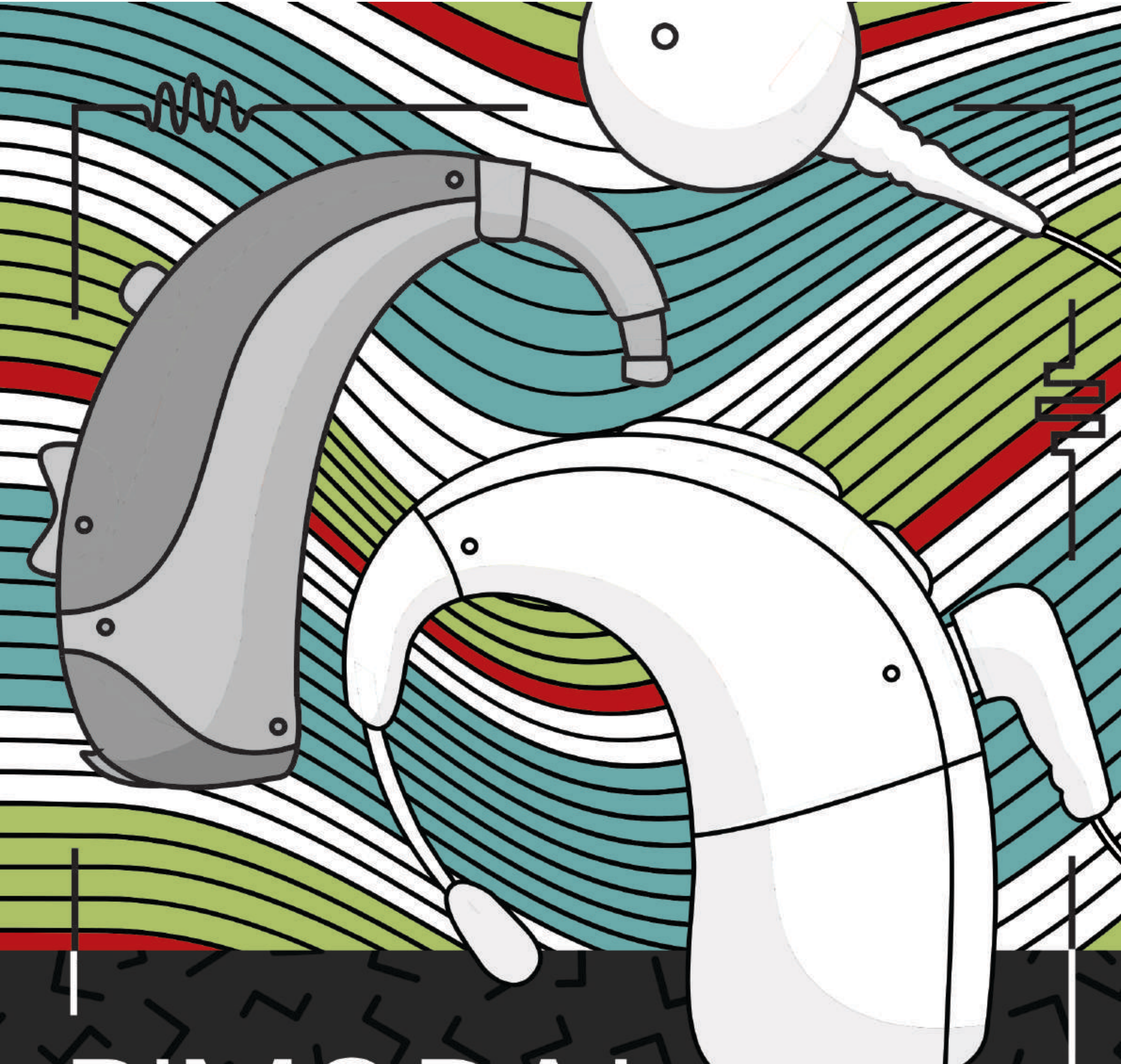

BIMODAL STIMULATION

TOWARDS BINAURAL INTEGRATION

Lidwien Veugen

DONDERS SERIES

BIMODAL STIMULATION

TOWARDS BINAURAL INTEGRATION

Lidwien Veugen

Printed by: Ipskamp Drukkers
ISBN: 978-94-6284-121-5

This thesis was financially supported by Advanced Bionics. The research presented in this thesis was carried out at the Department of Biophysics of the Radboud University Nijmegen and Hearing & Implants at the Donders Institute for Brain, Cognition and Behaviour in Nijmegen, the Netherlands.

BIMODAL STIMULATION

TOWARDS BINAURAL INTEGRATION

Proefschrift

ter verkrijging van de graad van doctor

aan de Radboud Universiteit Nijmegen

op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,

volgens besluit van het college van decanen

in het openbaar te verdedigen op

woensdag 1 november 2017

om 14:30 uur precies

door

Lidwien Catharina Elisabeth Veugen

Geboren op 13 december 1987
te Eindhoven

PROMOTOREN

Prof. dr. A.J. van Opstal
Prof. dr. ir. A.F.M. Snik

COPROMOTOREN

Dr. L.H.M. Mens
Dr. M.M. van Wanrooij

MANUSCRIPTCOMMISSIE

Prof. dr. R.J.A. van Wezel (voorzitter)
Prof. dr. P.W.M. Desain
Prof. dr. J.H.M. Frijns (Leiden UMC)

CONTENTS

CHAPTER 1

GENERAL INTRODUCTION

HEARING

THE AUDITORY SYSTEM

Our sense of hearing is mediated by the auditory system that converts acoustic waves (rapid pressure changes) into meaningful auditory information. The visible part of the ear, the pinna, directs incoming sound pressure waves into the ear canal towards the eardrum (Figure 1). There, the sound vibrations are passed to the chain of three small bones in the middle ear, called the ossicles (malleus, incus, and stapes, respectively), that together function as a lever-hydraulic system, which significantly amplifies the acoustic energy at the stapes. Their function is to transfer the amplified sound vibrations from the air (with its low acoustic impedance) to the fluid of the inner ear (with its very high acoustic impedance) via a membrane, called the oval window.

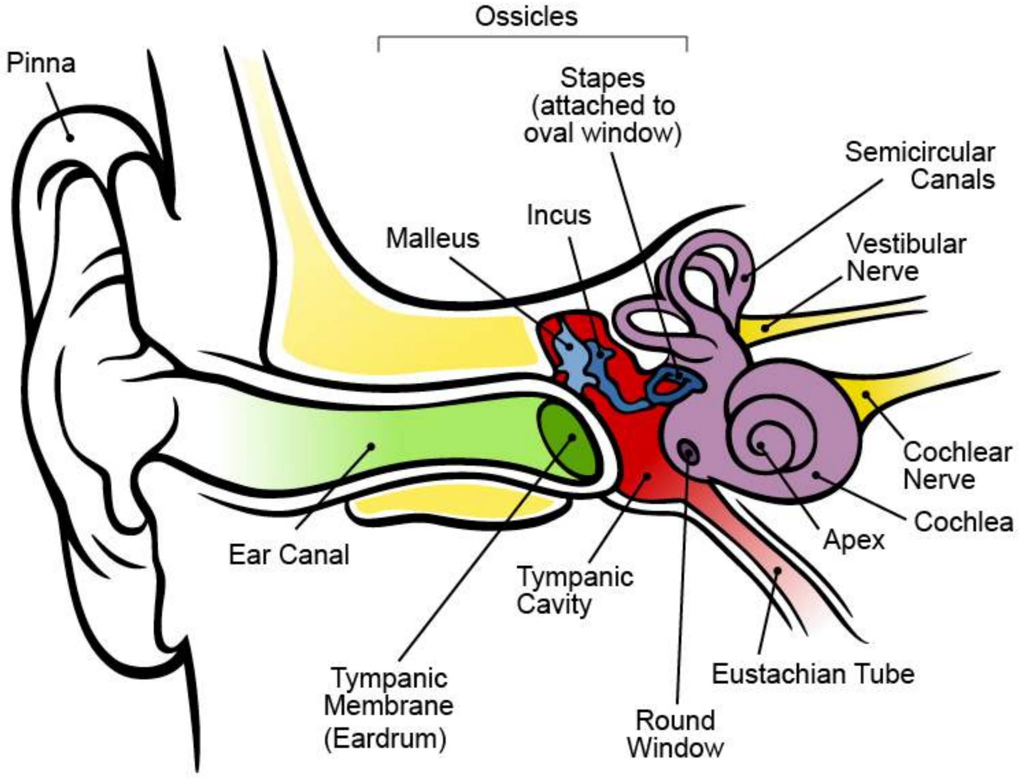

Figure 1. *Anatomical representation of the auditory system [Retrieved from Wiki Commons]*

The spirally-shaped inner ear (the cochlea), consists of three fluid-filled chambers (scala vestibuli, scala media, and scala tympani, respectively). Scala vestibuli and scala tympani are separated over the entire length of the cochlea by the scala media that contains Corti's organ (Figure 2A) and the basilar membrane (BM), but connect at the apex via the helicotrema, which acts as a pressure short-circuit. The BM vibrates in response to the time-varying pressure difference that is caused by the inward/outward motion of the oval and round windows as the stapes delivers its acoustic energy. This pressure difference decreases from high (at the oval/round windows) to zero (at the helicotrema), and causes a mechanical vibration of the elastic BM that is in resonance with the driving frequency of the sound. Due to the varying elastic properties of the BM across its length, and the coupling with the surrounding fluid, the pressure difference leads to a traveling wave along the BM that peaks at a frequency-specific location. In humans high frequencies up to 20 kHz peak at basal sites of the BM (near the oval window), low frequencies down to about 20 Hz at more apical locations, and mid-frequencies at intermediate locations. This response forms the basis of the auditory system's tonotopy. Although the relation between frequency and anatomic location originates in the cochlea, it is preserved in the different stages of the ascending auditory pathway, up to the core auditory cortex.

The BM motion activates the auditory receptors of the organ of Corti, the inner and outer hair cells (Figure 2A). The inner hair cells rectify and low-pass filter the oscillatory bending motion of their stereocilia, which is subsequently transformed into electrical impulses that are sent through the auditory nerve to the brain, where the information is subsequently processed. The outer hair cells effectively function as a precisely tuned dynamic and highly nonlinear mechanical feedback system. This mechanism further amplifies and sharpens the local pressure-induced vibration patterns of the BM at low sound levels, which underlies the extreme sensitivity of the auditory system to weak acoustic inputs (Figure 2B). At high sound levels this mechanical feedback is actively shut down (compressive attenuation), which allows the auditory system to respond over a tremendous range of acoustic input levels (about 120 dB).

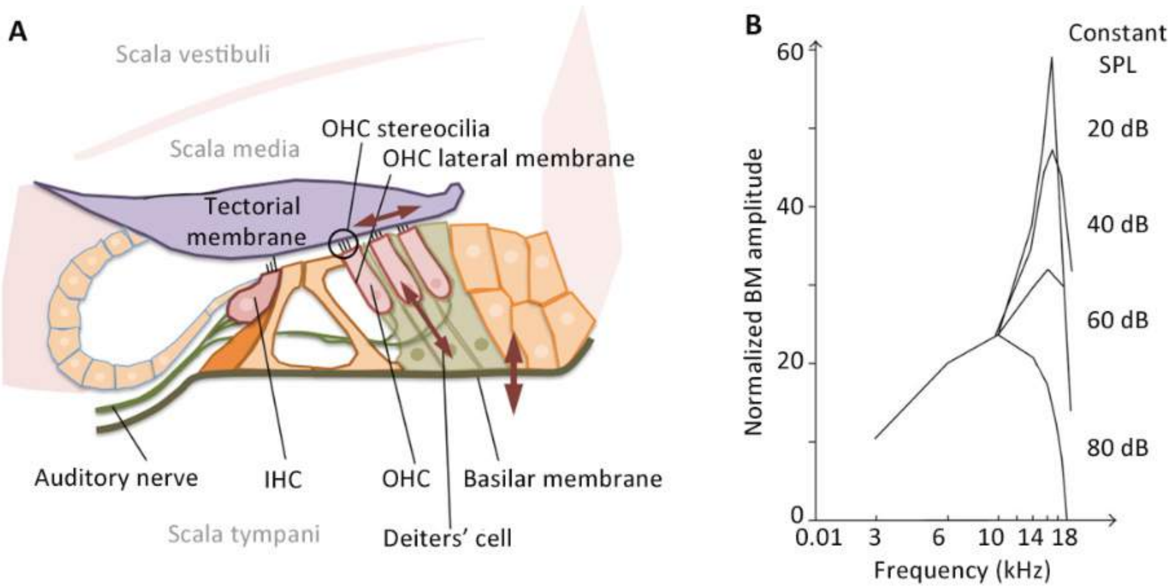

Figure 2. A) *Anatomy of the Organ of Corti in the cochlea containing the inner- (IHC) and outer hair cells (OHC).* ***B)*** *Result of feedback from the outer hair cells: the normalized amplitude of the basilar membrane (BM) over frequency for different sound levels. [image retrieved from Van Opstal (2016), with permission]*

LOUDNESS

Loudness, measured in phons, is a subjective measure that refers to the perception of a sound's intensity in dB or W/m^2. Both the number of recruited auditory nerve fibers, and their firing rate determine loudness perception. But also the sound's frequency, bandwidth, and its duration influence loudness. As described by the equal-loudness contours of Figure 3, the human ear is most sensitive to frequencies between 2-4 kHz (corresponding to the resonance frequency of the ear canal) and less sensitive at the extremes. The perception of loudness increases for sounds with broader bandwidths and with increasing sound pressure level, but not in a linear way. A doubling in loudness requires an increase of 10 dB, which corresponds to a tripling in actual sound pressure level. Hearing loss often affects loudness growth, which measures the dynamic range of acoustic sensitivity of the listener from 'very soft' to 'painfully loud' as function of acoustic power (see 'Hearing Impairments').

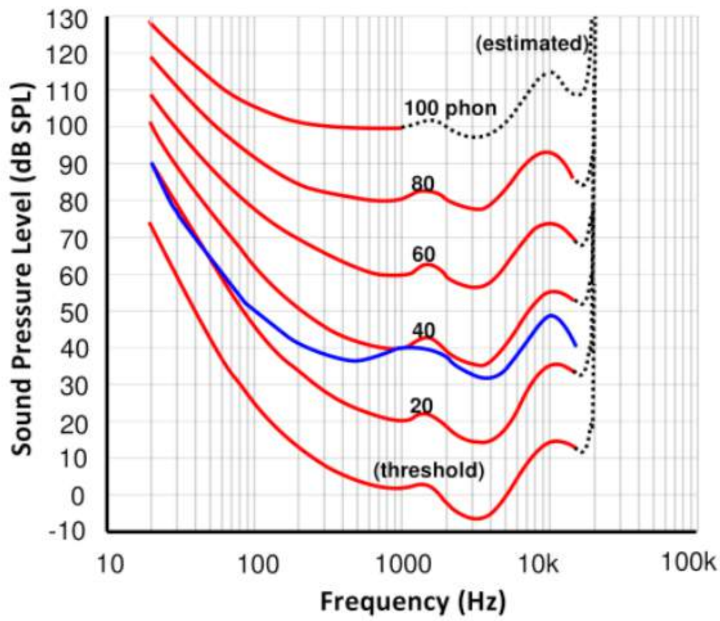

Figure 3. *Equal-loudness contours (red) from ISO 226:2003 revision and the original ISO standard for 40-phons (blue) [retrieved from Wiki commons]*

BINAURAL HEARING

The primary advantage of having two ears, apart from an increased signal-to-noise ratio, is the ability to determine the location of sounds, which can support sound-source segregation and speech understanding in cluttered, noisy environments (Blauert, 1997). Sounds from the left arrive earlier and louder at the left ear than at the right ear. These two acoustic cues are known as interaural time (ITD) and level differences (ILD), respectively, and are required to localize sound sources in the horizontal plane of the head (described by the azimuth angle). ILDs are created because the head acts as an obstruction, thereby causing an acoustic shadow (Figure 4A). The attenuation becomes larger and more useful with increasing frequency, resulting in ILDs up to 20 dB at 6 kHz (Shaw, 1974). At low frequencies, the sound's wavelength is of similar size as the head, and the waves can thus bend around the head, resulting in small ILDs. The path length difference from the source to both ears produces an interaural timing (ITD), and thus phase (IPD), difference. The latter depends on the sound's frequency and angle of incidence (Figure 4B). At high frequencies, IPDs become ambiguous as soon as the wavelength is smaller than the maximum path-length difference (phase ambiguity). Therefore, according to Rayleigh's duplex theory, timing differences are the dominant cue for low frequencies (< 1.5 kHz), whereas level differences dominate for the higher frequencies (Rayleigh, 1907). The brain is able to integrate these interaural differences to provide an accurate estimate of the sound's azimuth angle.

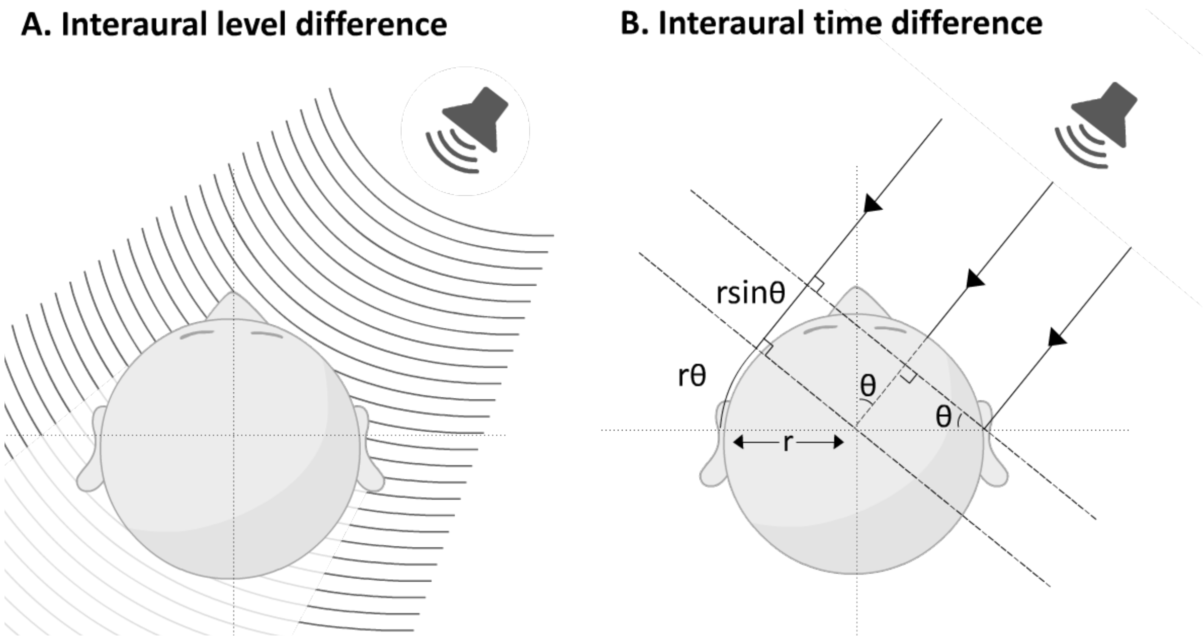

Figure 4. *Principles of interaural* ***A)*** *level difference for a high frequency sound, and* ***B)*** *time difference for a sound source with incident angle θ. The additional distance between both ears is r(θ +sinθ).*

Sound localization in the vertical plane of the head (the medial plane, in which directions are described by the elevation angle) is based on a very different mechanism, not requiring binaural hearing. The irregular shape of the pinna causes reflections and diffractions of incoming sounds that vary in a systematic, yet complex way with the elevation angle (Figure 5). Experiments have shown that the brain needs to learn the mapping of these cues to the vertical location of the sound source (Hofman et al., 1998; Van Wanrooij et al., 2005).

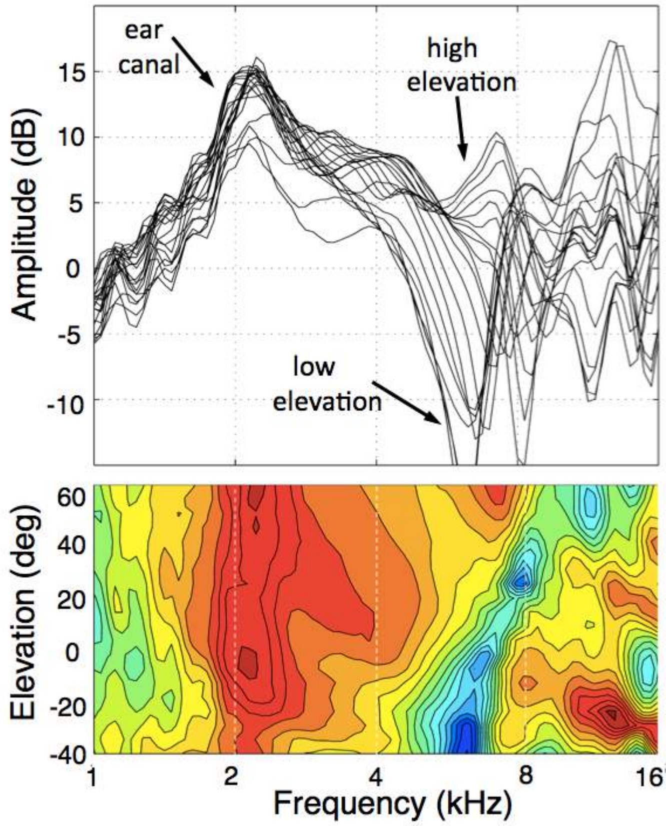

Figure 5. *Spectral cues as a function of frequency for sounds from the front with different elevation angles. Red means amplification, blue means attenuation. [image retrieved from Van Opstal (2016), with permission]*

HEARING LOSS

HEARING IMPAIRMENTS

Hearing problems with a peripheral origin can arise from different types of damage in the auditory system. Conductive hearing loss occurs when sounds are not efficiently transmitted from the outer ear to the eardrum and/or ossicles. Such hearing impairments can often be restored by hearing aids, or by bone conduction devices, that transfer sound vibrations to the cochlea, bypassing the middle ear. This thesis will mainly focus on a different type of hearing loss, called sensorineural hearing loss, which is caused by damage to the inner ear (cochlea).

