Stickney, and F.-G. Zeng (2005). "Speech and melody recognition in binaurally combined acoustic and electric hearing". In: *J. Acoust. Soc. Am.* 117.3, 1351–1361.

Lamping, W., S. Santurette, and J. Marozeau (2017). "Verbal attribute magnitude estimates of pulse trains across electrode places and stimulation rates in cochlear implant listeners". In: *Proceedings of the International Symposium on Auditory and Audiological Research. Vol. 6: Adaptive Processes in Hearing.*

Ed. by S. Santurette, T. Dau, J. C. Dalsgaard, L. Tranebjærg, T. Andersen, and T. Poulsen. The Danavox Jubilee Foundation, pp. 215–222.

Landsberger, D. M., A. C. K. Vermeire, V. V. Rompaey, and P. V. d. Heyning (2016). "Qualities of single electrode stimulation as a function of rate and place of stimulation with a cochlear implant". In: *Ear Hear.* 37.3, 149–159.

Lazard, D. S., J. Marozeau, and H. J. McDermott (2012). "The sound sensation of Apical electric stimulation in cochlear implant recipients with contralateral residual hearing". In: *J. PLoS One* 7.6, 2–6.

Liang, D. H., H. S. Lusted, and R. L. White (1999). "The nerve-electrode interface of the cochlear implant: current spread". In: *IEEE Transact. Biomed. Eng.* 46.1.

Litovsky, R. Y. and K. Gordon (2016). "Bilateral cochlear implants in children: Effects of auditory experience and deprivation on auditory perception". In: *Hear. Res.* 338, pp. 76–87.

Litovsky, R. Y., M. J. Goupell, S. M. Misurelli, and A. Kan (2017). "Hearing with cochlear implants and hearing aids in complex auditory scenes". In: *The Auditory System at the Cocktail Party*. Ed. by J. C. Middlebrooks, J. Z. Simon, A. N. Popper, and R. R. Fay. New York: Springer, 261–291.

Lopez-Poveda, E. A., A. Eustaquio-Martín, J. S. Stohl, R. D. Wolford, R. Schatzer, and B. S. Wilson (2016). "A Binaural Cochlear Implant Sound Coding Strategy Inspired by the Contralateral Medial Olivocochlear Reflex". In: *Ear Hear.* 37, e138–e148.

Luo, X., M. Padilla, and D. M. Landsberger (2012). "Pitch contour identification with combined place and temporal cues using cochlear implants". In: *J. Acoust. Soc. Am.* 131.2, 1325–1336.

Ma, N., S. Morris, and P. T. Kitterick (2016). "Benefits to speech perception in noise from the binaural integration of electric and acoustic signals in simulated unilateral deafness". In: *Ear Hear.,* 37, pp. 248–259.

McDermott, J. H. (2009). "The cocktail party problem". In: *Curr. Biol.* 19.22, 1024–1027.

McKay, C. M., J. H. McDermott, and R. P. Carlyon (2000). "Place and temporal cues in pitch perception: are they truly independent?" In: *Acoust. Res. Lett. Online* 1.1, p. 25.

Mcdermott, H. J. (2004). "Music Perception with Cochlear Implants : A Review". In: *Trends in Amplification* 8.2, pp. 49–82.

Mehta, A. H., I. Yasin, A. J. Oxenham, and S. Shamma (2016). “Neural correlates of attention and streaming in a perceptually multistable auditory illusion”. In: *J. Acoust. Soc. Am.* 140.4, pp. 2225–2233.

Mehta, A. H., N. Jacoby, I. Yasin, A. J. Oxenham, and S. A. Shamma (2017). “An auditory illusion reveals the role of streaming in the temporal misallocation of perceptual objects”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1714, pp. 1–10.

Mertens, G. and P. van de Heyning (2013). “Self-assessment of hearing disabilities in cochlear implant users using the SSQ and the reduced SSQ5 version”. In: *Otol. Neurotol.* 34, 1622–1629.

Micheyl, C. and A. J. Oxenham (2010b). “Pitch, harmonicity and concurrent sound segregation: Psychoacoustical and neurophysiological findings”. In: *Hear. Res.* 266.1-2, 36–51.

Moore, B. C. J. and J. I. Alcantara (2001). “The use of psychophysical tuning curves to explore dead regions in the Cochlea”. In: *Ear Hear.* 22.4, pp. 268–278.

Moore, B. C. J. and H. E. Gockel (2002). “Factors influencing sequential stream segregation”. In: *Acta Acustica united with Acustica* 88.3, 320–333.