Most frequently, sensorineural hearing loss is caused by damage to outer hair cells, leading to a dramatic degradation of hearing sensitivity (Figure 2B). As a result, sound information cannot be adequately transmitted by the auditory nerve, resulting in a distorted or impoverished perception. This type of hearing loss most often starts at the higher frequencies (at the base of the cochlea), but can increase and worsen progressively. Typical causes include aging, prolonged exposure to loud sounds or sudden explosive sounds (gun shots), infections, or genetic disorders. The damage is usually irreversible, but can often be corrected (partly) by amplification of the acoustic energy through an acoustic hearing aid.

Pure-tone audiometry is the gold-standard to measure the degree of hearing loss. It measures the softest sound and the loudest comfortable sound a person can hear by presenting tones at frequencies ranging from 250 Hz to 8 kHz. For normal hearing listeners these values are typically 0 and 120 dB hearing level (HL) respectively. The measured hearing thresholds are charted in an audiogram as a function of frequency (Figure 6). For hearing-impaired listeners, poor audibility of low-intensity sounds is often accompanied by an abnormally rapid loudness growth, causing discomfort for intense sounds. The range between the hearing threshold and the loudness discomfort level of the hearing-impaired ear is therefore considerably smaller when compared to normal hearing.

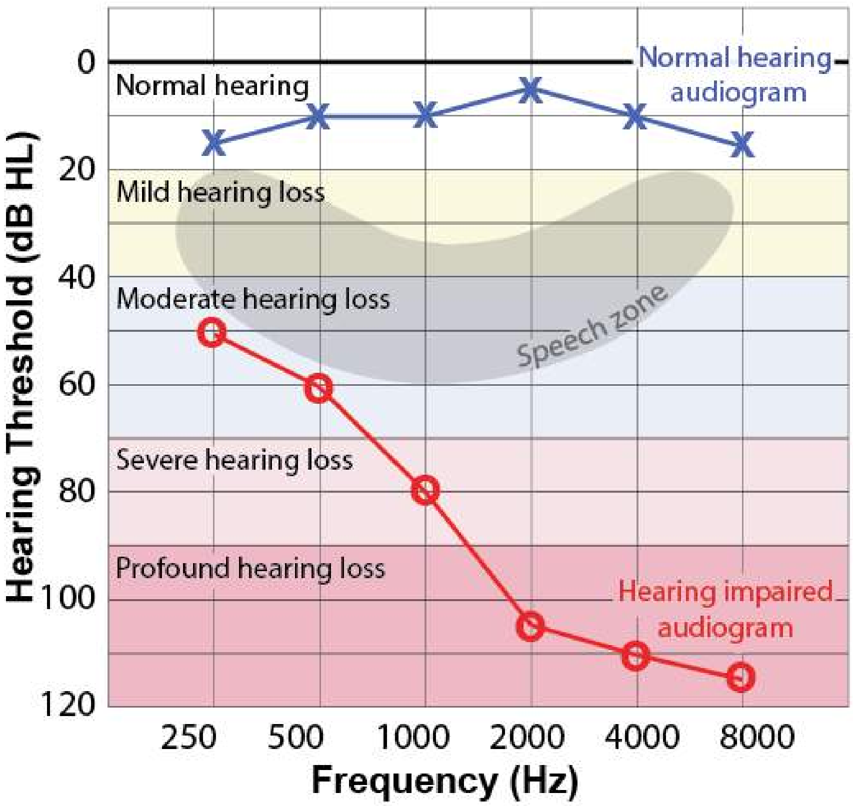

Figure 6. *Audiogram with classifications of hearing loss and with hearing thresholds of the softest audible tones for an ear with normal hearing and a hearing-impaired ear.*

HEARING AIDS

Hearing aids are available in a wide range of types and technologies that all aim to provide better audibility of incoming sounds without making them uncomfortably loud.

This thesis focuses on the most widely used hearing aids that aim to compensate sensorineural hearing loss.

Currently, hearing aids can be small enough to fit and hide entirely in the ear canal. However, the patient group suffering from severe to profound hearing loss, which participated in this research, typically used a digital behind-the-ear hearing aid (Figure 7A). The microphone on this hearing aid collects sounds from the environment, which are digitally processed and adjusted by a processor. The resulting amplified signal is then sent from the receiver (speaker), via a tube, to a custom-made ear mold inside the ear canal.

Hearing devices use nonlinear compression algorithms to squeeze the large range of input levels from the environment into the restricted dynamic hearing range of the patient. Besides this so-called automatic gain control, modern digital hearing aids offer complex signal processing with many different features like noise reduction, feedback cancellation, multiple programs and directional microphones. The HA is usually programmed by the audiologist to match the patient's needs and degree of hearing loss.

When hearing loss is too profound for amplification by the hearing aid, the patient may be considered as a candidate for cochlear implantation.

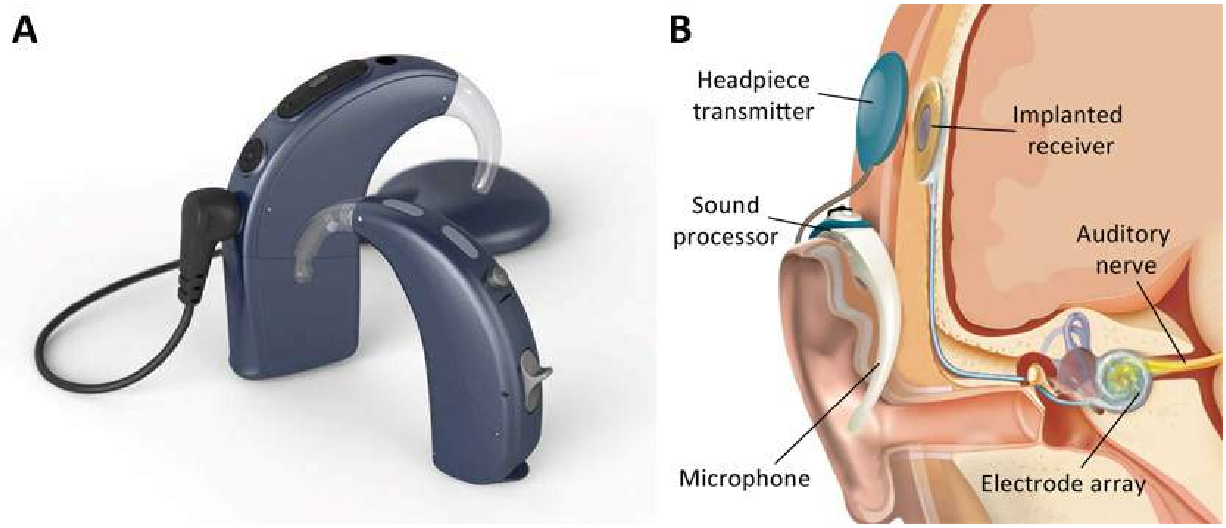

Figure 7. A) *Advanced Bionics cochlear implant processor (on the left) and a Phonak Naida hearing aid (on the right)* ***B)*** *Placement of the cochlear implant with the electrode array in the cochlea [used with permission of Advanced Bionics]*

COCHLEAR IMPLANTS

In most cases of sensorineural hearing loss, the auditory nerve is still intact. This opens possibilities for restoring hearing through direct electrical stimulation of this nerve, bypassing the damaged hair cells in the cochlea. The first human trials exploring this technique took place around the early sixties (Wilson et al., 2008). Twenty years later, in the mid-1980s, cochlear implantation became an approved treatment for deafness (House et al., 1991).

To date, a cochlear implant (CI) consists of an external behind-the-ear processor connected to a headpiece coil, and an implanted internal component (Figure 7B). During surgery, a receiver is placed on the skull bone above the ear to hold the external headpiece in place through a magnet. The receiver is connected to an array containing 16-22 electrodes that is inserted into the cochlea (typically, through the round window), to provide direct electrical stimulation of the auditory nerve.

Incoming sound waves are picked up by the microphone and sent to the CI processor for digital signal processing. This mainly involves a band-pass filterbank that separates the sound in different frequency bands, equal to the number of implanted electrodes. The energy in each frequency band is estimated at each time instant, to determine the stimulus strength of the corresponding electrode. Note that this frequency-electrode relationship mimics the natural tonotopic organization of the cochlea. Using a radio-frequency signal, the processed information is passed from the transmitting headpiece coil to the implanted receiver, which delivers current pulses to the appropriate electrodes in the electrode array in the cochlea. This way, sounds are sent as coded impulses to the brain, where they are interpreted as meaningful sounds.

During therapy after CI implantation, the brain should adapt to the new perceptual situation. Cochlear implantation in post-lingual deaf subjects generally results in remarkably good speech understanding, up to monosyllabic word scores of 55% (Krueger et al., 2008), regardless the brand. Because of these successful outcomes, CIs become more and more an option for hearing impaired individuals. At present, about 7000 people in the Netherlands use a cochlear implant (OPCI, 2016).

Like in many countries, bilateral implantation was recently approved for children up to 18 years of age by the Dutch health insurance systems. Strict inclusion criteria do not exist, but a second CI is generally provided when it is expected to give more benefit compared to a conventional contralateral hearing aid. Other countries, including Germany, Austria, Switzerland, Spain and Sweden also offer bilateral cochlear implantation as a standard therapy for adults (Vickers et al., 2016; www.zorginstituutnederland.nl, 2014). In some countries (like Australia) cochlear implantation even became a treatment option for single-sided deafness. In The Netherlands there has been a long debate about the cost-effectiveness for a second CI and the additional benefit compared to a conventional contralateral HA. Several studies have shown a benefit of bilateral CIs compared to a single CI for sound localization and speech understanding in noise (Brown et al., 2007), but the average costs for cochlear implantation in the Netherlands are 80.000 Euro, including lifelong aftercare (www.opciweb.nl, 2012).

In this thesis, we worked with CI processors from Advanced Bionics, one of the leading developers for CI systems, besides Cochlear, MED-EL and Oticon. Although the general concept is the same for all manufacturers, differences reside in electrode design, sound signal processing algorithms and stimulation schemes.

BIMODAL STIMULATION

With expanding inclusion criteria for cochlear implantation, more and more CI users can benefit from 'bimodal stimulation'. In this situation, the listener uses a CI in one ear and an acoustic hearing aid in the non-implanted contralateral ear. In the Nijmegen Audiological Center, about half of the postlingual adult CI recipients use a conventional HA in the opposite ear, one year after implantation. Several studies showed that bimodal listeners often benefit from the complementary types of information delivered by both devices, as well as from the fact that sounds become available at both ears, often resulting in improved speech, voice pitch and music perception (for a review see (Ching et al., 2007)).

The CI is often the most dominant device in bimodal stimulation as it conveys a large range of input frequencies, from 250 Hz to 8 kHz for the Advanced Bionics CI processor used in this thesis. However, the CI's spectral resolution is thought to be poor as a result of the band-pass filtering mechanism. Residual hearing in the non-implanted ear is typically reduced to the low frequencies (< 1 kHz). Therefore, the HA predominantly provides low-frequency acoustic information, but with higher spectral resolution and typically below the lowest frequency transmitted by the CI. The complementarity of sounds from the CI and HA brings additional information, but could possibly also hamper the integration of both sounds as useful listening cues. Below, an overview of the possible benefits and limitations in bimodal stimulation is described.

BENEFITS

CI users with an additional HA in the non-implanted ear often experience better speech understanding in quiet, but especially in background noise (Ching et al., 2007). An important cue in these complex listening situations is voice pitch, that can help segregating different auditory sources (Brokx et al., 1982). Therefore, the HA could possibly provide important information since it conveys the low frequencies containing the speaker's fundamental frequency and (first) formants, that are not well transmitted by the CI. Several studies confirmed this theory, showing that bimodal benefit for speech understanding is at least partly the result of a better audible fundamental frequency through the HA (Qin et al., 2006; Zhang et al., 2010). Probably because of this reason, bimodal listeners also experience better melody recognition and overall sound quality (Kong et al., 2005; Sucher et al., 2009).

The amount of benefit for speech understanding in noise largely depends on the spatial configuration of target speaker and noise sources. Besides the benefit from the complementary types of information, listening with two ears can give advantages through different mechanisms (Ching et al., 2007; Van Deun et al., 2010). The brain usually compares the input across both ears, resulting in an overall better perception, even when both ears receive similar information, e.g. from a frontal speaker, known as the binaural redundancy or summation effect. In the case of spatially separated sources, part of the sounds are attenuated by obstruction of the head (head shadow), creating a favorable signal-to-noise ratio at the ear contralateral to the noise source. In that case, 'glimpsing' or 'listening in the dips' of a fluctuating noise signal could also result in better speech understanding. The squelch effect that involves the perception of interaural time or level differences can also improve speech perception, if available in bimodal stimulation. The different mechanisms and the complementarity of CI and HA inputs can add up to a total bimodal benefit of 1 to 3 dB (Ching et al., 2007), corresponding to 7 to 21 % better speech perception (van Hoesel, 2012).

Studies that investigated sound localization in bimodal listeners found significantly better performance with both devices compared to the CI alone in about half of the subjects (Ching et al., 2007). On average, only a modest improvement in root-mean-squared error was found for bimodal hearing, with localization errors around 30° for the best performers reported (van Hoesel, 2012). Unlike benefits in speech understanding, sound localization of short stimuli is fully based on the central auditory processing of interaural time and level differences. However, most studies used longer sounds like words that can introduce other localization cues, e.g. because of head movements. It therefore remains the question if bimodal listeners can use binaural cues (discussed below).

Even though many CI users benefit from a contralateral HA, the actual improvement in performance varies considerably across subjects. In some cases, the HA can even introduce negative interference (Mok et al., 2006). A clear predictor for the amount of bimodal benefit has so far not been found. A number of potential causes have been suggested, briefly summarized here. First, and most likely, aided and unaided hearing thresholds play a role (Cullington et al., 2011; Tyler et al., 2002), but so far this has not been confirmed on a larger scale (Ching et al., 2005; Luntz et al., 2005). Second, the duration of bimodal experience has been considered, but it does not seem to explain the amount of bimodal benefit in a given listener (Ching et al., 2007). Third, there is a relationship between bimodal benefit and the difference in performance between the unimodal devices, with smaller differences more likely resulting in better interaural integration (Yoon et al., 2015). Finally, the pitch match across ears could play a role in the success rate of bimodal hearing (Reiss et al., 2014b).

LIMITATIONS

An important question is whether bimodal listeners can use binaural cues. In order to perceive interaural differences, the CI and HA should transmit an overlapping range of frequencies for either ear, as the binaural comparison in the auditory system is frequency-specific. This range is typically restricted by the amount of residual hearing in the non-implanted ear, which is often confined to the lower frequencies (< 1 kHz). Even though loudness-growth differs for the acoustically and electrically stimulated ear (Blamey et al., 2000), the brain may be sufficiently plastic to still extract (and learn to interpret) interaural level differences (Francart et al., 2008). However, ILDs are best pronounced in the higher frequencies (> 1.5 kHz) that often fall beyond the residual low-frequency hearing in the non-implanted ear. It is unknown whether bimodal listeners can learn to perceive ILD cues in the lower frequencies, when there is spectral overlap. Note that because signal-processing in the CI destroys the temporal fine-structure, interaural time differences are thought to be lost in bimodal stimulation (Francart et al., 2011).

Bimodal listeners can possibly compare the temporal onset or envelope of sounds across both ears to derive information about sound-source direction (Francart et al., 2009; Henning, 1974). However, signal processing and type of stimulation are completely different in a CI and HA, resulting in different processing delays (Francart et al., 2011). The HA used in this thesis (Phonak Naida S IX UP) typically needs 6 ms to process incoming sounds (based on own measurements). On top of that, propagation time of the traveling wave in the cochlea adds an additional 5-20 ms, depending on the frequency. The CI has a fixed delay between sound capture and stimulation of 15 ms for the Advanced Bionics CI processors used in this thesis (Chalupper et al., 2014). In bimodal stimulation, the total processing delay thus differs between both ears and even varies per incoming frequency. This probably challenges the perception of possible interaural timing cues, although the brain may be able to adapt to these unnatural delays.

Besides the processing delay, several other steps in the signal processing of sounds could create a mismatch across the CI and HA. Modern hearing devices are equipped with numerous features, including noise reduction systems, different microphone modes and gain and frequency compression programs. Until recently, CIs and HAs have typically been developed independently, not optimized for combined use. Most combinations of bimodal devices do therefore differ in sound processing characteristics and options. Besides an uncomfortable listening situation, this likely impairs the integration of interaural differences for sound localization and improved speech perception.

In contrast to the HA side, the natural place-frequency relationship in the cochlea may not be preserved in the CI ear. With a CI, the place of stimulation in the cochlea is determined by the insertion depth of the electrode array that often does not reach the apex. Low frequencies are therefore encoded at locations in the cochlea that used to be tuned to higher frequencies. As a result, incoming sounds that are processed by the CI and HA activate different places in the cochlea, resulting in a different pitch. This could possibly disturb the fusion of CI and HA input and the perception of interaural cues. However, it has been shown that over time, the perceived pitch in the CI ear of several bimodal listeners shifted in frequency, reducing the interaural pitch mismatch (Reiss et al., 2015).

SCOPE OF THIS THESIS

Because of its success, cochlear implants are offered at earlier and earlier stages of hearing loss (Gifford, 2011), leading to a growing interest in bimodal stimulation over the last years. However, currently available CIs and HAs are not optimized for combined use yet, likely resulting in mismatches across both ears. Also from a clinical perspective, well-validated fitting guidelines for bimodal devices do, so far, not exist. These problems could also underly the large variability in benefit seen across bimodal listeners. At the same time, it opens a world of possibilities in bimodal research.

The aim of this thesis was to assess and optimize bimodal benefit, and possible binaural cues, by adjusting the settings and programs of the HA, to create a better match with the CI. This research idea arose from the scientific view of the Biophysics and Hearing & Implants research group in Nijmegen together with the bimodal interests of Sonova, holding HA company Phonak and CI manufacturer Advanced Bionics. Fortunately, we found a group of 15-20 bimodal listeners that were willing to participate in a large set of experiments, stretching out over approximately two years.

In *Chapter 2* the first important goal in bimodal stimulation was investigated, intended to create an equal loudness percept at both ears. Besides a more comfortable listening situation, an adequate loudness balance could possibly facilitate the perception of binaural cues, aiding sound localization and speech perception in noise. Loudness across the CI and HA was balanced in three separate frequency bands to avoid over-amplification in the frequency range of residual hearing. This frequency-dependent loudness balancing procedure was compared to a simple broadband balancing procedure, reflecting current clinical practice.

After creating a balanced loudness for stationary sounds based on subjective loudness judgments, the next goal was to maintain this balance for fluctuating speech sounds. Given the restricted dynamic range of hearing in our bimodal subject group, automatic gain compression played an important role in the processing of incoming sounds. This could possibly induce distortions between the percept at both ears, since the CI and HA typically use other compression systems, differing in knee point, speed of reaction and the number of compression channels. In *Chapter 3*, we succeeded to improve speech understanding in noise by matching the compression settings of the HA to the CI.

It is unsure whether the distinct types of input from the CI and HA can be successfully integrated in the brain in order to perceive interaural difference cues. In bimodal research it is therefore often not clear if benefits arise from the complementary types of information from both devices or from real binaural cues. An appropriate task to separate these two is horizontal sound localization, since this is solely based on binaural cues (for short sounds in a dark environment). *Chapter 4* investigated sound localization in our group of bimodal listeners to test the availability and contribution of interaural time and level difference cues.

A prerequisite for the perception of interaural differences is that the information transmitted by both devices overlaps in frequency. However, the majority of our bimodal subject group only had residual hearing in the lower frequencies, where interaural level differences are small. The aim of *Chapter 5* was to enlarge the range of spectral overlap by a novel frequency lowering algorithm in the HA, compressing high frequency information, beyond residual hearing, into the audible range of hearing. This was hypothesized to enhance interaural level difference cues, which was tested again in a sound localization task, after subjects acclimatized to the new HA program.

Better speech perception is often the most important motivation to use a hearing aid. For this reason, research outcomes and clinical assessments are usually expressed as a certain improvement or decrement in speech understanding. This often involves a percentage words correct or signal-to-noise ratios at criterion performance. However, speech scores could possibly be influenced by the limited set of words of the test, the cognitive ability of the subject and his linguistic knowledge. Therefore, *Chapter 6 and 7* describe an alternative performance test using non-linguistic stimuli, ripples, that contain modulations over time and over frequency, as present in speech sounds.

In *Chapter 6,* listening performance using ripple stimuli was evaluated in normal hearing listeners under normal hearing conditions, as well as under hearing impaired conditions, using simulations of a hearing aid and a cochlear implant. Thereafter, in *Chapter 7*, we systematically tested ripple detection in bimodal listeners for a large set of spectral, temporal and spectro-temporal modulations. Performance of both monaural listening conditions was compared with bimodal stimulation, to investigate the bimodal benefit for various modulations. Furthermore, we correlated ripple results with speech perception performance.

Lastly, *Chapter 8* summarizes and discusses the main results in this thesis.

CHAPTER 2

FREQUENCY-DEPENDENT LOUDNESS BALANCING IN BIMODAL COCHLEAR IMPLANT USERS

Published as *Veugen, L.C., Chalupper, J., Snik, A.F., van Opstal, A.J. and Mens, L.H., 2016. Frequency-dependent loudness balancing in bimodal cochlear implant users. Acta oto-laryngologica, 136(8), pp.775-781*

ABSTRACT

CONCLUSION. In users of a cochlear implant (CI) and a hearing aid (HA) in contralateral ears, frequency-dependent loudness balancing between devices did, on average, not lead to improved speech understanding as compared to broadband balancing. However, nine out of fifteen bimodal subjects showed significantly better speech understanding with either one of the fittings.

OBJECTIVES. Suboptimal fittings and mismatches in loudness are possible explanations for the large individual differences seen in listeners using bimodal stimulation.

METHODS. HA gain was adjusted for soft and loud input sounds in three frequency bands (0 - 548, 548 - 1000 and >1000 Hz) to match loudness with the CI. This procedure was compared to a simple broadband balancing procedure that reflected current clinical practice. In a three-visit cross-over design with four weeks between sessions, we tested speech understanding in quiet and in noise and administered questionnaires to assess benefit in real world.

RESULTS. Both procedures resulted in comparable HA gains. For speech in noise, a marginal bimodal benefit of 0.3 ± 4 dB was found, with large differences between subjects and spatial configurations. Speech understanding in quiet and in noise did not differ between the two loudness balancing procedures.

INTRODUCTION

Beneficial effects for the combined use of a cochlear implant (CI) and a contralateral hearing aid (HA) have repeatedly been reported (for review see: Ching et al., 2007). Advantages provided through such 'bimodal' stimulation can include improved speech understanding in noise, voice pitch perception, localization abilities and music perception (Litovsky et al., 2006; Straatman et al., 2010; Tyler et al., 2002).

Despite the overall positive findings, large individual differences in bimodal benefit have been found (Ching et al., 2007; Litovsky et al., 2006; Straatman et al., 2010). The factors that underlie these differences are not fully investigated, but may include, for example, tonotopical mismatches, degree of hearing loss and binaural fusion, sub-optimal fittings and mismatches in loudness. An essential part of fitting is to obtain equal loudness percepts at both ears, but no well-validated bimodal fitting procedure is available to achieve this. Francart and McDermott (2012a) have proposed a model to normalize loudness through the CI and HA, aiming to achieve a loudness that is comparable to that in normal hearing. In the present study, we aimed at the more modest goal of achieving balanced loudness between both ears considering that, even if true loudness normalization cannot be achieved, the brain may still be plastic enough to extract binaural cues if loudness is balanced across input levels and frequencies (Dorman et al., 2014a).

Ching (2007) formulated a recommendation to achieve loudness balance between both devices in bimodal listeners for soft and loud input sounds, based on a pairwise comparison procedure with different HA frequency responses. However, other studies (Blamey et al., 2000; Hoth, 2007) have shown that not only the dynamic range differs between acoustically and electrically stimulated ears of bimodal listeners, but also the shapes of iso-loudness curves. Therefore, ideally, equal overall loudness would not only be achieved for soft and loud complex sounds, but for all signals across frequency bands.

Residual hearing in bimodal listeners is typically limited to the low frequencies, while most CI devices encode signals between 250 and 8000 Hz. Thus, adjusting HA gain to simply match the overall loudness of a broadband signal may result in loudness mismatches in frequency bands that are transmitted by only one device. In cases of low frequency residual hearing, over-amplification of the low frequencies by the HA is a likely result and a low frequency signal may be more appropriate for balancing. Ching et al. (2001b) balanced loudness in bimodally fitted children using a 65 dB SPL warble tone at 500, 1000 and 2000 Hz next to a paired comparison of several frequency responses deviating from the NAL prescription. Their findings highlight the importance of individual HA fine-tuning for optimized bimodal benefit, but the exact contribution of loudness balancing at different frequencies remained unclear.

In the present study, we investigated the effect of loudness balancing in three separate frequency bands: (1) low frequencies up to 500 Hz that contain voicing cues and complement information provided by the CI; (2) the middle frequency band (500 Hz – 1 kHz) that potentially contains segmental information; (3) the third band with frequencies above 1 kHz. We aimed to compare this frequency-dependent fitting procedure with a simple broadband loudness balancing procedure that reflected current clinical practice. A possible improvement in bimodal benefit of the new fitting method over the conventional approach was examined by testing speech understanding in quiet and in noise, and through subjective judgments.

METHODS

SUBJECTS

Fifteen postlingual deaf subjects (10 male; mean age 61 ± 12 years) participated in the study (Table 1). All used a Harmony speech processor (Advanced Bionics, Valencia, CA) in one ear, for at least a year. Subjects were selected with thresholds in the non-implanted ear better than 110 dB HL at 500 Hz. To eliminate variability due to devices, all subjects were fitted with one and the same type of compressive HA (Naida S IX UP, Phonak, Stäfa, Switzerland). However, the loudness balancing procedures investigated here were not specific for this product pair, but aimed to contribute to the general knowledge of bimodal fitting. Only six out of the 15 subjects were already bimodal users at the start of this study, but then used other types of HAs. The study was approved by the Local Ethics Committee of the Radboud University Nijmegen (protocol number 40327.091.12).

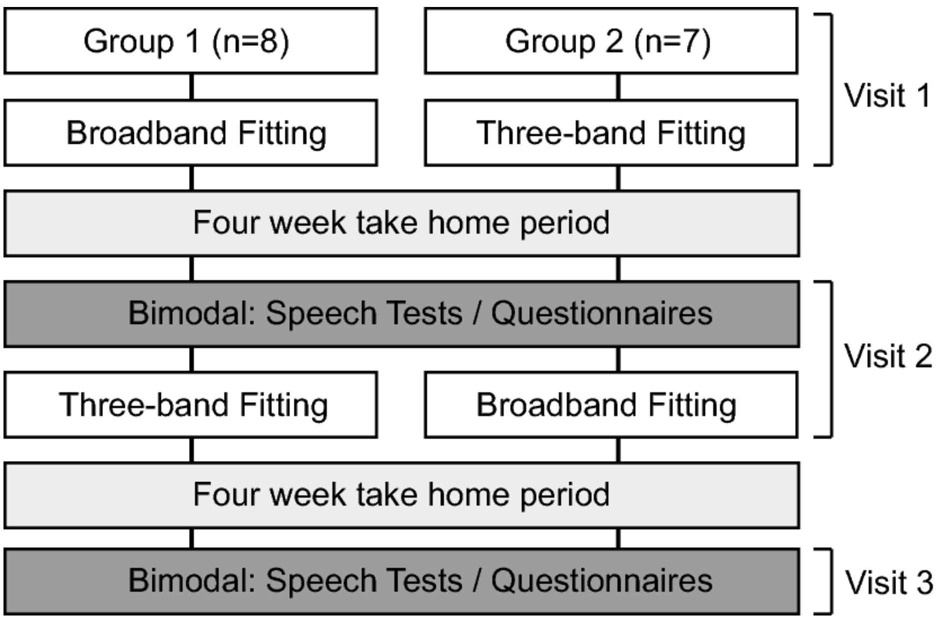

Figure 1. *Overview of the cross-over design of the study.*

Table 1. *Demographic details of the participating subjects. The 'group' column indicates the distribution of the subjects over the cross-over design with group A starting with broadband balancing and group B starting with three-band balancing. The asterisk marks experienced bimodal users, indicating that HA use overlaps with the full amount of CI experience.*

Participant	Age (yr)	Gender	Implanted ear	Etiology	Bimodal user	CI experience (yr)	HA experience non-implanted ear (yr)	Group
1	63	M	L	Unknown	Y	6.3	9*	A
2	46	M	L	Genetic	Y	3.7	16*	A
3	79	F	L	Heriditary	N	8.9	26	A
4	58	M	R	Ototoxicity	Y	5.3	50*	B
5	74	M	R	Genetic	Y	1.3	28*	A
6	52	F	R	Genetic	N	5.6	27	B
7	44	M	R	Congenital	Y	3.8	32*	B
8	75	M	L	Unknown	N	1.1	18	B
9	64	M	R	Unknown	N	2.8	20	A
10	78	F	R	Unknown	N	6.8	10	B
11	64	F	R	Heriditary	Y	5.6	30*	A
12	58	M	R	Genetic	N	4.8	32	B
13	42	M	L	Perinatal	N	9.9	28	A
14	70	M	R	Meniere's	N	2.3	13	B
15	53	F	R	Heriditary	N	4.3	9	A

Note. CI=cochlear implant; HA=hearing aid

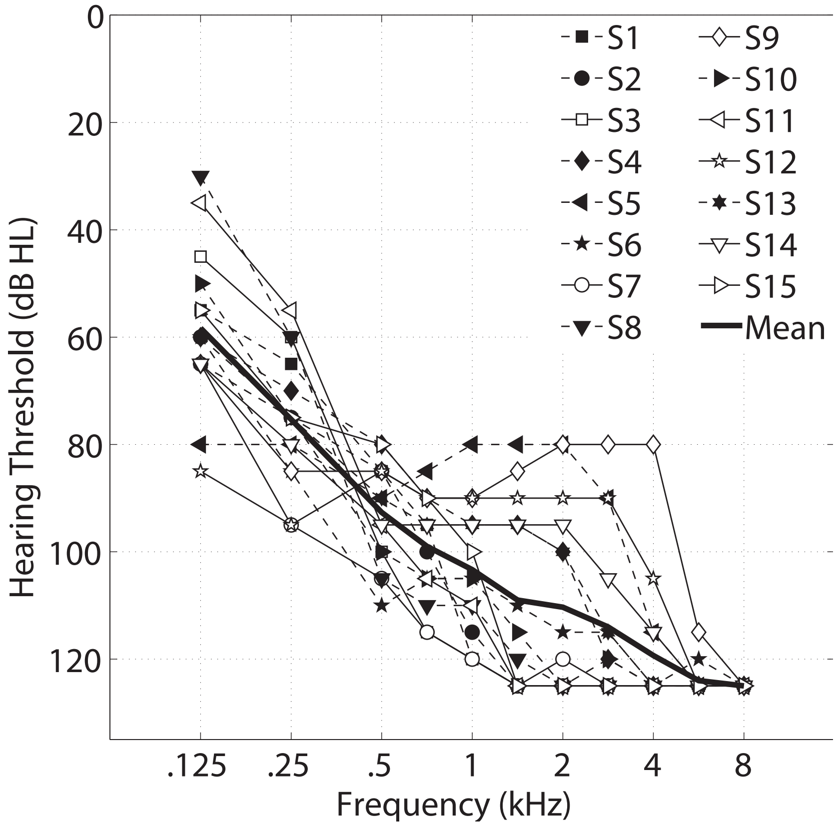

Figure 2. *Mean and individual hearing thresholds in the non-implanted ear as assessed with standard pure tone audiometry. Subjects with black markers started with the broadband fitting and subjects with white markers started with the three-band fitting. Thresholds beyond the audiometer limit (120 dB HL) were assigned a value of 125 dB HL.*

STUDY DESIGN

We compared a new bimodal hearing aid fitting protocol, which is referred to as 'three-band fitting', with the current clinical standard, referred to as 'broadband fitting'. The study consisted of a three-visit cross-over design with four weeks between sessions (Figure 1). Subjects were randomly distributed over the two arms of the cross-over study in two groups without significant differences in hearing thresholds (Figure 2), age, gender, post-implant duration and side of implant. Visits were scheduled on average 35 ± 6 days apart, with a minimum of 26 days, to allow for HA acclimatization during a take-home period.

SETUP FREE FIELD

Sound stimuli for the broadband fitting and speech understanding tasks were delivered through an external soundcard (RME Babyface, Audio AG, Haimhausen, Germany), a main amplifier (MPA 4-80, Ecler, Spain) and JBL loudspeakers (Control 1, Harmon International Industries, Washington DC, USA). Free field presentation levels were calibrated using a sound level meter (Brüel & Kjaer 2260 investigator) at the position of the subject's head.

HEARING AID FITTING PROCEDURES

GENERAL SETTINGS

During the whole study, all adaptive features and settings were turned off in the CI and HA except for the Naida's adaptive feedback canceller in case of feedback problems. The CI was left fixed at the subjects' default everyday program and volume setting for both loudness balancing procedures. All HAs were fitted using in situ pure-tone audiometry through the HA. The HA fitting software (Phonak Target 3.0) allows to adjust HA gain in twenty small frequency bands. After loudness balancing, the HA gains in these bands were read out from the fitting software to compare both procedures. Both fitting procedures could be completed within a clinically acceptable 10-15 minutes and are described below.

BROADBAND FITTING

The broadband fitting procedure was based on local clinical practice. Subjects were seated in a sound-treated room, one meter in front of a loudspeaker. Nonsense natural running speech, adapted from the International Female Fluctuating Masker (IFFM) (EHIMA, Retrieved 17 April, 2013) was presented at 65 dB SPL. The fundamental frequency of the IFFM signal was lowered to 127 Hz, in-between a male and female voice (IFFM'), and a three-talker signal was created by overlaying different parts of the signal to make it slightly more steady-state (IFFM3'). Overall gain of the HA was increased or decreased, according to instructions of the subject, who was asked to find a balanced loudness with the CI, while listening to IFFM3'. Starting point was the conventional Phonak fitting rule (Adaptive Phonak Digital) (Phonak, 2013).

THREE-BAND FITTING

The three-band fitting involved loudness balancing at two intensities (40 and 80 dB SPL) and in three frequency bands that could be adjusted separately in the HA fitting software: low (< 548 Hz), middle (548 to 1000 Hz) and high frequencies (> 1000 Hz). Loudness balancing was performed using a speech-shaped steady state noise, filtered according to the three frequency bands. Noise bursts of 1.5 seconds with ramps of 40 ms were presented alternately to the CI and HA via direct audio input (DAI), with an inter-stimulus interval of 0.6 seconds. Because of the short stimulus duration, steady-state noise was chosen instead of a speech signal to avoid large variations in loudness. Starting point was a novel fitting formula for bimodal fitting, which reduced gain to zero if the hearing loss exceeded 120 dB HL (Chalupper et al., 2015).

The HA fitting software allows for separate gain adjustments at three input levels: 40, 60 and 80 dB SPL. The HA gain for the soft ("40 dB" in the HA fitting software) and loud level ("80 dB") was adjusted separately per frequency band, until the subject reported to perceive a loudness balance across both ears. Loudness balancing was conducted sequentially from low to high bands and from soft to loud levels. To achieve a loudness balance for the 80 dB SPL input signal, the compression ratio of the HA was altered. Because the compression knee-point was set between 30 and 50 dB SPL input, depending on the hearing loss, this could in some cases affect the balance for soft sounds. When a loudness balance was achieved for all bands and levels, the IFFM3' speech signal at 65 dB SPL was used in free field to check for a loudness balance using microphone input, which was always confirmed by the patient. Subjects were instructed not to use the volume control on their HA or CI during their take-home testing period.

To assess bimodal benefit after the take-home acclimatization period, measurements were performed as described below.

SPEECH TESTS

SETUP

Subjects were seated in a sound treated room with low reverberation. Three loudspeakers were positioned at a distance of one meter from the patient's head at 0°, +90°, and -90° azimuth. Target words and sentences were always played via the loudspeaker directly in front of the subject at 0°.

For testing speech understanding in quiet, the NVA (Nederlandse Vereniging voor Audiologie) Dutch monosyllabic word test was used (Bosman, 1992). Words were presented at 65 dB SPL in three conditions: CI-only, HA-only and CI+HA (bimodal). Per condition, three lists were presented and subjects were instructed to repeat the words as accurately as possible. The mean percentage phonemes correct was transformed to rationalized arcsine units (RAU) (Studebaker, 1985) for statistical analysis.

Speech understanding in noise was tested with the Leuven Intelligibility Sentence Test (LIST) (van Wieringen et al., 2008). Two types of noise were used: the speech-weighted noise provided by the LIST sentences, and the IFFM' signal, referred to as single-talker noise. Noise was presented either from the front (S0N0), from the implanted side (S0NCI), the non-implanted side (S0NHA) or simultaneously from +90° and -90° (S0N±90). The S0N±90 configuration was tested with both noise types; S0N0, S0NCI and S0NHA were only tested with single talker noise. All noise configurations were assessed under two listening conditions, CI+HA and CI-only, apart from S0N0 which was only evaluated for CI+HA.

The presentation level of the target sentences was held constant at 65 dB SPL, while the noise level was varied according to the standard scoring procedure of the LIST (van Wieringen et al., 2008). The first sentence of each list was presented at a 0 dB signal to noise ratio (SNR) and was repeated with decreasing noise level until correctly identified by the subject. Then, the noise level was changed adaptively in steps of two dB to obtain the 50% speech reception threshold (SNR50), which was calculated as the mean of the last six SNRs (including one level beyond the last presented). Two lists per condition were tested and the mean SNR50 was our final outcome measure.

The CI-only measurement was performed after another study in our lab with the same subjects, two months after the end of this study. Bimodal and HA-only conditions were measured for all sessions after the four weeks acclimatization period. Bimodal benefit was calculated by subtraction of the bimodal SNR50 from the CI-only SNR50, with higher values indicating more benefit. Listening conditions were always presented in random order within sessions.

QUESTIONNAIRES

For each fitting, subjective evaluations were collected using the Speech, Spatial and Qualities of hearing scale (SSQ) questionnaire (Gatehouse et al., 2004). In addition, every week during the four-week take home period, subjects were asked to fill in seven basic questions concerning everyday listening situations, to monitor adaptation. Both questionnaires were rated on a 0 (not good) to 10 (perfect) scale.

STATISTICS

Paired t-tests were used to compare HA gains after loudness balancing between both fittings, for each frequency band in each of the 40, 60 and 80 dB SPL input levels. Speech understanding and questionnaire results were analyzed using a Linear Mixed Model (LMM) treating 'Subject' as a random factor. For speech understanding in quiet, we used a fixed factor 'Device' (CI-only, bimodal-broadband, bimodal-three-band, HA-broadband, HA-three-band). Speech understanding in noise was tested per noise configuration with the factor 'Fitting' (broadband, three-band). Analysis per individual subject was done with a 'Noise Configuration' and 'Fitting' factor, testing the two SNR50's measured per speaker configuration. For the questionnaires, mean results per subscale were tested with the factor 'Fitting'. For the questionnaire monitoring adaptation, an extra factor 'Week' (four weeks) was included.

RESULTS

HA GAIN

The mean difference in gain between fittings, after loudness balancing, is visualized in Figure 3 by the black dots. Broadband balancing resulted in more gain than three-band fitting with 40 dB SPL input for the middle ($t(14)=2.87$, $p=0.01$; BB: 62±5 dB; 3B: 59±5 dB) and high ($t(14)=3.89$, $p=0.002$; BB: 36±4 dB; 3B: 20±14 dB) frequency band, but not for the low band. A similar result was found with 60 dB SPL input (middle band: $t(14)=2.24$, $p=0.04$; BB: 54±6 dB; 3B: 51±5 dB, high band: $t(14)=2.63$, $p=0.02$; BB:

25±4 dB; 3B: 15±13 dB). For the 80 dB SPL input level there were no differences in gain between fittings. Little gain was expected in the high frequency band of the three-band fitting because we eliminated gains at frequencies >120 dB HL. This effect is excluded in Figure 3 by the open symbols that represent gain differences only for frequencies where hearing loss was better than 120 dB HL. The 120 dB HL cut-off frequency was on average 3.1 ± 1.7 kHz and 1 kHz or higher for all subjects.

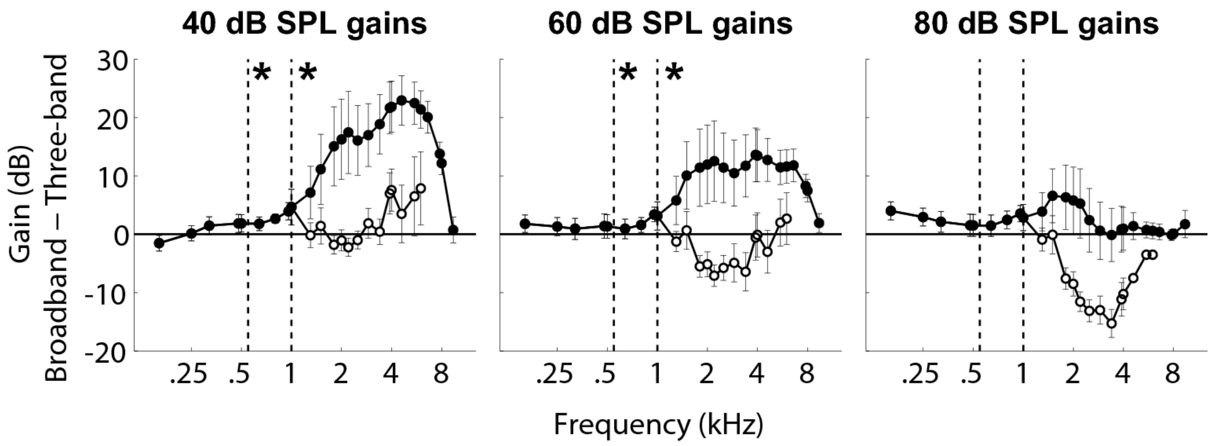

Figure 3. *Mean (± SE) difference in the resulting gains of the broadband fitting minus the three-band fitting at 40, 60 and 80 dB SPL input levels. Open symbols represent differences only for gains where hearing loss is better than 120 dB HL. Dashed lines represent boundaries of the frequency bands used in the three-band fitting. The star indicates a significant difference (p<0.05) in mean gains between fittings in that frequency band.*

SPEECH UNDERSTANDING IN QUIET

Percentages phonemes correct, transformed to rationalized arcsine units, differed significantly between listening conditions ($F(4,56)=122.44$, $p<0.001$) (Figure 4). The two HA-only scores (broadband: 19 ± 19 RAU; three-band: 20 ± 18 RAU) were worse than CI-only (84 ± 10 RAU, both $p<0.001$) and both bimodal configurations (broadband: 97 ± 12 RAU; three-band: 84 ± 11 RAU, both $p<0.001$).

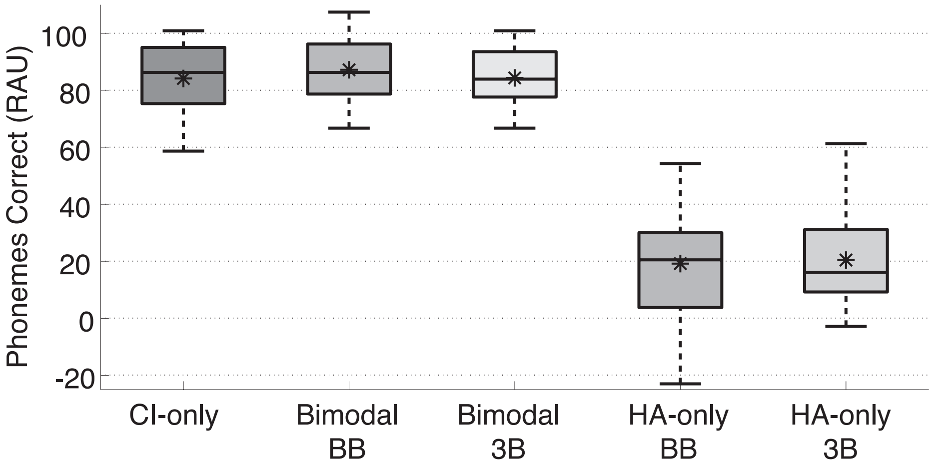

Figure 4. Box-whisker plots of the monosyllable understanding scores in quiet (percent phonemes correct transformed to RAU) for CI-only, bimodal and HA-only conditions for both the three-band (3B) and the broadband fitting. Stars indicate mean values.