Morera, C. et al. (2012). “Contralateral hearing aid use in cochlear implanted patients: Multicenter study of bimodal benefit”. In: *Acta Oto-Laryngologica* 132.10, pp. 1084–1094.

Murtagh, F. and P. Legendre (2014). “Ward’s hierarchical agglomerative clustering method: Which algorithms implement Ward’s criterion?” In: *J. Classif.* 31, pp. 274–295.

Nelson, P. B., S.-H. Jin, A. E. Carney, and D. A. Nelson (2003). “Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners”. In: *J. Acoust. Soc. Am.* 113.2, 961–968.

Noorden, L. P. A. S. van (1975). *Temporal coherence in the perception of tone sequences, PhD thesis.* Technical University Eindhoven.

Noorden, L. P. A. S. van (1977). “Minimum differences of level and frequency for perceptual fission of tone sequences ABAB*”. In: *J. Acoust. Soc. Am.* 61.4, 1041–1045.

Offeciers, E et al. (2005). “International consensus on bilateral cochlear implants and bimodal stimulation”. In: *Acta Oto-Largyngologica* 125.9, pp. 918–919.

Oxenham, A. J. (2008). “Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants”. In: *Trends Amplif.* 12.4, 316–331.

Oxenham, A. J. (2018). “How we hear: The perception and neural coding of sound”. In: *Annu. Rev. Psychol.* 69.1, 27–50.

Paredes-Gallardo, A., H. Innes-Brown, S. M. K. Madsen, T. Dau, and J. Marozeau (2018a). “Auditory Stream Segregation and Selective Attention for Cochlear Implant Listeners: Evidence From Behavioral Measures and Event-Related Potentials”. In: *Frontiers in Neuroscience.*

Paredes-Gallardo, A., S. M. K. Madsen, T. Dau, and J. Marozeau (2018b). “The role of place cues in voluntary stream segregation for cochlear implant users”. In: *Trends Hear.* 22.

Paredes Gallardo, A., S. M. K. Madsen, T. Dau, and J. Marozeau (2018c). “The role of temporal cues on voluntary stream segregation in cochlear implant users”. In: *Trends Hear.* Pp. 1–13.

Plack, C. J. (2018). *The Sense of Hearing.* New York: Routledge.

Polonenko, M. J., B. C. Papsin, and K. A. Gordon (2019). “Cortical plasticity with bimodal hearing in children with asymmetric hearing loss”. In: *Hear. Res.* 372, pp. 88–98.

Quin, M. K. and A. J. Oxenham (2005). “Effects of envelope-vocoder processing on F0 discrimination and concurrent-vowel identification”. In: *Ear Hear.* 26.5, pp. 451–460.

Reiss, L. A. J., R. A. Ito, J. L. Eggleston, and D. R. Wozny (2014a). “Abnormal binaural spectral integration in cochlear implant users”. In: *J. Ass. Res. Otolaryngol.* 15.2, 235–248.

Reiss, L. A. J. et al. (2014b). “Pitch Adaptation Patterns in Bimodal Cochlear Implant Users: Over Time and After Experience”. In: *Ear Hear.* 36.2, e23–e34.

Reiss, L. A. J., J. R. Fowler, C. L. Hartling, and Y. Oh (2017). “Binaural pitch fusion in bilateral cochlear implant users”. In: *Ear Hear.* 39, pp. 390–397.

Reiss, L. A. J., J. L. Eggleston, E. P. Walker, and Y. Oh (2016). “Two Ears Are Not Always Better than One: Mandatory Vowel Fusion Across Spectrally Mismatched Ears in Hearing-Impaired Listeners”. In: *JARO - Journal of the Association for Research in Otolaryngology* 17.4, pp. 341–356.

Reyes, C. (2016). *Oticon Bimodal Fitting Guideline 2016.* URL: `http://ipaper.ipapercms.dk/Oticon/Whitepaper/WP{\_}Bimodal{\_}fitting{\_}H22015/{\#}/` (visited on 06/12/2019).

Rose, M. M. and B. C. J. Moore (2005). “The relationship between stream segregation and frequency discrimination in normally hearing and hearing-impaired subjects”. In: *Hear. Res.* 204, 16–28.

Sach, A. J. and P. J. Bailey (2004). “Some characteristics of auditory spatial attention revealed using rhythmic masking release”. In: *Percept. & Psychophys.* 66.8, 1379–1387.

Sataloff, R. T. and J. Sataloff (2005). *Hearing loss.* Boca Raton: Taylor & Francis.