SPEECH UNDERSTANDING IN NOISE

Figure 5A shows the SNR50 of the CI-only and bimodal listening conditions and Figure 5B shows the difference between these values, called bimodal benefit. For none of the speaker configurations, a significant amount of benefit was found (S0N±90 stationary noise: $F(1,14)=0.01$, $p=0.91$; S0NHA: $F(1,14)=0.17$, $p=0.69$; S0NCI: $F(1,14)=0.79$, $p=0.39$; S0N±90 single-talker noise: $F(1,14)=0.05$, $p=0.84$). We also did not find a significant difference between broadband and three-band fitting (S0N±90 stationary noise: $F(1,14)=0.36$, $p=0.56$; S0NHA: $F(1,14)=0.06$, $p=0.80$; S0NCI: $F(1,14)=0.07$, $p=0.79$; S0N±90 single-talker noise: $F(1,14)=0.02$, $p=0.89$). Thus, overall, the HA did not bring an advantage nor disadvantage for speech understanding in noise, regardless its fitting. The bimodal benefit, averaged over fittings, was 0.1±4 dB for S0N±90 in stationary noise, -0.2±5 dB for S0N±90 in single-talker noise, 0.9±4 dB for S0NCI and 0.3±4 dB for S0NHA.

Figure 6 shows the individual bimodal benefit for the two fittings, averaged over the two SNR50's measured per condition. Significantly more benefit for one of the fittings was found for nine out of fifteen subjects. Five subjects did 3.1 ± 1.0 dB better on average with the broadband fitting (P1, P4, P7, P11, P15) and four subjects showed 3.3 ± 1.2 dB more benefit with the three-band fitting (P2, P6, P9, P12). Of note, after Bonferroni correction for multiple comparisons, reducing the level of significance to $p = 0.05/15$, the results of only one subject remained significant (P6, $p=0.003$). The trends towards individual differences seemed not related to the subject's binaural experience; most subjects reported to perceive only minor differences between both fittings. However, we noted that the subjects that performed better with the broadband fitting had more bimodal experience than the four subjects that performed better with the three-band fitting (see Discussion).

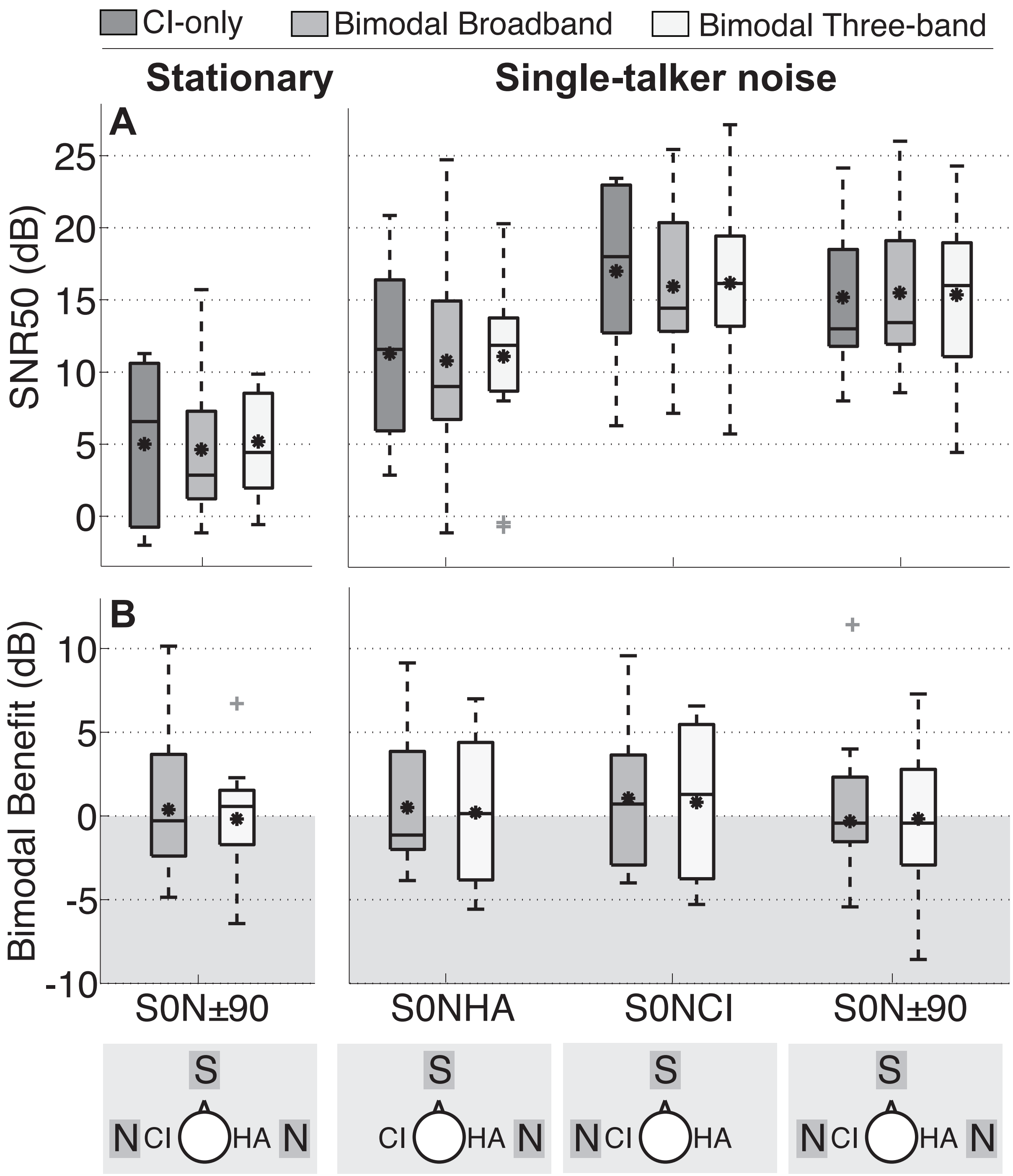

Figure 5. *Box-whisker plots of the speech reception threshold **(A)** and bimodal benefit **(B)** in noise in four speaker configurations. **A)** SNR50 represents the signal to noise ratio (SNR) at 50% correct. Lower values indicate better performance. Results of the CI-only and the bimodal conditions with the broadband and three-band fitting are shown. **B)** Bimodal benefit for the broadband and three-band fitting. Single-talker noise was the IFFM' nonsense natural speech material. Bars represent 25-75% quartiles. Stars indicate mean values. Outliers beyond the 1.5 interquartile range are represented by gray plus symbols.*

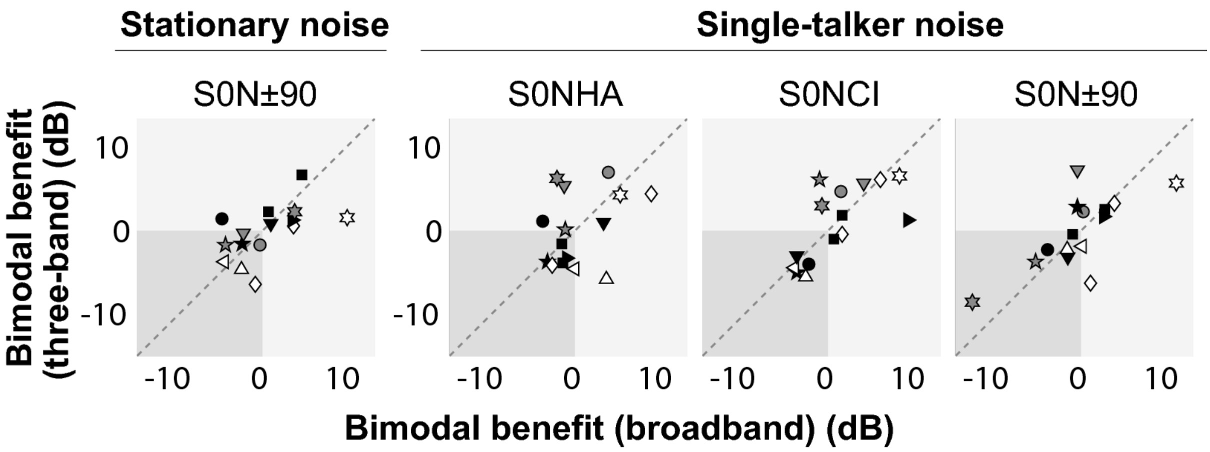

***Figure 6**. Individual scores of bimodal benefit for speech understanding in noise for four speaker configurations comparing the broadband and three-band fitting procedures. Data points in the dark area indicate that a negative effect (bimodal interference) was obtained with both HA fittings. Subjects are indicated by the same marker in all subplots. White markers identify subjects that benefitted significantly more from broadband balancing, and grey markers identify subjects that benefitted significantly more from three-band balancing.*

SPATIAL RELEASE FROM MASKING

Spatial release from masking (SRM) represents the benefit from spatially separating the target and noise signals as compared to the S0N0 configuration. Only for S0NHA we found SRM (3.0 dB for both fittings) ($F(1,14)=12.43$, $p=0.001$). For S0NCI, a significant decrement was found: -2.1 dB ($F(1,14)=15.21$, $p=0.002$). We obtained no significant difference between fittings (S0NHA: $F(1,14)=0.06$, $p=0.80$; S0NCI: $F(1,14)=0.07$, $p=0.79$).

QUESTIONNAIRES

We administered 30 questionnaires (two per subject), two of which were not returned. There were no significant differences between the broadband and three-band fitting in subsections of the SSQ; speech in quiet was given the highest ranking (7.0 for three-band fitting) but low rankings were given to speech in noise (4.3, broadband fitting) and other situations requiring the separation of multiple sources. Mean ratings, averaged over the two fittings, were 4.9±1.7, 4.3±1.6 and 5.6±1.4 for respectively the Speech, Spatial and Quality domain. The questionnaire that was filled in every week to monitor the take-home period did not show any effects over time, implying no adaptation; three-band fitting was ranked significantly better for understanding one person in quiet ($F(1,93.54)=7.00$, $p=0.01$) and the timbre of the sound ($F(1,89.10)=7.74$, $p=0.007$); no difference was found for the other five questions (control of own voice, sound localization, perceiving speech intonation, understanding a group in quiet and understanding one person at a party).

FACTORS CORRELATING WITH BIMODAL BENEFIT

No significant correlations were found between subject characteristics (the low (250, 500, 750 and 1000 Hz) and high (1, 2, 4 kHz) frequency pure tone average in the HA ear, age, bimodal experience, CI experience, electrical dynamic range, HA-only phonemes correct scores) versus the amount of bimodal benefit averaged over fittings or versus the difference in benefit of both fittings.

DISCUSSION

In this study, a loudness balancing procedure was tested for the combined use of a CI and HA, which was intended to balance loudness between devices across the dynamic range and across frequencies. This three-band fitting procedure was compared to a simple balancing procedure using a broadband signal. The two procedures differed in stimulus type (speech versus stationary noise), frequency band (broadband versus three bands), input level (65 dB SPL versus both 40 and 80 dB SPL) and stimulus presentation (free field versus DAI). By these means, we aimed to compare the most simple balancing procedure with a more detailed procedure. No significant differences in bimodal benefit were found between the two fittings, as assessed with measures of speech understanding after a 4-week acclimatization period. However, we found a trend towards more benefit with either one of the fittings for several subjects, , suggesting the importance of testing both procedures in each subject.

Interestingly, the four subjects that seemed to perform better with three-band fitting had better hearing at 1, 2 and 4 kHz than subjects that seemed to perform better with broadband fitting (pure tone averages of 102 and 115 dB HL respectively). This finding was supported by a trend towards more benefit from three-band as opposed to broadband fitting with better high frequency thresholds (S0NCI: $r=-0.55$, $p=0.032$) (Figure 7). Possibly, an accurate loudness balance can contribute to improved speech understanding in the case of usable high-frequency residual hearing. Another reason for better performance with the three-band fitting is that we decreased gains for frequencies with > 120 dB HL, possibly corresponding to dead regions that can sometimes hamper speech understanding when amplified (Moore, 2004). The additional value of frequency-dependent fitting may become more obvious when tested in a population with better hearing thresholds.

Another finding regarding individual differences is that all five subjects who seemed to perform better with the broadband fitting did have experience with wearing a HA beside their CI and four of these five were still bimodal users at the start of this study. Of the other group, better with three-band, only one subject was an experienced bimodal user. Possibly, bimodal performance was biased towards the broadband fitting, since HA fittings of the experienced bimodal users corresponded more with the broadband than with the three-band fitting. The four-week adaptation period we applied in our cross-over design was possibly not long enough for full acclimatization to the new bimodal listening situation, with a fully fused sound percept and pitch match with the CI (Reiss et al., 2014a).

The absence of large differences in bimodal benefit between both fitting procedures is possibly explained by the similarity in HA gains after loudness balancing (Figure 3), especially for the lower frequencies. This implies that the standard Phonak fitting formula we used for broadband fitting resulted in a frequency response with a low frequency slope that agreed well with loudness judgments of our subjects. We did find a significant difference in HA gain between fittings for higher frequencies, but these probably played a minor role in speech understanding, given the severe hearing loss of our subjects.

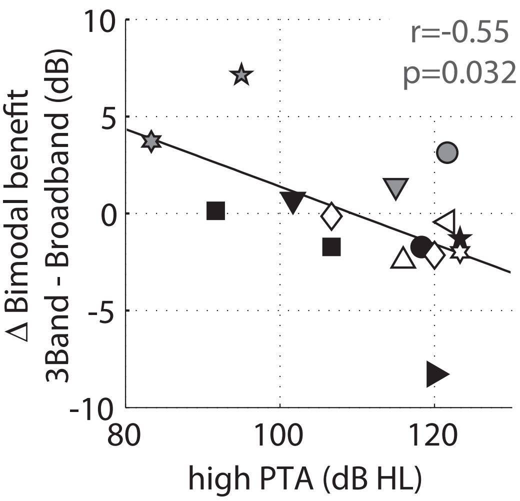

Figure 7. *The correlation between the high-frequency pure tone average (1,2,4 kHz) in the non-implanted ear versus the additional benefit from three-band balancing (three-band minus broadband) in the S0NCI noise configuration. White markers identify subjects that seemed to benefit significantly more from broadband balancing, and grey markers identify subjects that seemed to benefit significantly more from three-band balancing.*

BIMODAL BENEFIT

Bimodal benefit was on average 0.4 dB with noise from the HA side (S0NHA) and 0.9 dB for S0NCI. For adult bimodal listeners, values of respectively 1 dB and 1 to 3 dB have been reported in the literature (Ching et al., 2007; Dunn et al., 2005; Tyler et al., 2002) and our results are therefore comparatively small. However, so far we have not mentioned that we performed a CI-only measurement before we provided subjects with their new HA, besides the CI-only measurement at the end of the study. Thresholds from the first measurement were on average 2 dB poorer than the CI-only measurement at the end of the study; bimodal benefits referenced against this first CI-only measurement was therefore on average 2.2±4 dB. Assuming that improved performance was caused by more experience with the task, we based our analysis on the last CI-only measurement to reduce overestimation of the absolute level of bimodal benefit (but at the risk of underestimation, e.g. in the case of a content-learning effect of the sentences that had to be re-used for the final CI measurement).

Single-talker non-sense speech was a more effective masker than stationary speech-weighted noise. . The absence of a positive effect of masker modulations in speech understanding in noise in CI users has recently been interpreted as the result of spectral smearing caused by current spread, leveling out masker modulations (Oxenham et al., 2014). Considerable spatial release from masking was only present when noise was shifted from the front to the HA side and a deteriorating effect occurred by moving the noise source to the CI side, consistent with findings from others (Litovsky et al., 2006). This was not an unexpected finding since the head shadow protects the CI ear from high frequency noise components in the S0NHA configuration, while low frequencies remain more or less audible at the HA ear. For speech in quiet we did not find a change in performance by adding a HA, in line with several other studies testing bimodal subjects with comparable poor residual hearing (Dunn et al., 2005; Tyler et al., 2002).

Subjects reported to perceive the pitch of the signal used for loudness balancing differently in each ear, especially for the highest frequency band in three-band fitting. The maximum possible amplification of the HA was insufficient for establishing a loudness balance in this frequency band (> 1 kHz) for five subjects.(no link could be found with the bimodal performance of these subjects). For future studies, we would therefore recommend to reduce the bandwidth to frequencies with hearing loss below 120 dB HL, ensuring better audibility. The procedure can furthermore be improved by an additional compression knee point in between the soft and loud level, to allow for free gain adjustments.

CHAPTER 3

MATCHING AUTOMATIC GAIN CONTROL ACROSS DEVICES IN BIMODAL COCHLEAR IMPLANT USERS

Published as *Veugen, L.C., Chalupper, J., Snik, A.F., van Opstal, A.J. and Mens, L.H., 2016. Matching automatic gain control across devices in bimodal cochlear implant users. Ear and hearing, 37(3), pp.260-270*

ABSTRACT

OBJECTIVES. The purpose of this study was to improve bimodal benefit in listeners using a cochlear implant (CI) and a hearing aid (HA) in contralateral ears, by matching the time constants and the number of compression channels of the automatic gain control (AGC) of the HA to the CI. Equivalent AGC was hypothesized to support a balanced loudness for dynamically changing signals like speech and improve bimodal benefit for speech understanding in quiet and with noise presented from the side(s) at 90°.

DESIGN. Fifteen subjects participated in the study, all using the same Advanced Bionics Harmony CI processor and HA (Phonak Naida S IX UP). In a three-visit cross-over design with four weeks between sessions, performance was measured using a HA with a standard AGC (syllabic multichannel compression with 1 ms attack time and 50 ms release time) or an AGC that was adjusted to match that of the CI processor (dual AGC broadband compression, 3 and 240 ms attack time, 80 and 1500 ms release time). In all devices, the AGC was activated above the threshold of 63 dB SPL. We balanced loudness across the devices for soft and loud input sounds in three frequency bands (0 - 548, 548 - 1000 and >1000 Hz). Speech understanding was tested in free field in quiet and in noise for three spatial speaker configurations, with target speech always presented from the front. Single-talker noise was either presented from the CI side or the HA side, or uncorrelated stationary speech-weighted noise or single-talker noise was presented from both sides. Questionnaires were administered to assess differences in perception between the two bimodal fittings.

RESULTS. Significant bimodal benefit over the CI alone was only found for the AGC-matched HA for the speech tests with single-talker noise. Compared to the standard HA, matched AGC characteristics significantly improved speech understanding in single-talker noise by 1.9 dB when noise was presented from the HA side. AGC matching increased bimodal benefit insignificantly by 0.6 dB when noise was presented from the CI implanted side, or by 0.8 (single-talker noise) and 1.1 dB (stationary noise) in the more complex configurations with two simultaneous maskers from both sides. In questionnaires, subjects rated the AGC matched HA higher than the standard HA for understanding of one person in quiet and in noise, and for the quality of sounds. Listening to a slightly raised voice, subjects indicated increased listening comfort with matched AGCs. At the end of the study, nine out of fifteen subjects preferred to take home the AGC-matched HA, one preferred the standard HA and five subjects had no preference.

CONCLUSION. For bimodal listening, the AGC-matched HA outperformed the standard HA in speech understanding in noise tasks using a single competing talker and it was favored in questionnaires and in a subjective preference test. When noise was presented from the HA side, AGC matching resulted in a 1.9 dB SNR additional benefit, even though the HA was at the least favorable SNR side in this speaker configuration. Our results possibly suggest better binaural processing for matched AGCs.

INTRODUCTION

Research on the combined use of a cochlear implant (CI) and a contralateral hearing aid (HA) is an area of growing interest and importance. Since inclusion criteria for CI selection are expanding (Gifford, 2011), more and more people with considerable residual hearing are implanted who can potentially benefit from 'bimodal' stimulation. According to a recent survey in our clinical center, about one third of the unilateral CI recipients uses a conventional HA in the opposite ear. This bilateral combination of acoustic and electric information delivered by the CI and the HA has been shown to result in a wide range of benefits over a unilateral CI, at least in some patients, including improved speech understanding in noise, music and voice pitch perception, and sound-source localization (Armstrong et al., 1997; Ching et al., 2007; Dorman et al., 2008; Firszt et al., 2008; Kong et al., 2005; Shpak et al., 2014; Straatman et al., 2010; Tyler et al., 2002; Zhang et al., 2010).

The HA typically provides access to some of the lower frequencies that are not available through the CI (Francart et al., 2013). This complementary information may support voice pitch perception and the perceived naturalness of sounds, among other benefits (Ching et al., 2007). In addition, bimodal stimulation possibly results in improved spatial hearing through binaural processing, provided that interaural level differences (ILDs) are available (Francart et al., 2013). These bimodal and binaural cues may improve horizontal sound localization and help segregating the target speech signal from a mixture of sounds or competing speakers, which is challenging for hearing impaired people, and CI users in particular (Ching et al., 2007; van Hoesel, 2012).

In current clinical practice, bimodal devices are often fit separately, lacking to ensure optimal perception of binaural cues. ILDs would be better preserved with a HA frequency response that creates equal loudness with the CI, which was the topic of several studies on bimodal fitting (Blamey et al., 2000; Ching et al., 2007; Francart et al., 2012a; Potts et al., 2009). However, even with a perfect loudness balance across the dynamic range for stationary sounds, ILD cues may still be disrupted for inputs with dynamically changing intensities, like speech sounds, because of differences in signal processing. The automatic gain control (AGC) in both devices can react differently on changes in sound level, generating unstable binaural cues (Moore, 2007).

AGC circuits in hearing devices aim to map incoming sounds into the reduced dynamic range of the hearing impaired ear by controlling the gain as a function of signal level to optimize audibility of low-level sounds and avoid discomfort of high-level sounds (Dillon, 2001). In most systems, linear amplification is applied up to a certain input level, called the compression threshold or knee point. Above this input level, the signal is compressed. The speed with which the AGC reacts to sudden increases and decreases of the input sound level is determined by the attack and release times, respectively (Moore, 2008). Typical settings are fast "syllabic" compression (< 10 ms attack and 10-50 ms release time) to reduce intensity differences between speech sounds, or slow compression (> 100 ms attack and > 400 ms release time) to adapt to the overall level of speech and other sounds (Dillon, 2001). Very slow compression systems (attack and release > 1 s) are often referred to as 'automatic volume control'. Both approaches can be combined, for instance in a dual front-end AGC system, incorporating a slow AGC loop to control overall signal level and a fast AGC loop to reduce sudden changes in sound level (Moore, 2008). There may be one AGC reacting to all frequencies in the input signal to control the overall gain, or multiple AGCs that work independently in different frequency bands. In effect, a wide variety of AGC circuits are available in modern hearing devices (Dillon, 2001).

Currently, there are no commercially available CI processors and HAs specifically designed for combined use, so AGCs differ in design and therefore in operation. The amount of mismatch in AGC characteristics most likely differs widely across combinations of devices, possibly explaining part of the inter-individual differences in bimodal benefit often observed, and also of the success rate and acceptance of the HA by CI users (Fitzpatrick et al., 2010; Mok et al., 2006). To our knowledge, the effect of dissimilar AGCs compared to identical AGCs in bilateral devices has never been studied.

In this study, we aimed to increase bimodal benefit by using a similar design of the AGC in the CI and HA. Bimodal listeners were included who used the Advanced Bionics Harmony processor in one ear and the Phonak Naida IX UP HA in the other ear. Because the CI is the main source of auditory information in most bimodal patients, no alterations were done to the AGC circuit of the CI. To determine the effect of the compression characteristics of the HA used in conjunction with a CI, bimodal performance was compared with the Naida HA with standard multichannel fast-acting compression and with its compression matched to the dual-loop AGC in the Harmony processor (Boyle et al., 2009). The effect of matching AGC characteristics on bimodal benefit was examined by testing speech understanding in quiet and in noise with target speech presented from the front and noise presented from the sides, as well as through subjective judgments. Although all subjects in this study had at least two months of bimodal experience, they differed in aspects such as the amount of residual hearing and duration of deafness, which were considered factors associated with bimodal performance.

METHODS

SUBJECTS

Fifteen postlingually deaf subjects (10 male, 5 female; mean age 61 ± 12 years) were recruited that all used the same type of CI speech processor (Harmony, Advanced Bionics, Sylmar, CA) in one ear and an acoustic HA (Naida S IX UP, Phonak, Stäfa, Switzerland) in the other ear for the two months preceding this study. Six subjects were already bimodal users before that time, but then used other types of HAs. Subjects were selected with thresholds in the non-implanted ear better than 110 dB HL at 500 Hz. Audiometric thresholds and demographic details of the subjects are reported in Table 1 and 2, respectively. The study was approved by the Local Ethics Committee of the Radboud University Nijmegen (40327.091.12).

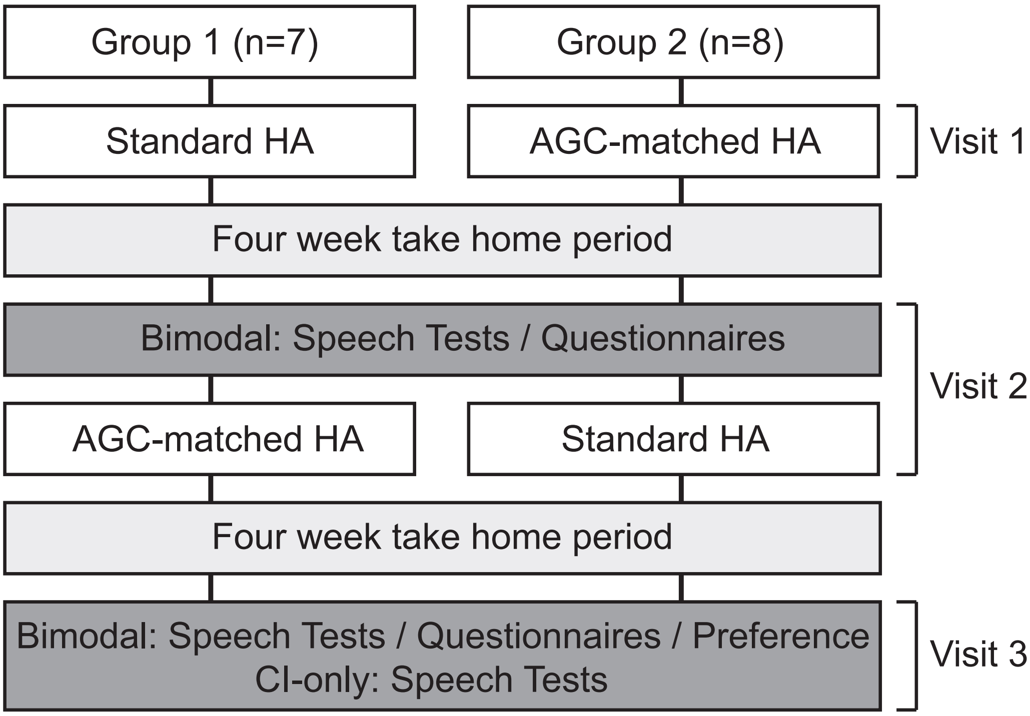

Figure 1. *Overview of the cross-over design of the study.*

STUDY DESIGN AND CONDITIONS

The study consisted of a three-visit cross-over design with four weeks between sessions (Figure 1). We compared the AGC as programmed in a commercially available HA (Naida S IX UP; 'standard HA'), with an experimentally programmed version of the same aid, featuring an AGC that was adjusted to match the Harmony CI's AGC (same device, 'AGC-matched HA'). In this AGC-matched HA, AGC time constants and compression channels were similar to the CI processor. For the cross-over design, subjects were divided in two groups, without significant differences in hearing thresholds (Figure 2), age (62 ± 11 and 61 ± 13 years), gender (4 and 6 male), post-implant duration of CI use (5.1 ± 2.2 and 5.5 ± 2.9 years) and side of implant (6 and 4 were subjects implanted on the left side). Visits were scheduled on average 31 days apart, with a minimum of 22 days, to allow for HA acclimatization during a take home period. Subjects were not told which type of AGC was programmed in the HA. For all subjects, we fitted HAs according to the same procedure that aimed to create a loudness balance with the CI across loudness levels and across frequency bands. After the acclimatization period, bimodal benefit was evaluated by assessments of speech understanding in quiet and in noise and by questionnaires. Speech reception thresholds at the signal-to-noise ratio for 50% performance (SNR-50) were determined for several speaker configurations with target sentences always presented from the front and noise presented from the side(s). We compared speech perception performance in the bimodal conditions with the CI-only listening situation.

Table 1. *Hearing thresholds (dB HL) for all subjects at the non-implanted ear as assessed with standard pure tone audiometry.*

Subject	125 Hz	250 Hz	500 Hz	750 Hz	1 kHz	1.5 kHz	2 kHz	3 kHz	4 kHz	6 kHz	8 kHz
1	55	65	85	95	120	120	120	120	120	120	120
2	60	75	85	100	115	↓	↓	↓	↓	↓	↓
3	45	60	100	115	120	↓	↓	↓	↓	↓	↓
4	60	70	80	90	95	95	100	120	120	120	120
5	80	80	90	85	80	80	80	90	115	↓	↓
6	60	85	110	105	105	110	115	115	120	120	120
7	65	95	105	115	120	120	120	120	120	120	120
8	30	60	105	110	110	120	120	120	120	120	120
9	65	85	85	90	90	85	80	80	80	115	↓
10	50	75	100	105	105	115	120	120	120	120	120
11	35	55	95	105	110	↓	↓	↓	↓	↓	↓
12	85	95	85	90	90	90	90	90	105	120	120
13	65	75	90	95	95	95	100	115	↓	↓	↓
14	65	80	95	95	95	95	95	105	115	120	120
15	55	75	80	90	100	↓	↓	↓	↓	↓	↓

↓ indicates no response at audiometer limit of 120 dB HL.

Table 2. *Demographic details of the participating subjects. The 'group' column indicates the distribution of the subjects over the cross-over design with group 1 starting with the standard HA and group 2 starting with the AGC-matched HA. The asterisk marks experienced bimodal users, indicating that HA use overlaps with the full amount of CI experience.*

Participant	Age (yr)	Gender	Implanted ear	Etiology	Bimodal user	CI experience (yr)	HA experience non-implanted ear (yr)	Group
1	63	M	L	Unknown	Y	6.3	9*	2
2	46	M	L	Genetic	Y	3.7	16*	2
3	79	F	L	Genetic	N	8.9	26	1
4	58	M	R	Ototoxicity	Y	5.3	50*	2
5	74	M	R	Genetic	Y	1.3	28*	2
6	52	F	R	Genetic	N	5.6	27	2
7	44	M	R	Prenatal	Y	3.8	32*	1
8	75	M	L	Unknown	N	1.1	18	2
9	64	M	R	Unknown	N	2.8	20	1
10	78	F	R	Genetic	N	6.8	10	2
11	64	F	R	Genetic	Y	5.6	30*	1
12	58	M	R	Genetic	N	4.8	32	1
13	42	M	L	Perinatal	N	9.9	28	2
14	70	M	R	Meniere's	N	2.3	13	1
15	53	F	R	Genetic	N	4.3	9	1

Note. CI=cochlear implant; HA=hearing aid

3

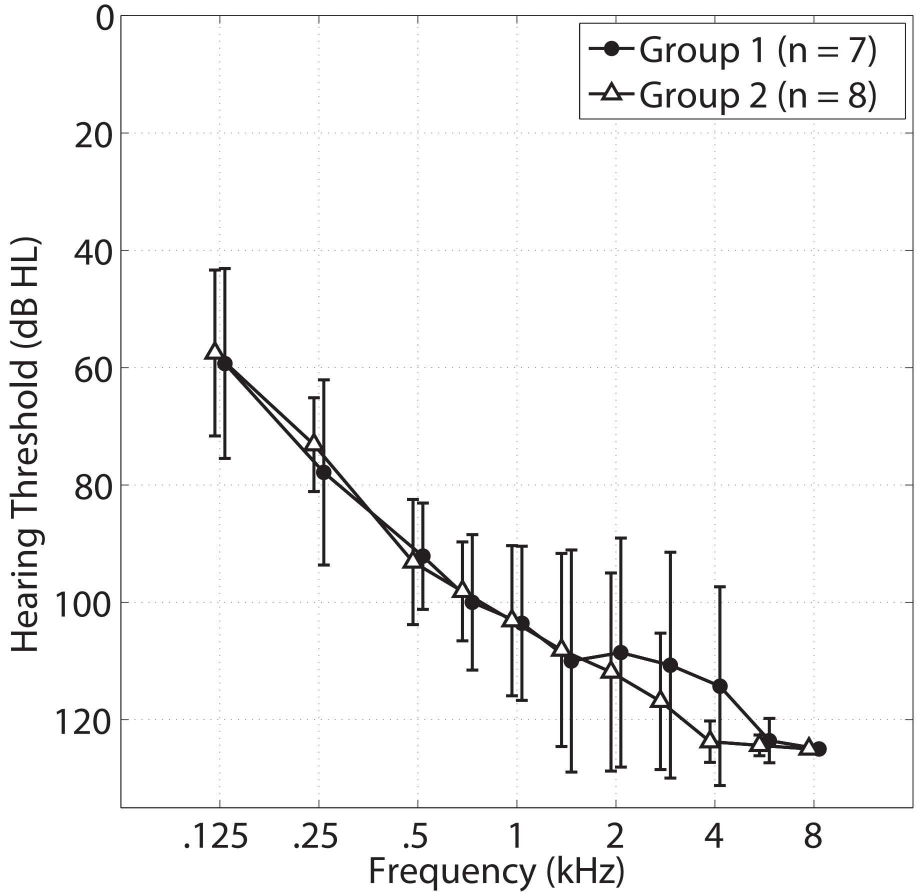

Figure 2. *Pure tone hearing thresholds (mean ± 1 SD) in the non-implanted ear for the two subgroups of the cross-over design. Group 1 started with the standard HA and group 2 started with the AGC-matched HA. Thresholds beyond the audiometer limit (120 dB HL) were assigned a value of 125 dB HL.*

AGC CHARACTERISTICS

The Harmony CI processor has a single-channel dual-loop AGC system incorporating both slow and fast attack and release time constant circuits, and a compression ratio of 12:1 (Boyle et al., 2009). Operation is normally controlled by the slow AGC loop; the fast AGC loop rapidly reduces gain in case of sudden increases. This compression system was implemented as close as possible for speech signals in the AGC-matched HA as follows: 1) Slow (240 and 1500 ms) and fast (3 and 80 ms) time constants were programmed into the HA. 2) Compression channels in the HA were coupled to mimic the single channel broadband compression as present in the CI processor. For comparison, the standard HA uses multichannel compression that operates independently in twenty different frequency bands with a 1 ms attack time and 50 ms release time. Given the differences in loudness growth between the CI and HA, the compression ratio in both HAs was set during the loudness balancing procedure, usually resulting in strong compression above 63 dB SPL

(≥ 10:1 in twelve subjects, 5:1 in two subjects and 2.5:1 in one subject), approaching the 12:1 ratio of the CI processor. Below 63 dB SPL input, amplification approximated linear behavior with a compression ratio of 1.5 ± 0.4 on average across subjects.

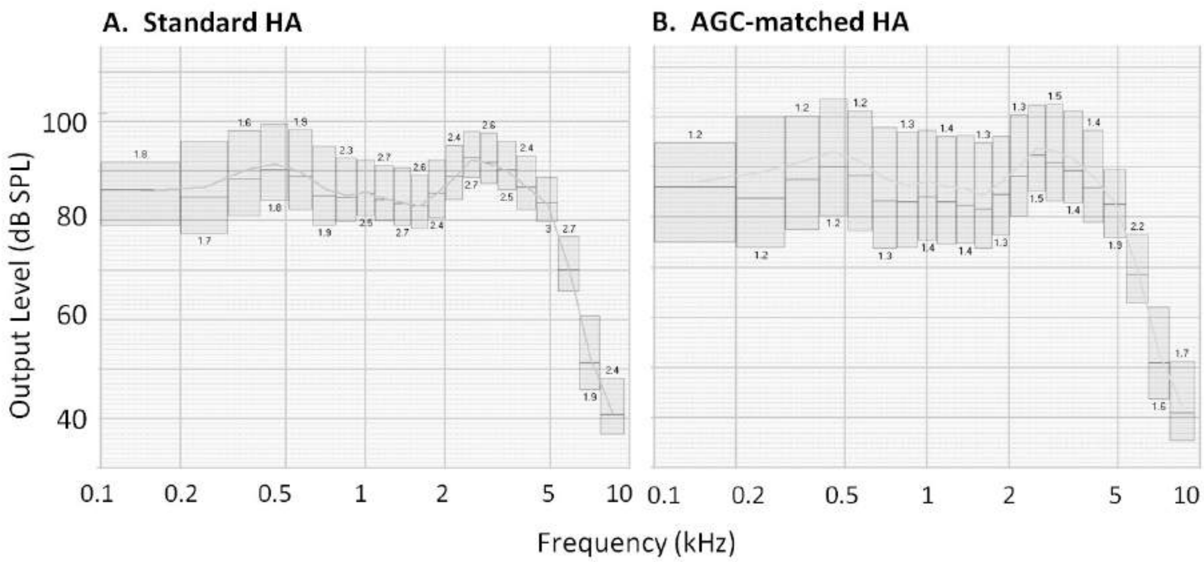

Figure 3. *Output level per critical band of the International Speech Test Signal (ISTS) at 80 dB SPL after processing by the standard* ***(A)*** *and AGC-matched HA* ***(B)****. Boxes represent the percentiles for 30, 65 and 99% of the short-term levels. The line shows the long-term average speech spectrum level. Numbers around bars represent the effective compression ratios from the input to output in the 30-90 % range.*

In all devices (both HAs and the CI), the compression knee point was fixed at 63 dB SPL, considering the long-term average speech spectrum. Above this knee point, the slow AGC loop (in the CI and AGC-matched HA) or syllabic compression (in the standard HA) was activated. The fast compressor of the dual-loop AGC was activated above 71 dB SPL or when the output increased by more than 8 dB on top of the level as determined by the slow AGC loop. Attack and release times were identical below and above 63 dB SPL. Both HAs used a fast-acting compression circuit to limit output levels according to the maximum power output around 134 dB SPL and frequency dependent expansion for input levels below 30 dB SPL to attenuate microphone noise.

In Figure 3 we visualized how speech is processed by both HAs, using standard procedures for percentile analysis (IEC 60118-15). In this figure, the 30th percentile denotes the level that is exceeded by 70% of the short-term (125 ms) speech levels. The 30th, 65th and 99th percentiles thus represent soft, moderate and loud components in speech. The AGC-matched HA resulted in a large output dynamic range compared to the standard HA, as a result of its small effective compression ratio due to longer AGC time constants. Syllabic compression in the standard HA reduces the output dynamic range by making soft sounds louder and loud sounds softer.

DEVICE SETTINGS

All subjects kept using their 'everyday' CI program. Adaptive features and settings (directionality, noise reduction, etc.) were turned off in the CI and HA for the entire duration of the study, except for the HA's adaptive feedback canceller ('WhistleBlock') in case of feedback problems. Disabling the adaptive features ensured that signal processing in both devices was only controlled by the AGC mechanisms that were under investigation. Subjects exclusively used the omnidirectional microphone mode of the CI and the HA.

HEARING AID FITTING

HAs were fitted according to the same procedure for all subjects. Fittings were based on in situ pure-tone audiometry ('AudiogramDirect'), for which sounds were presented through the HA and levels controlled by the fitting software. Our aim was to establish a loudness balance between the CI and the HA across the dynamic range and across frequencies. First, the Advanced Bionics/Phonak bimodal fitting formula was used to prescribe HA gain, eliminating gain if hearing thresholds exceeded 120 dB HL and optimizing audibility by providing more gain at the low frequencies than the conventional fitting rule (Chalupper et al., 2015); the latter was applied to maximize the effective audibility (Ching et al., 2001a).

We adjusted the HA gain to match loudness with the CI for input signals at two intensities (45 dB SPL and 80 dB SPL) and in three frequency bands (250 – 548 Hz, 548 - 1000 Hz and 1000 Hz up to the frequency where hearing loss exceeded 120 dB HL). The CI was left at the subjects' default everyday volume setting. Loudness balancing was performed using a speech-shaped steady state noise (male speaker, track one of the International Collegium of Rehabilitative Audiology (ICRA) noises (Dreschler et al., 2001), which was transformed to the Fourier domain to obtain the three frequency bands. Noise bursts of 1.5 seconds with ramps of 40 ms were presented alternately to the CI and HA via direct audio input, with an inter-stimulus interval of 0.6 seconds. The stimulus continued playing while we adjusted the HA gain per frequency band and input level with the HA fitting software, until the patient confirmed equal loudness in both ears. Balancing input sounds at 80 dB SPL only affected the compression ratio above 63 dB SPL. Below this knee point, the HA was linear after initial fitting with the prescriptive formula, but balancing the soft 45 dB SPL level could result in weak compression, as explained above. The whole fitting procedure was performed at the first visit in approximately 10-15 minutes; the same gain settings were used for the matched and unmatched AGC HAs.

For loudness balancing of the 80 dB SPL presentation level, a continuous high-level tone of 6 kHz was fed into the CI processor to keep the AGC circuit in steady state. This tone was made inaudible by setting the upper stimulation level of the channel corresponding to 6 kHz to 0 μA. The 40 ms ramps were enough to avoid transient overshooting of the syllabic compressor of the HA.

SPEECH TESTS

Subjects were seated in a sound treated room with reverberation times measured at the location of the listener in one-third octave bands according to ISO 354:2003. A time constant (RT 60) of 240 ms was found in the frequency range of 90 to 500 Hz, and 70 ms for 500 to 6000 Hz. Three loudspeakers (JBL Control 1, Harmon International Industries, Washington DC, USA) were positioned at a distance of one meter from the patient's head at 0°, +90°, and -90° azimuth. Sound stimuli used in this study were delivered through an external soundcard (RME Babyface, Audio AG, Haimhausen, Germany) and a main amplifier (MPA 4-80, Ecler, Spain). Free field presentation levels were calibrated using a sound level meter (Brüel & Kjaer 2260 investigator) at the position of the subject's head. Target words and sentences were always played via the loudspeaker directly in front of the patient at 0°.

The NVA (Nederlandse Vereniging voor Audiologie) Dutch monosyllabic word test was used to assess speech understanding in quiet (Bosman, 1992). In every session, three lists of 12 words were presented at 65 dB SPL in each CI-only, HA-only, and bimodal condition. For statistical analysis, the percentage phonemes that was correctly repeated was transformed to rationalized arcsine units (RAU) to stabilize the error variance in the presence of floor and ceiling effects associated with scores reaching 0% or 100% (Studebaker, 1985).

For speech understanding in noise, we used the Leuven Intelligibility Sentence Test (LIST; (van Wieringen et al., 2008), consisting of 35 lists with 10 sentences. Target sentences were presented at 65 dB SPL and the noise was varied in steps of 2 dB to obtain the signal-to-noise ratio (SNR) for 50% performance (SNR50), which was calculated as the mean presentation level (in SNR) of the last five sentences and one level beyond. The first sentence of each list was presented at a 0 dB SNR and was repeated while decreasing the noise level until correctly identified by the patient. The SNR50 was always determined twice for each condition in every session and averaged for statistical analysis.

We used the stationary speech-weighted noise that is provided by the LIST sentences (F0 at 210 Hz, determined by Praat software, (Boersma et al., 2001), having a frequency spectrum similar to the average target speech. Furthermore, tests were performed with a single competing talker (single-talker noise), generated from the International Female Fluctuating Masker (IFFM) with F0 adjusted to 127 Hz (EHIMA, Retrieved 17 April, 2013). Noise was either presented from the front (S0N0), from the implanted side (S0NCI), from the HA side (S0NHA) or uncorrelated noise was presented from both sides simultaneously (S0N±90). S0N±90 was tested with both types of noise; S0N0, S0NCI and S0NHA were only tested with single-talker noise. In the S0N±90 conditions, noise levels were reduced by three dB to achieve the same overall RMS level as for the unilateral signals.

During the two sessions after each HA acclimatization period, speech understanding in noise was evaluated for all speaker configurations in the bimodal condition. CI-only performance was only tested at the end of this study for all subjects in all speaker configurations except for S0N0, which was only measured in eight subjects. Bimodal benefit was calculated as the difference in SNR50 of the CI-only minus the bimodal listening condition (Ching et al., 2001b; Dorman et al., 2012), with higher values indicating more benefit. Spatial release from masking (SRM) was calculated by subtracting the threshold in a spatially separated noise configuration (S0NHA and S0NCI) from that in S0N0.

QUESTIONNAIRES

Every week during the take home period, subjects filled in seven basic questions concerning everyday listening situations, to monitor acclimatization to the new settings. After acclimatization, subjects were asked to fill in the Speech, Spatial and Qualities of Hearing Scale (SSQ) to assess subjective experience of bimodal hearing (Gatehouse et al., 2004). For both questionnaires, ratings were given on a zero (not good) to ten (perfect) scale. Questionnaires were used to screen for differences between bimodal fittings in the perceived handicap on several binaural hearing functions. All questionnaires were returned at the time of speech testing. One subject forgot to fill in the questionnaires for the AGC-matched HA.