Scherf, F. W. a. C. and L. P. Arnold (2014). “Exploring the clinical approach to the bimodal fitting of hearing aids and cochlear implants: results of an international survey.” In: *Acta oto-laryngologica* 134.11, pp. 1151–1157.

Seeber, B. (2004). “Localization ability with bimodal hearing aids and bilateral cochlear implants”. In: *J. Acoust. Soc Am.* 116, 1698–1709.

Shamma, S., A. M. Elhilali, and C. Micheyl (2011). “Temporal coherence and attention in auditory scene analysis”. In: *Trends Neurosci.* 34.3, 114–23.

Shamma, S. et al. (2013). “Temporal Coherence and the Streaming of Complex Sounds”. In: *Basic Asp. Hear. Physiol. Percept.* Ed. by B. C. J. Moore, R. D Patterson, I. M. Winter, R. P. Carlyon, and H. E. Gockel. Vol. 787, 535–543.

Shannon, R. (1983a). “Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics”. In: *Hear. Res.* 11.2, 157–189.

Shannon, R. (1983b). “Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction”. In: *Hear. Res.* 12.1, 1–16.

Sheffield, S. W., M. Simha, K. N. Jahn, and R. H. Gifford (2016). “The effects of acoustic bandwidth on simulated bimodal benefit in children and adults with normal hearing”. In: *Ear Hear.,* 37, pp. 282–288.

Shinn-Cunningham, B., V. Best, and A. K. C. Lee (2017). “Auditory object formation and selection”. In: *The auditory system at the cocktail party.* Ed. by J. C. Middlebrooks, J. Z. Simon, A. N. Popper, and R. R. Fay. New York: Springer, 7–40.

Singh, P. G. and A. S. Bregman (1997). “The influence of different timbre attributes on the perceptual segregation of complex-tone sequences”. In: *J. Acoust. Soc. Am.* 102.4, 1943–1952.

Smeds, K., F. Wolters, and M. Rung (2015). “Estimation of Signal-to-Noise Ratios in Realistic Sound Scenarios”. In: *Journal of the American Academy of Audiology* 26.2, pp. 183–196.

“Speech recognition with primarily temporal cues” (1995). In: 80, 303–304.

Stainsby, T. H., C. Fu, H. J. Flanagan, S. K. Waldman, and B. C. J. Moore (2011). "Sequential streaming due to manipulation of interaural time". In: *J. Acoust. Soc. Am.* 130.2, 904–914.

Staisloff, H. E., D. H. Lee, and J. M. Aronoff (2016). "Perceptually aligning apical frequency regions leads to more binaural fusion of speech in a cochlear implant simulation". In: *Hearing Research* 337, 59–64.

Steel, M. M., B. C. Papsin, and K. A. Gordon (2015). "Binaural fusion and listening effort in children who use bilateral cochlear implants: a psychoacoustic and pupillometric study". In: *PLoS One* 10, e0117611.

Studebaker, G. A. (1985). "A 'rationalized' arcsine transform". In: *J. Speech Hear. Res.* 28, 455–462.

Suneel, D., H. Staisloff, C. S. Shayman, J. Stelmach, and J. M. Aronoff (2017). "Localization performance correlates with binaural fusion for interaurally mismatched vocoded speech". In: *J acoust. Soc. Am.* 142.3, EL276–EL280.

Vandali, A., D. Sly, R. Cowan, and R. van Hoesel (2013). "Pitch and loudness matching of unmodulated and modulated stimuli in cochlear implantees". In: *Hear. Res.* 302, 32–49.

Veugen, L. C. E., J. Chalupper, A. F. M. Snik, A. J. V. Opstal, and L. H. M. Mens (2016). "Matching Automatic Gain Control Across Devices in Bimodal Cochlear Implant Users". In: *Ear Hear.,* 37, pp. 260–270.

Vickers, D., P. Kitterick, C. Verschuur, C. Leal, F. Vickers, and J. Graham (2016). "Issues in Cochlear Implant Candidacy Issues in Cochlear Implant Candidacy". In: *Cochlear Implants Internat.* 17.1, pp. 1–2.

Vliegen, J, B. C. J. Moore, and A. J. Oxenham (1999b). "The role of spectral and periodicity cues in auditory stream segregation, measured using a temporal discrimination task". In: *J. Acoust. Soc. Am.* 106.2, 938–945.

Vroegop, J., A. Goedegebure, and M. van der Schroeff (2018). "How to Optimally Fit a Hearing Aid for Bimodal Cochlear Implant Users: A Systematic Review: A Systematic Review". In: *Ear Hear.* 39.6, pp. 1039–1045.