QUICK SUBJECTIVE PREFERENCE

At the end of the study, subjects were presented four HA fittings and we asked for their preference in a quick listening test. These fittings included the standard and AGC-matched HA from this study, as well as two older HA fittings that these subjects once tried in a previous study (AGCs not matched, (Veugen et al., 2016a)) intended as a diversion. Subjects were asked to rate "sound clarity" and "sound quality" on a 5-point Likert scale (ranging from 'very bad' to 'very good'), while listening to everyday sentences (Versfeld et al., 2000). Ratings were obtained for two loudness levels: four sentences were presented at a normal conversational level of 65 dB SPL and another four sentences were presented at 75 dB SPL, ensuring maximum effect of the AGCs. Fittings were presented according to a randomly chosen 4x4 latin square design. We only analyzed the ratings of the standard and AGC-matched HA.

STATISTICS

The data were reported as mean values ± 1 inter-subject standard deviation. Results were analyzed using Linear Mixed Model procedures (LMM) treating Subject as a random factor. For speech understanding in quiet, we used a fixed factor Device (CI-only, bimodal-standard AGC, bimodal-matched AGC, HA with standard AGC, HA with matched AGC). The bimodal benefit was tested per speaker configuration with the fixed factor AGC (standard, matched). Besides the AGC factor, we added a second factor Noise (stationary, single-talker) to test benefit as measured in the two S0N±90 conditions. SSQ scores were analyzed per AGC (standard, matched); for the weekly questionnaire we added the factor Week (1st, 2nd, 3rd, 4th week). Post-hoc analyses were performed using pairwise comparisons with Bonferroni-adjusted unweighted means. A Wilcoxon signed rank test was used to test the effect of AGC on the ratings for clarity and comfort for 65 and 75 dB SPL speech sounds.

We tried to explain individual differences in bimodal benefit by subject characteristics, using Pearson correlation analysis. For each of the four speech-in-noise speaker configurations, the benefit as averaged over the two HAs and the difference in bimodal benefit between HAs was compared with age, the low- and high frequency pure tone average, duration of CI use, electrical dynamic range across electrodes, and duration of bimodal use.

RESULTS

SPEECH UNDERSTANDING IN QUIET

Percentages phonemes correct, transformed to rationalized arcsine units, are shown in Figure 4 for the CI-only, HA-only and bimodal conditions with both the standard and AGC-matched HA. The RAU scale ranges from -23 to 123 RAU, corresponding to 0 and 100% (Studebaker, 1985), explaining the values in Figure 4; within the range from 20 to 80 RAU, values are equivalent to percentages. RAU scores significantly differed between the different listening modes ($F(4,56)=100.02$, $p<0.001$); CI-only and bimodal scores were significantly better than HA-only scores ($p<0.001$; $p=1$ for all other pairwise comparisons).

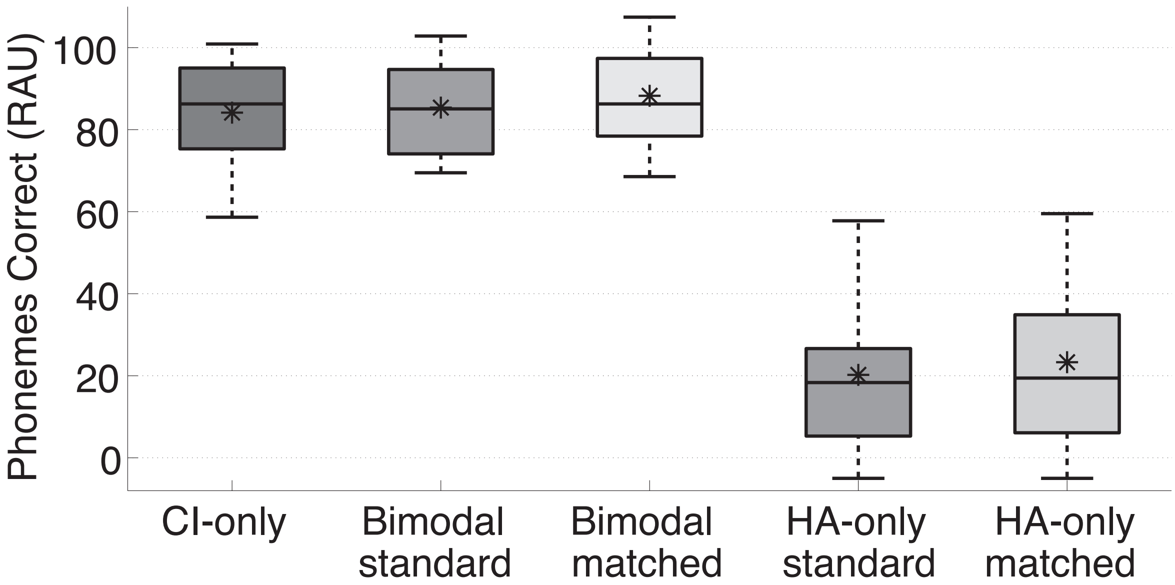

Figure 4. *Box-whisker plots of the monosyllable understanding scores in quiet (percent phonemes correct transformed to RAU) for CI-only, bimodal and HA-only conditions for both the HA with standard AGC ("standard") and the HA with matched AGC ("matched"). Stars indicate mean values.*

SPEECH UNDERSTANDING IN NOISE

We found a significant improvement of 0.69 ± 4.03 dB over the two SNR50 thresholds measured per listening condition ($F(1,322)=4.03$, $p=0.046$). However, as this improvement showed no interaction with speaker configuration or condition (CI-only and both bimodal fittings), values for SNR50 were averaged per condition for statistical analyses. For each speaker configuration, Figure 5A shows the SNR50 of the CI-only and the bimodal conditions and Figure 5B shows the bimodal benefit for the standard and AGC-matched HA, with higher values indicating more benefit. We analyzed bimodal benefit per speaker configuration, as described below.

S0NHA

Tested across types of AGC, no significant overall bimodal benefit was found when single-talker noise was presented from the HA side ($F(1,14)=3.80$, $p=0.071$), but the AGC-matched HA resulted in significantly higher benefit (3.0 ± 4.2 dB) than the standard HA (1.1 ± 4.3 dB, $F(1,14)=6.86$, $p=0.02$). A separate test for each of the bimodal fittings only showed a significant benefit for the AGC-matched HA ($F(1,14)=7.58$, $p=0.016$) and not for the standard HA ($F(1,14)=0.95$, $p=0.346$).

S0NCI

With single-talker noise presented from the CI side, a significant bimodal benefit was found when tested across HA types ($F(1,14)=6.32$, $p=0.025$). We found no difference between HAs ($F(1,14)=0.75$, $p=0.402$). A separate test for each of the HAs showed a significant bimodal benefit for the AGC-matched HA (3.1 ± 3.6 dB, $F(1,14)=10.81$, $p=0.005$) and not for the standard HA (2.5 ± 5.2 dB, $F(1,14)=3.44$, $p=0.085$).

S0N±90

For both single-talker and stationary noise, we obtained no significant bimodal benefit with two uncorrelated noise sources. Furthermore, the benefit did not differ between HAs either. When testing the HAs separately, we found a significant amount of bimodal benefit for the AGC-matched HA in single-talker noise (2.4 ± 3.9 dB, $F(1,14)=5.58$, $p=0.033$), but not for the standard HA (1.5 ± 5.2 dB, $F(1,14)=1.31$, $p=0.27$). No significant bimodal benefit was found in the test condition with stationary speech-weighted noise (standard HA: 0.7 ± 3.1 dB, $F(1,14)=3.58$, $p=0.079$; AGC-matched HA; 1.8 ± 3.7 dB, $F(1,14)=0.82$, $p=0.38$). We did not find a significant difference in bimodal benefit between the single-talker and stationary noise conditions ($F(1,42)=1.68$, $p=0.20$).

S0N0

This configuration was only measured in eight subjects due to time constraints and is therefore not displayed in Figure 5. On average, the benefit was 1.0 ± 4.5 dB for the standard and 0.7 ± 3.2 dB for the AGC-matched HA. These averages did not significantly differ from zero ($F(1,7)=0.45$, $p=0.52$), or from each other ($F(1,7)=0.05$, $p=0.83$).

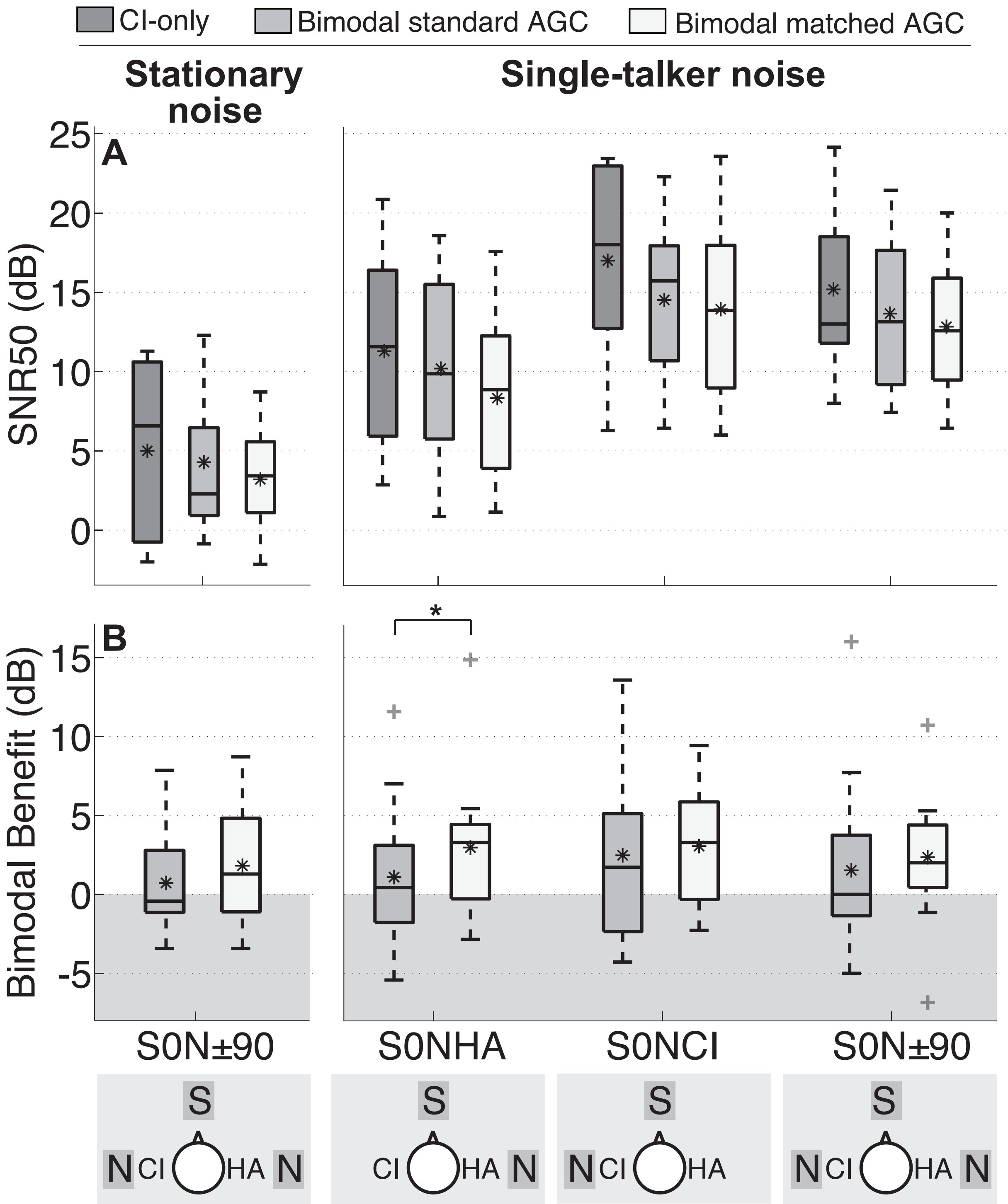

*__Figure 5__. Box-whisker plots of the speech reception threshold __(A)__ and bimodal benefit __(B)__ in noise in four speaker configurations. __A__) SNR50 represents the signal to noise ratio at 50% correct responses. Lower values indicate better performance. Results of the CI-only and the bimodal conditions with the standard and AGC-matched HA are shown. __B)__ Bimodal benefit for the standard and AGC-matched HA. Single-talker noise was the IFFM' nonsense natural speech material. Stars indicate mean values. Outliers beyond 1.5 interquartile range are represented by gray plus symbols. *p < 0.05*

3

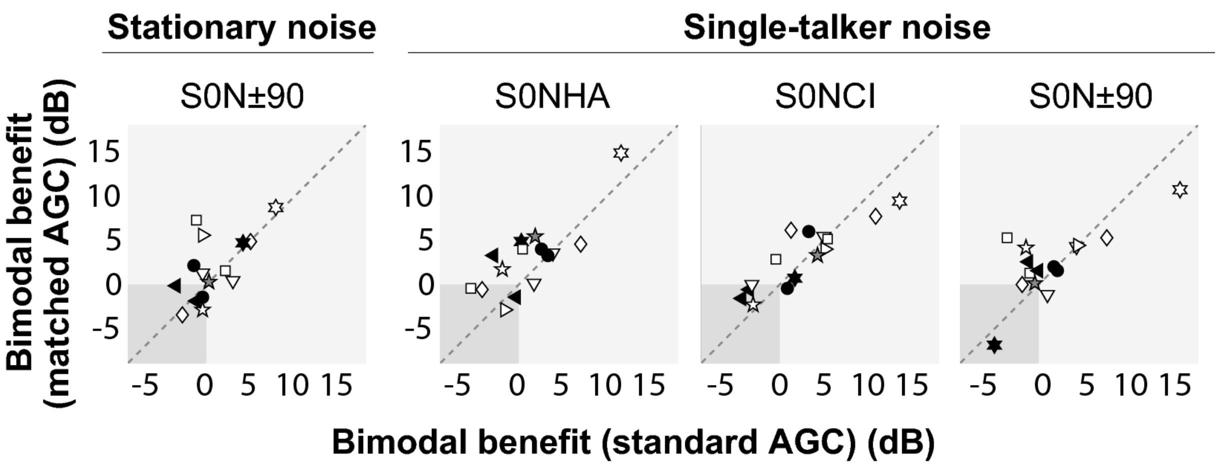

Figure 6. *Individual scores of bimodal benefit for speech understanding in noise in four speaker configurations, comparing the HAs with standard and matched AGC. Marker color indicates HA preference after the study (white for the AGC-matched HA, black for the standard HA, grey means no preference). Data points in the dark area indicate that a negative effect (bimodal interference) was obtained with both HAs. Subjects are indicated by the same marker in all subplots.*

Individual results of bimodal benefit for the two HAs are displayed in Figure 6. Points above the diagonal indicate that the benefit with the AGC-matched HA was larger than with the standard HA. For all speaker configurations, moderate to strong correlations were obtained between the benefits resulting from both HAs; S0N±90 with stationary noise ($r=0.60$, $p=0.017$), S0N±90 with single-talker noise ($r=0.74$, $p=0.001$), S0NCI ($r=0.88$, $p<0.001$) and S0NHA ($r=0.79$, ($p<0.001$). Subjects performed similarly for all speaker configurations.

Variability of performance, as assessed by the inter-subject standard deviation, was biggest for the standard HA in the single-talker S0N±90 and S0NCI configuration (both 5.2 dB); for S0N±90 (stationary noise) the standard deviation was 3.1 dB and for S0NHA it was 4.3 dB. For the AGC-matched HA the standard deviations ranged between 3.6 dB and 4.2 dB, suggesting less variability between subjects.

SPATIAL RELEASE FROM MASKING (SRM)

SRM was calculated as the increase in performance by moving the single-talker noise from a position coincident with the target (S0N0) to a spatially separated one (S0NHA and S0NCI in Figure 7). Overall SRM, tested across both types of HA, was significant only for the S0NHA configuration ($F(1,14)=35.58$, $p<0.001$), but did not differ significantly between HA types. Also when testing the HAs separately for S0NHA, both showed significant SRM: 3.8 ± 3.0 dB for the standard HA ($F(1,14)=23.53$, $p<0.001$) and 4.4 ± 3.5 dB for the AGC-matched HA ($F(1,14)=24.13$, $p<0.001$). For S0NCI, no significant SRM was found (-0.6 ± 2.5, standard HA; -1.2 ± 3.1 dB, AGC-matched HA) and also the difference between HA types was not significant.

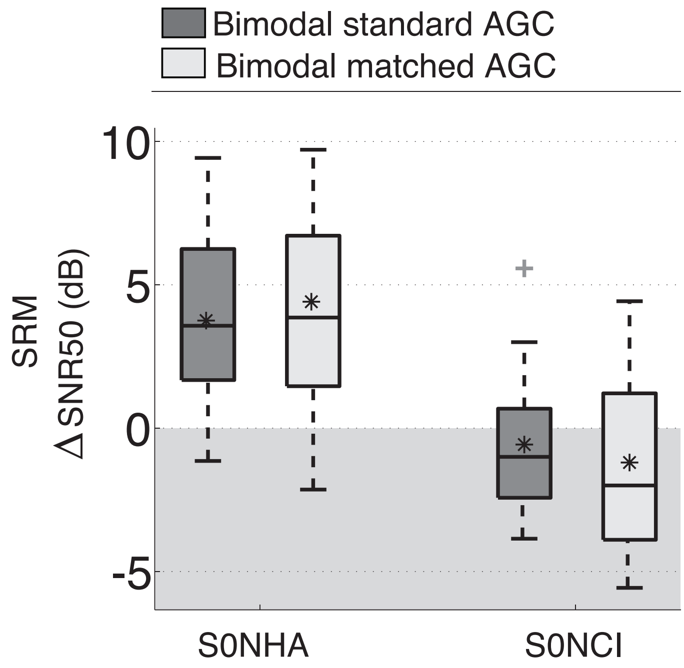

Figure 7. *Box-whisker plots of spatial release from masking (SRM), in each case compared to S0N0. Stars indicate mean values. Positive values are improvements. The dark area means that a negative effect was obtained from spatially separating speech and noise. SRM was measured using single-talker IFFM' noise. Outliers beyond 1.5 interquartile range are represented by gray plus symbols.*

QUESTIONNAIRES

Mean ratings to the subscales of both questionnaires are listed in Table 3 (ratings for the questionnaire that monitored acclimatization were averaged over the four weeks). No significant differences between HA types were found for any of the three general domains of the Speech, Spatial and Qualities of Hearing Scale (SSQ) or it's subsections (Gatehouse et al., 2006). The bimodal listening questionnaire that was filled in every week to monitor the take-home period did not show any effects of time, suggesting no effect of acclimatization. In these questionnaires, the AGC-matched HA was ranked significantly better for understanding one person in quiet ($F(1,94.11)=10.20$, $p=0.002$), understanding one person in noise ($F(1,95.09)=7.92$, $p=0.006$) and the timbre of sounds ($F(1,91.62)=13.99$, $p<0.001$). No differences were found for the other questions (Table 3).

Table 3. *Mean (± SD) ratings given to the Speech, Spatial and Qualities of Hearing Scale and the bimodal listening questionnaire (averaged over the four weeks) for the standard and AGC-matched HA. Questionnaire scales ranged from zero (not good) to ten (perfect). *$p < 0.05$*

	Subscale	**Bimodal fitting**	
		Standard	**AGC-matched**
SSQ	**Speech**	4.7 ± 1.5	4.6 ± 1.4
	Speech in quiet	7.0 ± 1.7	6.9 ± 1.4
	Speech in noise	4.4 ± 1.9	4.4 ± 1.6
	Speech in speech contexts	4.8 ± 1.7	4.8 ± 1.6
	Multiple speech-stream processing and switching	3.2 ± 1.8	3.0 ± 1.8
	Spatial	4.6 ± 1.7	4.5 ± 1.6
	Localization	4.5 ± 2.1	4.4 ± 1.9
	Distance and movement	4.7 ± 1.6	4.7 ± 1.6
	Quality	5.6 ± 1.5	5.7 ± 1.4
	Sound quality and naturalness	5.9 ± 1.8	6.4 ± 1.7
	Identification of sounds and objects	5.7 ± 1.5	5.5 ± 1.7
	Segregation of sounds	6.3 ± 2.0	6.1 ± 1.9
	Listening effort	4.7 ± 2.4	4.5 ± 2.4
	Total average SSQ	5.1 ± 1.4	5.0 ± 1.3
Bimodal	Can you control your own voice?	6.9 ± 2.2	7.2 ± 2.1
	Can you hear the difference between a question and confirmation?	7.1 ± 1.6	7.0 ± 1.6
	Can you understand one person in quiet? *	7.6 ± 1.5	8.0 ± 1.3
	Can you understand a small group in quiet?	6.1 ± 1.6	6.3 ± 1.8
	Can you understand one person at a party? *	4.4 ± 1.9	5.0 ± 1.9
	Can you localize sounds?	4.7 ± 2.5	4.6 ± 2.3
	How do you judge the timbre of sounds? *	5.4 ± 2.1	6.4 ± 1.8
	Average	6.0 ± 1.2	6.4 ± 1.2

PREFERENCE

At the end of the study, we performed a test to assess the subjective preference of the subjects for the two bimodal conditions tested in this study (Table 4). A significant difference between bimodal fittings was only observed in the ratings given to listening comfort of 75 dB SPL sounds (Z=-2.454, p=0.014): the AGC-matched HA was rated better (3.8) than the standard HA (3.4).

At the end of the study, nine subjects (1, 3, 4, 5, 6, 8, 11, 13, 14) chose to take home the AGC-matched HA, one subject (12) preferred the standard HA and the other five subjects (2, 7, 9, 10, 15) did not have a preference. Averaged across HAs and speaker configurations, the nine subjects preferring the AGC matched HA showed a bimodal benefit for speech understanding in noise of 2.9 dB, compared to only 0.5 dB for subjects without preference and 1.8 dB for the subject that preferred the standard HA. Interestingly, bimodal benefit for speech understanding in noise was on average 1.0 dB higher with the AGC-matched HA than with the standard HA for the nine subjects preferring the AGC-matched HA, comparable to the 1.3 dB found for the group of five subjects without a preference.

3

Table 4. *Mean (± SD) ratings given in the subjective preference test at the end of the study for the standard and AGC-matched bimodal fitting. *p < 0.05*

Subjective preference	**Bimodal fitting**	
	Standard	AGC-matched
Clarity 65 dB SPL speech	4.0 ± 0.6	4.0 ± 0.7
Comfort 65 dB SPL speech	3.8 ± 0.5	3.9 ± 0.6
Clarity 75 dB SPL speech	3.9 ± 0.5	3.9 ± 0.7
Comfort 75 dB SPL speech*	3.4 ± 0.9	3.8 ± 0.5

FACTORS CORRELATING WITH BIMODAL BENEFIT

We did not find a significant correlation between the difference in benefit between both HAs or the average benefit per speaker configuration and any of the subject characteristics.

DISCUSSION

To our knowledge, this is the first study to demonstrate additional bimodal benefit for speech understanding in single-talker noise, achieved by matching the AGC characteristics of the HA to that of the CI. The commercially available HA (syllabic compression only) resulted on average in 0.7 to 2.5 dB bimodal benefit over the four speaker configurations tested, which was improved by an additional 0.6 (S0NCI) to 1.9 dB (S0NHA) when using a HA that was especially engineered to mimic the dual AGC-loop of the CI processor. Questionnaires and a subjective preference test were also in favor of the HA with matched AGC. These findings of improved bimodal benefit with AGC-matched devices were furthermore supported by the fact that the majority of the subjects preferred the AGC-matched HA for use after the study.

AGC MATCHING

Matching AGC characteristics across devices minimizes differences in the sound processing of dynamically changing signals like speech. Therefore, the positive results we found with the AGC-matched HA may partly be the result of reduced interaural mismatches in loudness, causing less conflicting binaural information and increased listening comfort. AGC matching possibly improved binaural processing, since an extra bimodal benefit of 1.9 dB was found when single-talker noise was presented from the HA side, which is the least favorable SNR side when comparing bimodal stimulation to unilateral CI use. Possible binaural cues likely depend on the overlapping input frequency range between devices and the loudness balance between ears. Because temporal fine-structure ITD cues are highly distorted or absent after CI signal processing, ILDs are thought to be the most important localization cue for bimodal listeners, even though ILDs are much less pronounced for the lower frequencies that are audible through the HA in these subjects (Francart et al., 2013). A prospective study is needed to investigate the exact contribution, or availability, of ITDs and ILDs in these subjects, e.g. in a sound-localization task.

The AGC time constants of the standard Naida HA are comparable to the fast time constants of the Harmony CI processor (1 versus 3 ms attack time, and 50 versus 80 ms release time, for the HA and CI, respectively). However, the dual-loop AGC system of the CI processor also comprises a slow component (240 and 1500 ms attack and release time), not matched by the standard HA (Boyle et al., 2009). Speech at a conversational level, without sudden increases in signal level, normally does not trigger the fast loop of the CI processor, but it does trigger the syllabic compressor in the HA. Therefore, conversational speech and other soft to moderate signals will create unmatched binaural input, possibly leading to conflicting ILD cues and increased listening effort. Since compression is applied on the total mixture of incoming sounds, syllabic compression may reduce perceptual segregation by applying a common component of modulation to independent sound sources (Stone et al., 2007). This problem is reduced for slow compression, which preserves most temporal fluctuations in speech signals. From that point of view, the impact of turning syllabic compression into a slow-acting system would be higher in our speech tests with single-talker noise than with stationary noise. Apart from matching, subjects possibly also may have benefitted from the more veridical temporal envelope received with the dual-AGC HA. On the other hand, fast-acting compression can improve the SNR by amplifying speech during temporal dips of the noise. Since percentages phonemes correct were not significantly different between the two HA-only conditions, we believe that most benefit can be attributed to matching the AGCs.

Apart from reengineering time constants, the HA's multichannel compression (operating independently in different frequency bands), was converted to single-channel compression as is present in the CI processor. When using single- and multi-channel compression in contralateral ears, mismatches in ILDs could be different in each frequency band of the multichannel system. However, this probably played a minor role because the frequency range of residual hearing in most of our subjects only spanned a small number of compression channels of the original HA, and speech signals usually do not exhibit many low-frequency fluctuations.

Choosing equivalent AGCs in both devices is a promising first step in improving bilateral hearing. Real-time synchronization of the compression in the CI and HA may further help in preserving ILDs across the dynamic range, by ensuring equal gain at both ears at all times. Wireless ear-to-ear communication has already been shown to improve sound localization and SSQ scores in the speech, spatial and quality domain of bilateral HA users (Kreisman et al., 2010; LB, 2008; Smith P, 2008; Sockalingam R, 2009). With synchronized compression, speech understanding in noise has been reported to improve by 8 to 14 percent in a simulation study with normal hearing subjects (Wiggins et al., 2013). Other studies also involved synchronized noise-reduction systems or microphone directionality modes, which resulted in better sound-localization abilities and equivalent or even improved speech understanding in noise (Ibrahim I, 2012; Kreisman et al., 2010). Possibly, the benefit is larger for bimodally fitted subjects who have little or no access to ITDs, or the advantage may be smaller than in bilateral HA fittings because of the mismatches in frequency ranges and different signal traveling paths in acoustic and electrical hearing.

If the ILDs indeed improved because of AGC matching, an accurate loudness balance between the CI and HA could potentially lead to even more benefit. We made an attempt by matching stimuli with 45 and 80 dB SPL input in different frequency bands, but ideally loudness should be matched over the whole dynamic range. However, developing a procedure to establish equal loudness growth in both devices across the whole input frequency range would also require adjustments in the CI fitting, which fell beyond the scope of the present study.

BIMODAL BENEFIT

Averaged across subjects, the combined use of a HA and CI resulted in improved performance compared to CI-only measurements for all speaker configurations. The benefit presently found agrees with results reported in earlier studies, both qualitatively and quantitatively. More benefit is observed when noise is presented from the CI side than from the HA side, and more benefit is found with competing talker than with stationary speech-weighted noise (Morera et al., 2005; Spriet et al., 2007; Tyler et al., 2002). Measured benefits with the standard HA were consistent with earlier reports: 1.1 dB in S0NHA and 2.5 dB in S0NCI (Ching et al., 2007). With the AGC-matched HA, a bimodal benefit of up to 3.1 and 3.0 dB was found in the S0NCI and S0NHA configurations, which fits well with results from the better bimodal users reported in other studies (Ching et al., 2007). Note that 50% speech reception thresholds were found at positive SNRs for all subjects and speaker configurations, indicating the task difficulty for these subjects. For speech understanding in quiet, we did not find an improvement for bimodal hearing over the CI-only condition, in line with several other studies (Ching et al., 2001b; Tyler et al., 2002). However, subjects with better residual hearing have been shown to benefit from using a HA in conjunction with the CI for understanding words in quiet (Mok et al., 2006; Potts et al., 2009). In addition, in these earlier studies there was often more room for improvement in the CI-only condition than in our study.

Because of time constraints we measured speech understanding in noise for the CI-only condition only at the end of our study. Since all subjects were fully familiarized with the task, we do not think that procedural learning could have influenced our results. It should, however, be kept in mind that at the end of our study, subjects may have been unfamiliar with the CI-only listening situation, after having used bimodal stimulation on a daily basis for at least four months. This could have overestimated the amount of bimodal benefit, a recurring issue in bimodal research. Note however, that this could not have affected our main outcome on the difference in bimodal benefit between the standard and the AGC-matched HA.

The benefit of bimodal stimulation over unilateral CI use can be explained by several underlying mechanisms. The redundancy effect states that there is an advantage of listening with both ears, even when speech and/or noise come from the same direction (Ching et al., 2007), possibly compensating for the noisiness in each auditory/cognitive pathway. Secondly, the HA adds acoustic low-frequency information that is not available through the CI, often referred to as complementarity of information. This is thought to aid in voice pitch perception, which can help to segregate sound streams of different voices, as present in several of our speech in noise tasks (Cullington et al., 2010). Improvements in speech recognition can also arise because of 'better ear glimpsing', the ability to detect the target signal during spectral or temporal dips of the masker, creating short periods with a favorable SNR (Li et al., 2008). In our study, we presented two types of noise to investigate the effect of different mechanisms. Single-talker noise was used to simulate a realistic listening environment, including different cues to segregate multiple speech sounds, e.g. differences in voice pitch, spatial separation and better-ear glimpsing. For comparison, energetic masking with the stationary speech-weighted noise included no other factors than spatial separation. However, we did not observe a difference in performance between stationary noise and single-talker noise in the S0N±90 speaker configurations. Possibly, the two single-talker noise sources, each presented from one side, were together too dense to allow for listening in the dips.

In both S0NHA and S0NCI, "squelch" implies that the impact of the masker is diminished in binaural listening by comparing interaural time, level and/or spectral differences (Tyler et al., 2003). Furthermore, the head shadow effect (HSE) plays a role when noise is presented from the CI or HA side, by creating a more favorable SNR at the contralateral side. In the S0NHA speaker configuration, we found significantly more benefit with the AGC-matched HA than with the standard HA, possibly indicating more squelch when assuming that the amount of redundancy and complementarity remains the same. The near absence of additional benefit from AGC matching when the noise source was positioned at the CI ear was possibly caused by the poor residual hearing in our subjects; since the HSE mainly attenuates the higher frequencies that are less audible through the HA, ILD cues may become unavailable. It should be further investigated how redundancy, complementarity, the HSE and squelch contribute to the bimodal benefit in each of the speaker configurations and if these are indeed simple additive processes.

In the S0NHA configuration, we found significant spatial release from masking, when moving the noise source away from the signal as in S0N0. The SRM of 3.8 (standard HA) and 4.4 dB (AGC-matched HA) found in this speaker configuration is consistent with another bimodal study (Gifford et al., 2014) and can be attributed to a combination of the HSE and squelch. When noise was moved from the front to the CI side, subjects did not benefit from spatial release from masking. A possible explanation for the absence in SRM is the asymmetrical hearing performance between the CI and the HA, since the CI was the dominant device for speech understanding, being masked more in S0NCI than S0N0.

In questionnaires, we found a small but significant improvement for the AGC-matched HA compared to the standard HA for the timbre of sounds and for understanding one person in quiet and in noise. These findings were not confirmed with the SSQ, which did not show differences between both HAs. Even though nine out of fifteen subjects chose the AGC-matched HA for use after the study, the SSQ was not sensitive enough to capture these subjective preferences. Mean SSQ ratings were comparable to another study with bimodal subjects, for all subscales (Noble et al., 2008).

Although only shown for the combination of one CI processor and one type of HA, our results agree with the assumption that equivalent compression systems in general are superior to unmatched systems. Because of the overall additional benefit and the smaller inter-subject variability found in this study, we would recommend to fit AGC-matched devices in bimodal stimulation. At this point, we cannot give any recommendations on other combinations of bimodal devices, in the absence of sufficient details of AGC processing in CI processors and HAs.

CONCLUSION

This study showed that a HA with AGC characteristics matched to that of the CI processor resulted in better speech understanding in noise as compared to a standard HA, a finding that was supported by positive questionnaire responses and usage preferences. Depending on the speech in noise speaker configuration, matching AGCs improved bimodal benefit insignificantly by 0.6 dB when noise was presented from the CI side to a significant 1.9 dB for noise from the HA side, adding up to a total 3 dB bimodal benefit in both S0NCI and S0NHA.

Questionnaires showed better results for the AGC-matched HA for speech understanding with one speaker in quiet and in noise, and for the quality of sounds. The AGC-matched HA was ranked best for listening comfort of a slightly raised voice. After the study, nine out of fifteen subjects preferred to continue to use the AGC-matched HA, one preferred the standard HA and five patients did not have a preference for either one of the HAs.

Our findings encourage the use of a CI processor and HA with matched AGC characteristics for bimodal use. A limitation of the present study is that our subjects had rather poor residual hearing. Possibly, larger benefits can be achieved in bimodal listeners who have better audiometric thresholds across a wider frequency range.

CHAPTER 4

HORIZONTAL SOUND LOCALIZATION IN COCHLEAR IMPLANT USERS WITH A CONTRALATERAL HEARING AID

Published as *Veugen, L.C., Hendrikse, M.M., van Wanrooij, M.M., Agterberg, M.J., Chalupper, J., Mens, L.H., Snik, A.F. and van Opstal, A.J., 2016. Horizontal sound localization in cochlear implant users with a contralateral hearing aid. Hearing research, 336, pp.72-82*

ABSTRACT

Interaural differences in sound arrival time (ITD) and in level (ILD) enable us to localize sounds in the horizontal plane, and can support source segregation and speech understanding in noisy environments. It is uncertain whether these cues are also available to hearing-impaired listeners who are bimodally fitted, i.e. with a cochlear implant (CI) and a contralateral hearing aid (HA).

Here, we assessed sound localization behavior of fourteen bimodal listeners, all using the same Phonak HA and an Advanced Bionics CI processor, matched with respect to loudness growth. We aimed to determine the availability and contribution of binaural (ILDs, temporal fine structure and envelope ITDs) and monaural (loudness, spectral) cues to horizontal sound localization in bimodal listeners, by systematically varying the frequency band, level and envelope of the stimuli.

The sound bandwidth had a strong effect on the localization bias of bimodal listeners, although localization performance was typically poor for all conditions. Responses could be systematically changed by adjusting the frequency range of the stimulus, or by simply switching the HA and CI on and off. Localization responses were largely biased to one side, typically the CI side for broadband and high-pass filtered sounds, and occasionally to the HA side for low-pass filtered sounds.. HA-aided thresholds better than 45 dB HL in the frequency range of the stimulus appeared to be a prerequisite, but not a guarantee, for the ability to indicate sound source direction.

We argue that bimodal sound localization is likely based on ILD cues, even at frequencies below 1500 Hz for which the natural ILDs are small. These cues are typically perturbed in bimodal listeners, leading to a biased localization percept of sounds. The high accuracy of some listeners could result from a combination of sufficient spectral overlap and loudness balance in bimodal hearing.

INTRODUCTION

In normal hearing, sound localization in the horizontal plane relies predominantly on interaural differences in arrival time and intensity of the sound reaching our ears. According to Rayleigh's duplex theory, interaural time differences (ITDs) dominate at low frequencies below 1.5 kHz, and interaural level differences (ILDs) are most effective at high frequencies above 3 kHz (Blauert, 1997; Rayleigh, 1907).

It is unclear whether binaural cues are available to hearing-impaired users of a cochlear implant (CI) in one ear, and a conventional hearing aid (HA) in the other ear ("bimodal" stimulation). Binaural cues can only arise in a frequency range that is audible through both hearing devices, which is typically the range from the low-frequency cut-off of the CI, at about 250 Hz, up to the frequency where hearing in the non-implanted ear becomes too poor for amplification (often between 750 and 4000 Hz). Numerous benefits, including improved speech understanding and sound-source localization, have been reported for the distinct, but complementary combination of acoustic HA amplification and electrical stimulation from the CI (Beijen et al., 2010; Ching et al., 2007; Mok et al., 2006; Morera et al., 2005; Veugen et al., 2016b). However, it is unclear whether these benefits result from true binaural integration at the brainstem level, or from alternative processes that depend on essentially monaural cues.

Binaural cues are highly distorted in bimodal stimulation for a number of reasons (Francart et al., 2013). (i) The envelope-encoding algorithms used in CI processors eliminate access to temporal fine structure, thus abolishing the potential for low-frequency ITD processing. (ii) Devices typically operate independently, thereby distorting or even inverting ILDs when more gain is applied to the signal that is attenuated by the head shadow (Dorman et al., 2014b). (iii) In the common case of low-frequency residual hearing, the CI and HA only overlap in the lower frequencies, where natural ILD cues are minimal. As such, bimodal listeners might have to rely on other localization cues that are typically less important for normal-hearing listeners (Macpherson et al., 2002). For example, ITDs based on the envelope of a sound, rather than on the fine structure from its carrier, could potentially convey location information (Henning, 1974). Monaural spectral pinna cues may also provide spatial information, which has been demonstrated for listeners without access to reliable binaural cues (Agterberg et al., 2012; Van Wanrooij et al., 2004; Van Wanrooij et al., 2007). However, these high-frequency cues (4-12 kHz) are probably not useful for bimodal listeners, as they fall often beyond their residual hearing, and are poorly, or not at all, preserved by hearing devices with behind-the-ear and in-the-concha microphones (Otte et al., 2013). Although bimodal listeners could in principle rely on subtle low-frequency monaural loudness cues that are caused by the acoustic head shadow, these cues are ambiguous, as they contain mixed information of both sound-source azimuth and intensity (Van Wanrooij et al., 2004).

Even if ITDs and ILDs are highly distorted in bimodal hearing, the brain might be sufficiently plastic to use all available cues, provided that these are consistent and unique (Hofman et al., 1998; Van Wanrooij et al., 2005). Bimodal users are sensitive to both ILDs and envelope ITDs when stimuli are presented directly on the electrode array, or acoustically via inserted earphones (Francart et al., 2008; Francart et al., 2009). It is unclear, however, to what extent these cues are preserved in commercially available devices. In sound-localization experiments that used speech or broadband stimuli, bimodal benefit over unilateral CI use has been observed in about 50% of the listeners (Ching et al., 2007), with bimodal behavior ranging from chance level to near-normal localization.

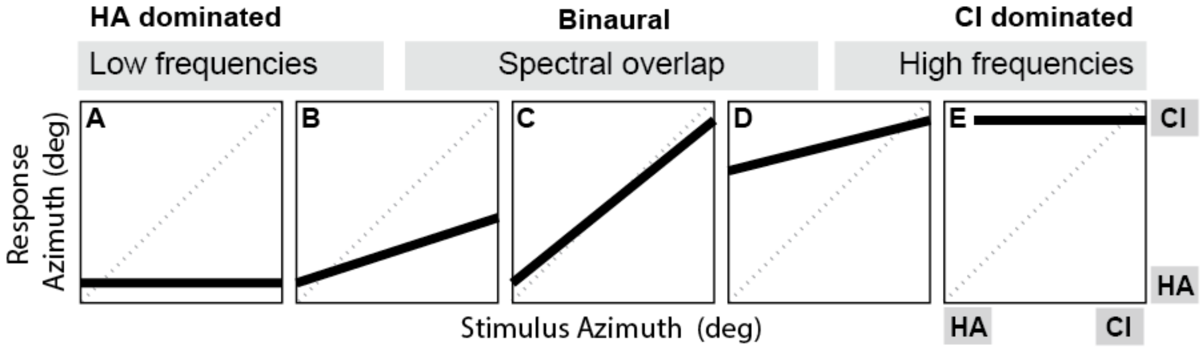

Figure 1. *Theoretical stimulus-response plots of bimodal listeners for sound stimuli of different frequency bands and for different listening conditions. Lines indicate the hypothesized stimulus-response regression line.* ***A), E)*** *Responses are lateralized to respectively the hearing aid (HA) or cochlear implant (CI) side, indicating strong spatial perception of sounds, yet perturbed by a mismatch in hearing.* ***B), D)*** *A residual stimulus-response localization relationship is present, but responses are still biased towards the stronger ear.* ***C)*** *Accurate localization behavior similar to normal hearing in the presence of sufficient spectral overlap between both ears and normal binaural cues. Transition from A-E might be observed through manipulation of the mismatch in hearing and the spectral overlap of sounds by varying the sound's frequency content from low, HA range to high, CI range, and in the extremes by turning off the CI or HA, respectively. Not shown: chance behavior (average response at 0 deg), and no perception (no response).*

The present study aimed to determine the contribution of binaural cues (ILDs, fine structure and envelope ITDs) and monaural cues (loudness, spectral) to the horizontal sound localization behavior of bimodal listeners, by systematically varying the sound's frequency band, level and envelope. If there is no contribution of any localization cue, or if bimodal listeners would rely on cues that cannot be transmitted by the devices (such as fine structure ITDs, or spectral pinna cues), one expects that bimodal listeners simply cannot report a spatial percept. This would lead them to report only one fixed location (e.g. at straight ahead), or to completely random localization behavior, independent of the actual sound location. Alternatively, bimodal listeners could fully rely on the contribution of ILDs for sound localization, as this cue could potentially be preserved by the hearing devices. In particular, if bimodal listeners perceive a sound's location by ILDs alone, we predict that localization responses will be biased towards the dominant device in the sound's frequency range (Dunn et al., 2005). Figure 1 schematically illustrates stimulus-response relationships for different stimulus conditions when the contribution of ILDs is dominant.

Likewise, for monaural listening conditions we predict that stimuli will be perceived on the aided side, similar to the localization behavior of single-sided deaf and normal-hearing listeners with one ear plugged (Agterberg et al., 2011; Kumpik, 2010; Van Wanrooij et al., 2004; Van Wanrooij et al., 2007). Responses are then expected on the CI side in the monaural CI condition, but also in the bimodal condition for high-pass filtered sounds that fall outside the range of residual hearing through the HA (Figure 1E). An opposite effect towards the HA side is expected for low frequencies that are well audible only through the HA (Figure 1A). Accurate localization behavior with a clear stimulus-response relationship (Fig. 1C) is only expected when there is considerable bimodal spectral overlap in hearing (when the sound contains frequencies transmitted by both devices) that allows for access to veridical ILDs or to envelope ITDs. Furthermore, we predict that observed differences in aided hearing thresholds will largely explain individual differences in bimodal localization behavior.

METHODS

SUBJECTS

A group of fourteen postlingually deaf bimodal listeners participated in this study (nine male, mean age 63 ± 11 years, range 45-81 years). All used on a daily basis a Harmony or Naida Q70 CI processor (Advanced Bionics, Valencia, CA) in one ear, and a Naida S IX UP hearing aid (Phonak, Stäfa, Switzerland) in the other ear, that was adapted in compression characteristics for research purposes (see below). Figure 2 shows the average aided and unaided hearing thresholds in the non-implanted ear, as determined by standard audiometry. For unaided thresholds, pure tones were presented through headphones; aided thresholds were measured for eleven subjects in a sound field with warble tones. To visualize the possible areas of binaural overlap, we also added CI-aided thresholds that were measured for nine subjects during their standard clinical examination. Subject and device characteristics are presented in Table 1. The study was approved by the local medical ethics committee (CMO) Arnhem-Nijmegen, the Netherlands (protocol number 40327.091.12).

At the time of the experiment, all subjects were bimodal users for at least one year. The CI and HA were matched in loudness and automatic gain control, according to a procedure described before (Veugen et al., 2016a), at least two months prior to this study, and used every day since then. Briefly, loudness matching was performed using steady-state speech-shaped noise, at two loudness levels (45 and 80 dB SPL) and in three frequency bands (250 – 548 Hz, 548 - 1000 Hz and 1000 Hz up to the frequency where hearing loss in the non-implanted ear exceeded 120 dB HL), therefore called 'three-band balancing'. Compression knee-points were the same in both devices, as well as the attack and release times (Veugen et al., 2016b). Adaptive features including noise reduction and directional microphones were turned off in both devices (only the adaptive feedback reduction in the HA was activated).