Vroegop, J., N. C. Homans, M. van der Schroeff, and A. Goedegebure (2019a). "Comparing Two Hearing Aid Fitting Algorithms for Bimodal Cochlear Implant Users". In: *Ear Hear.* 40.1, pp. 98–106.

Vroegop, J., G. Dingemanse, M. van der Schroeff, and A. Goedegebure (2019b). "Comparing the Effect of Different Hearing Aid Fitting Methods in Bimodal Cochlear Implant Users". In: *Am. J. Audiol.* 28.1, pp. 1–10.

Wickens, T. D. (2002). *Elementary Signal Detection Theory*. Oxford: Oxford University Press, 3–37, 93–110, and 237–241.

Williges, B., M. Dietz, V. Hohmann, and T. Jürgens (2015). "Spatial Release From Masking in Simulated Cochlear Implant Users With and Without Access to Low-Frequency Acoustic Hearing". In: *Trends Hear.* 19, pp. 1–14.

Yoon, Y. S., Y. R. Shin, and Q. J. Fu (2013). "Binaural Benefit With and Without a Bilateral Spectral Mismatch in Acoustic Simulations of Cochlear Implant Processing". In: *Ear and Hearing* 34, pp. 273–279.

Zamaninezhad, L., V. Hohmann, A. Büchner, M. R. Schädler, and T. Jürgens (2017). "A physiologically-inspired model reproducing the speech intelligibility benefit in cochlear implant listeners with residual acoustic hearing". In: *Hear. Res.* 344, pp. 50–61.

Zeng, F. G., S. Rebscher, W. Harrison, X. Sun, and H. Feng (2008). "Cochlear implants: system design, integration, and evaluation". In: *IEEE Rev. Biomed. Eng.* 1, 115–142.

Zeng, F. (2002). "Temporal pitch in electric hearing". In: *Hear. Res.* 174.1–2, 101–106.

Zhou, J. and J. D. Durrant (2003). "Effects of interaural frequency difference on binaural fusion evidenced by electrophysiological versus psychoacoustical measures effects of interaural frequency difference on binaural fusion evidenced by electrophysiological versus psychoacoustical". In: *J. Acoust. Soc. Am.* 114, pp. 1508–1515.

Zhu, M., B. Chen, J. Galvin, and Q. Fu (2011). "Influence of pitch, timbre and timing cues on melodic contour identification with a competing masker (L)". In: *J. Acoust.Soc. Am.* 130.6, pp. 3562–3565.

Contributions to Hearing Research

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Vol. 36: *Søren Fuglsang*, Characterizing neural mechanisms of attention-driven speech processing, 2019.

Vol. 37: *Borys Kowalewski*, Assessing the effects of hearing-aid dynamic-range compression on auditory signal processing and perception, 2019.

Vol. 38: *Helia Relaño Iborra*, Predicting speech perception of normal-hearing and hearing-impaired listeners, 2019.

Vol. 39: *Axel Ahrens*, Characterizing auditory and audio-visual perception in virtual environments, 2019.

Vol. 40: *Niclas A. Janssen*, Binaural streaming in cochlear implant patients, 2019.

The end.

To be continued…

Cochlear implants (CI), a form of neural prostheses, are used to aid patients with severe-to-profound hearing loss by representing sounds in a way fundamentally different compared to acoustic hearing. The healthy auditory system constantly analyses the complex sound environment around us. It forms sound objects from acoustic components, that, when linked over time, can form a so-called stream, such as a voice or melody. Normal-hearing listeners perform this auditory streaming with their two ears, unaware of the binaural processes that integrate the information from both ears into common sound objects. Especially for bimodal CI patients with a CI in one ear and hearing aid (HA) on the other ear, one sound can be perceived very differently across ears. Bilateral CI patients, implanted in both ears, may also perceive sounds differently in each ear. For CI patients the potential differences in a sound's representation across ears may influence how they process binaural sounds. Understanding the CI patients' binaural streaming better could help to guide the development of CI-candidacy criteria, clinical fitting of the devices, and new strategies for their simulation. The studies in this thesis center on the question of whether bimodal CI patients can use their devices effectively together and whether they can build common streams from binaural sounds like normal-hearing listeners would.

DTU Health Tech
Department of Health Technology

Ørsteds Plads
Building 352
DK-2800 Kgs. Lyngby
Denmark
Tel: (+45) 45 25 39 50
www.dtu.dk