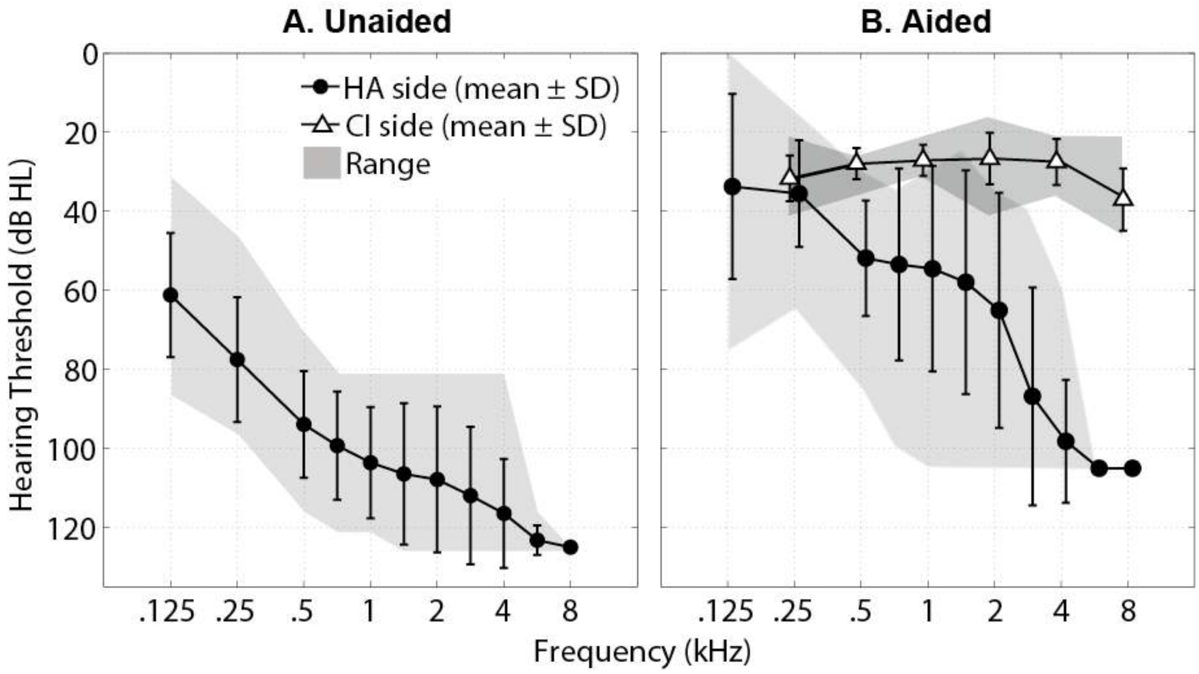

Figure 2. A) *Unaided pure tone hearing thresholds (mean ± 1 SD and range) in the non-implanted ear (n=14). Thresholds beyond the audiometer limit (120 dB HL) were assigned a value of 125 dB HL.* ***B)*** *Aided warble tone free-field thresholds using the hearing aid (HA) (n=11) and the CI (n=9). Thresholds beyond the sound field audiometer limit (100 dB HL) were assigned a value of 105 dB HL.*

APPARATUS

The experimental setup was the same as described before (Bremen et al., 2010). Briefly, all experiments took place in a completely dark, sound-attenuated room. Sounds were presented via a motorized hoop with 58 speakers that rotated around the subject's chair. Head movements were recorded using the magnetic search coil induction technique (Agterberg et al., 2011; Robinson, 1963), for which subjects wore a custom-built light-weight spectacle frame with a small search coil attached to the nose bridge.

STIMULI

Acoustic stimuli were digitally generated using Tucker-Davis System3 hardware (Tucker-Davis Technologies, Gainesville, FL) with a real-time processor (RP2.1 System3, 48828 Hz sampling rate). After attenuation by custom-built amplifiers, the stimulus was sent to one of the speakers (SC5.9, Visaton GmbH, Haan, Germany) on the hoop.

***Table 1.** Demographic details of the bimodal subjects.*

Subject	Age (years)	Gender	Ear implanted	Etiology	CI experience (years)	Bimodal experience (years)	Age onset HL (years)	CI processor
1	65	M	L	Unknown	8	8	38	Harmony
2	48	M	L	Genetic	5	5	23	Naida Q70
3	81	F	L	Hereditary	10	1	53	Naida Q70
4	60	M	R	Ototoxicity	7	7	9	Harmony
5	76	M	R	Genetic	3	3	47	Harmony
6	54	F	R	Hereditary	7	1	19	Harmony
7	45	M	R	Congenital	5	5	3	Naida Q70
8	66	M	R	Unknown	5	1	32	Harmony
9	80	F	R	Genetic	8	1	60	Harmony
10	65	F	R	Hereditary	7	7	30	Harmony
11	60	M	R	Genetic	6	1	14	Naida Q70
12	63	M	L	Hereditary	1	1	37	Naida Q70
13	61	F	R	Genetic	2	2	44	Naida Q70
14	61	M	R	Meningitis	2	2	58	Naida Q70

CI=cochlear implant, HL = hearing loss

4

PARADIGMS

COIL CALIBRATION

To obtain head-position data for the calibration procedure, subjects fixated on 18 LEDs with known locations from -90 to +90° in azimuth (-90°, -75°, -50°, 0°, 50°, 75°, 90°) and -40 to +60° in elevation (-40°, 0°, 40°, 60°), while the coil signals were recorded for 100 ms.

SOUND LOCALIZATION

For all sound-localization experiments, a trial started with the presentation of a central green LED on the wall, in front of the subject. Subjects were instructed to direct the laser pointer attached to the spectacle frame to this LED, to ensure that the subject's head was always oriented in the same starting position for each trial, and to press a button when ready. Then, the central LED was turned off within 100-300 ms and a sound stimulus was presented after a random delay between 300 and 500 ms. The subject was asked to accurately and rapidly orient the laser pointer to the perceived sound location. Head-movement recording started after LED extinction (200 ms before the onset of the sound) and lasted either for 2500 ms (for 150 ms noises and words) or 6000 ms (for words). For every experimental session and individual subject, stimulus locations were uniformly randomly selected in azimuth between -75°and +75° at a resolution of 1 deg and in elevation between -30° and +30° at a resolution of 2.5 deg. An elevation component was included to mimic a natural listening environment with small variations in spectral content. The data revealed (see Introduction) that all subjects localized sounds inaccurately in elevation; we therefore do not further report on sound-elevation localization, but retained the azimuth component for further analysis. Unless stated otherwise, 20 locations were tested for each combination of stimulus type and intensity.

STIMULI

Stimuli were presented in blocks, in which sound types, intensities, frequency bands and locations were interleaved pseudo-randomly.

BLOCK WITH VARIOUS STIMULI

Stimuli consisted of Gaussian white noise with a duration of 150 ms and 5 ms sine-squared on- and offset ramps, and were band-pass filtered in three ways: broadband noise (BB, 250 to 6000 Hz), high-pass filtered noise (HP, 2000 to 6000 Hz) and low-pass filtered noise (LP, 250 to 1500 Hz). A pulsed broadband noise with additional envelope timing cues compared to the standard broadband noise was presented with a pulse rate of 50 pulses per second, an exponential ramp with 2 ms flank and a width of 5 ms at half maximum (Monaghan et al., 2013), also with a total duration of 150 ms. For BB and HP noise, sounds were presented at 45, 55, 65 and 75 dB, A-weighted [dB(A)], for LP noise at 55 and 65 dB(A), and for pulsed BB at 55 and 75 dB(A). We also added 20 Dutch monosyllabic words at 65 dB(A), randomly taken from the standard Dutch speech recognition test NVA (Nederlandse Vereniging voor Audiologie) (Bosman, 1992), with additional envelope timing cues and a longer duration (500-1000 ms) compared to the standard noise bursts. All stimuli were tested with both the CI and HA on, which is the bimodal hearing condition. Two additional monaural conditions with only the HA on (HA-only) or the CI on (CI-only) were also tested, but only for BB noises at 45, 55, 65 and 75 dB(A).

For seven subjects, we shortened the duration of the experiment, by leaving out the monosyllabic words, omitting 45 dB(A) stimuli, presenting pulsed-BB stimuli only at 65 dB(A), and testing only twelve locations per intensity in the monaural conditions.

Table 2. *Overview of the presented stimuli. Availability of binaural cues to normal hearing subjects for these sounds is indicated as 0 (unavailable), ~ (weak) or + (available). All stimuli were tested in the bimodal condition; the CI- or HA-only condition was tested only when indicated.*

Stimulus	Subjects	Freq range (Hz)	ILD	TFS ITD	Env ITD
Broadband	14	250-6000	+	+	~
CI-only	13				
HA-only	12				
High frequencies	14	2000-6000	+	0	~
Low frequencies	14	250-1500	~	+	~
Pulsed broadband	14	250-6000	+	+	+
Monosyllabic words	7	100-4000	+	+	+

Freq=frequency; ILD = interaural level differences, ITD = interaural time differences. TFS = temporal fine structure, CI=cochlear implant, HA = hearing aid.

NARROW-BAND NOISE BLOCK

Because many subjects showed a distinct response bias towards the CI side for HP noise (as in Figure 1E) and to the HA side for LP noise (as in Figure 1A), we added an extra stimulus block for four subjects (S2, S7, S13 and S14). For this group, we systematically varied the frequency content of the stimuli, aiming to find a frequency range for which localization switched from the HA side to the CI side, potentially approaching high behavior localization as illustrated in Figure 1C. High-pass and low-pass cutoff frequencies were individually determined after analysis of the localization responses to the LP, BB and HP stimuli. For subjects 2, 7 and 14, localization of the LP stimuli (250 – 1500 Hz) was concentrated on the HA side, while localization of the BB stimuli (250 – 6000 Hz) was dominated by the CI. To investigate this transition, we presented additional stimuli in between these frequency ranges, thus with upper cutoff frequencies around 2 - 3 kHz. The lower cutoff frequency was always 250 Hz, except for subject 13. Because subject 13 could localize the BB sounds quite accurately, but lateralized the LP stimuli on the HA side, we presented several stimuli to this subject with cutoff frequencies in between the LP and BB sounds (2-4, 1-6, 2-6 and 3-6 kHz). For subject 14, we also presented stimuli from 250 – 1000 Hz, to see if these would be lateralized more towards the HA than the standard LP stimuli. We presented blocks with stimuli of two or three different frequency ranges, spanning the range between CI dominated and HA dominated responses. If needed, this procedure was repeated until the transitional frequency range was found. Stimuli were tested at two intensities, at 65 and 75 dB(A). We tested twelve locations per combination of frequency band and sound level.

Table 2 gives an overview of the different conditions and stimulus characteristics. Because of the large set of stimuli and time constraints of several subjects, we could not test all stimuli in all subjects.

DATA ANALYSIS

Head movement signals were analyzed in Matlab, after calibrating the head-position signals based on the calibration experiment, as described before (Van Barneveld et al., 2013). Using custom written software, head movements were automatically detected when the velocity exceeded 20 °/s. Onset and intercept markings were checked visually by the experimenter, without knowledge of the stimulus' location, and manually adjusted if necessary.

For ease of comparison, we defined azimuth as positive on the CI side and negative on the HA side, in all figures and analyses. To analyze the localization responses, we fitted linear regression lines through the data points of each subject and for each of the different stimuli, sound levels (although pooled for graphical purposes in Figure 3), and hearing conditions, as follows:

$$\alpha_R = intercept + slope \cdot \alpha_T \tag{1}$$

The azimuth components α_T and α_R represent stimulus location and response (in degrees) respectively. The intercept (in degrees) and slope (dimensionless) describe individual localization behavior. The slope (dimensionless) represents the change in response per degree change in stimulus location and is often termed the response gain, with values approaching one indicating accurate relationships; a slope of zero (horizontal flat line) indicates no systematic linear relationship between stimulus location and response. The intercept (in deg) is the response location where the regression line intersects the y-axis for the stimulus $\alpha_T = 0°$, thus representing the response bias. An intercept < 0 indicates that sounds were predominantly perceived on the HA side and, conversely, an intercept > 0 means that responses were dominated by the CI. Perfect localization would result in a slope of one and an intercept of zero degrees.

In eight cases very few data points could be collected (four times in HA-only, twice in CI-only and twice for pulsed BB; at different intensities) when subjects did not hear most of the stimuli. Because a reliable regression analysis was impossible for these cases, they were excluded from regression and further statistical analyses.

STATISTICS

Statistical analyses were performed using IBM SPSS Statistics (Version 20, IBM, NY, USA). To overcome the unbalanced structure of our data set (caused by missing values in a few subjects for certain intensities and stimuli), we used Linear Mixed Models (LMM). Values for the localization slope and intercept were tested separately to compare localization behavior between the different stimuli and conditions. A LMM allows modeling of measurable categorical levels as fixed factors and variation within subjects as random effects.

First, to analyze localization of BB stimuli in the different listening situations, testing was performed with fixed factors Listening condition (3 levels: bimodal, HA-only and CI-only) and Stimulus intensity (4 levels: 45, 55, 65 and 75 dB(A)). Secondly, testing was performed on bimodal data (not including unimodal data) with fixed factors Frequency band (3 levels: BB, HP, LP) and Stimulus intensity (4 levels) to analyze the effects of stimulus frequency. Lastly, testing was performed to investigate the effect of time envelope cues with fixed factors Envelope cue (3 levels: BB, pulsed BB and monosyllabic words) and Stimulus intensity (3 levels: 55, 65 and 75 dB(A)). In all tests, we included a random Subject factor to deal with idiosyncratic biases. Post-hoc pairwise comparisons were performed using Bonferroni-adjusted unweighted means. Given the number of tests, significance was accepted at a conservative level of 0.01. Data were always reported as mean values ± 1 inter-subject standard deviation. The F-statistic was reported with numerator and denominator degrees of freedom in parentheses.

Localization slopes were compared to the median HA-aided hearing thresholds in the frequency range of each stimulus type. Thresholds beyond the sound field audiometer limit (100 dB HL) were assigned a value of 105 dB HL. For the three subjects with unknown aided thresholds, we used estimations made by the HA fitting software (Phonak Target 3.0.3).

RESULTS

EXAMPLE LOCALIZATION BEHAVIOR

Bimodal users were instructed to localize sounds of varying frequency bands, with both devices turned on (bimodal hearing) and under HA-only or CI-only listening conditions. Figure 3 visualizes the stimulus-response relations and the fitted regression lines for broadband sounds for three different subjects in the different listening conditions, as well as for stimuli with different frequency contents in the bimodal listening condition. Typically, as exemplified for one subject in Figures 3A-C, many localization responses to broadband sounds (BB) were biased towards the CI side if both devices were turned on (Figure 3B,E). Turning off the HA had little effect compared to the bimodal stimulation, as responses were still biased towards the CI side (Figure 3C). In stark contrast, turning off the CI elicited a response bias towards the opposite, HA side (Figure 3A). Such a shift in response bias from CI to HA side, albeit to a lesser extent, could be elicited under bimodal conditions by manipulating the frequency content of the stimulus, as exemplified for another subject in Figure 3D-F. When the availability of high-frequency content was limited by presenting low-pass filtered sounds, responses were concentrated on the HA side (Figure 3D). Responses for high-pass filtered and broadband noises in the bimodal condition were often biased towards the CI side (Figure 3E,F). Notably, for certain sound bands, when both devices were turned on, some subjects' localization was quite good; responses of the best bimodal performer (subject 4, Fig. 3G-I) changed in a systematic fashion with stimulus azimuth (BB slope = 0.8) with no appreciable response bias for BB sounds (Figure 3H; BB intercept = 2.5 deg), although responses were still quite variable.

OVERALL LOCALIZATION BEHAVIOR

To obtain the general, average localization behavior we determined the mean absolute error (MAE). On average, the MAE was 50 ± 18 ° for BB sounds, 49 ± 13 ° for LP sounds and 52 ± 18° for HP sounds, indicating a general poor behavior by our bimodal listeners.

However, since the MAE is an absolute, composite measure that confounds (absolute) intercept, slope and residual variance, it will not be able to show the differences in localization bias induced by the various stimuli (e.g. Fig. 3). Therefore, to quantify localization behavior for all subjects, conditions, frequency bands and sound levels, we applied linear regression. We determined the intercept as a measure of response bias (e.g. device dominance), and the slope as a measure of the stimulus-response relationship (see Methods, Eq. 1). These are visualized in Figure 4 (top and bottom rows, respectively) and will be elaborated upon in the next few sections.

MONAURAL VERSUS BIMODAL LISTENING

Responses in the monaural listening conditions were lateralized towards the aided side for all subjects (Figure 4B): the intercept for BB sounds was on average 52 ± 24° in the CI-only condition and -58 ± 18° in the HA-only condition. Responses to BB sounds in the bimodal listening condition were perceived in between, but more towards the CI side, at 33 ± 28°. As a result, we found a significant difference in the intercept for BB sounds between the bimodal, CI-only and HA-only listening conditions ($F(2,111)=212.37$, $p<0.001$). The intercept in the bimodal condition was significantly larger compared to HA-only ($p<0.001$) and significantly smaller than CI-only ($p=0.001$). Also the HA-only and CI-only condition differed significantly from each other ($p<0.001$). No effect of stimulus intensity was found ($F(3,110)=0.53$, $p=0.66$).

Slope values of the best-fit stimulus-response regression in the two monaural conditions were close to zero, 0.02 ± 0.11 for HA-only and 0.00 ± 0.15 for CI-only (Figure 4F). In the bimodal condition, the average slope was 0.18 ± 0.32. We found a significant difference in slope across conditions ($F(2,112)=11.17$, $p<0.001$). The slope for bimodal listening was significantly higher than HA-only ($p=0.002$) and CI-only ($p<0.001$). On a group level no effect of stimulus intensity was found ($F(3,112)=1.12$, $p=0.34$).

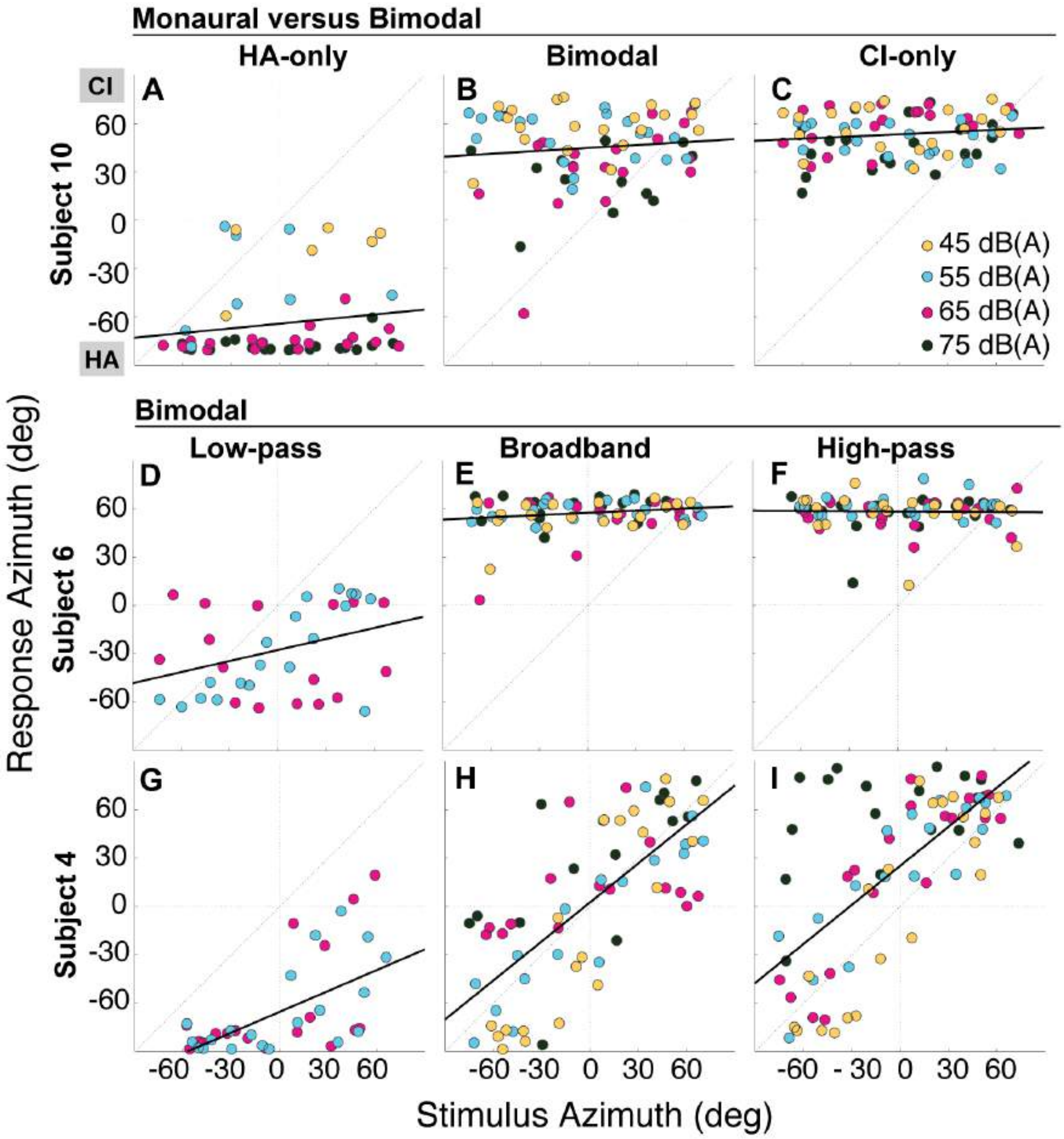

Figure 3. *Typical example stimulus-response relationships in azimuth, for **A-C)** broadband stimuli in the **A)** HA-only, **B)** bimodal and **C)** CI-only condition; and **D-I)** in the bimodal condition, for **D,G)** low-pass, **E,H)** broadband and **F,I)** high-pass filtered sounds; for **D-F)** a typical subject with large response intercepts and small slopes and for **G-I)** the subject with the highest response slope and smallest intercept for the broadband stimuli. Marker color represents stimulus intensity, including 45 (small markers), 55, 65 and 75 (large markers) dB(A). Bold black lines denote best-fit regression lines, pooled over all intensities.*

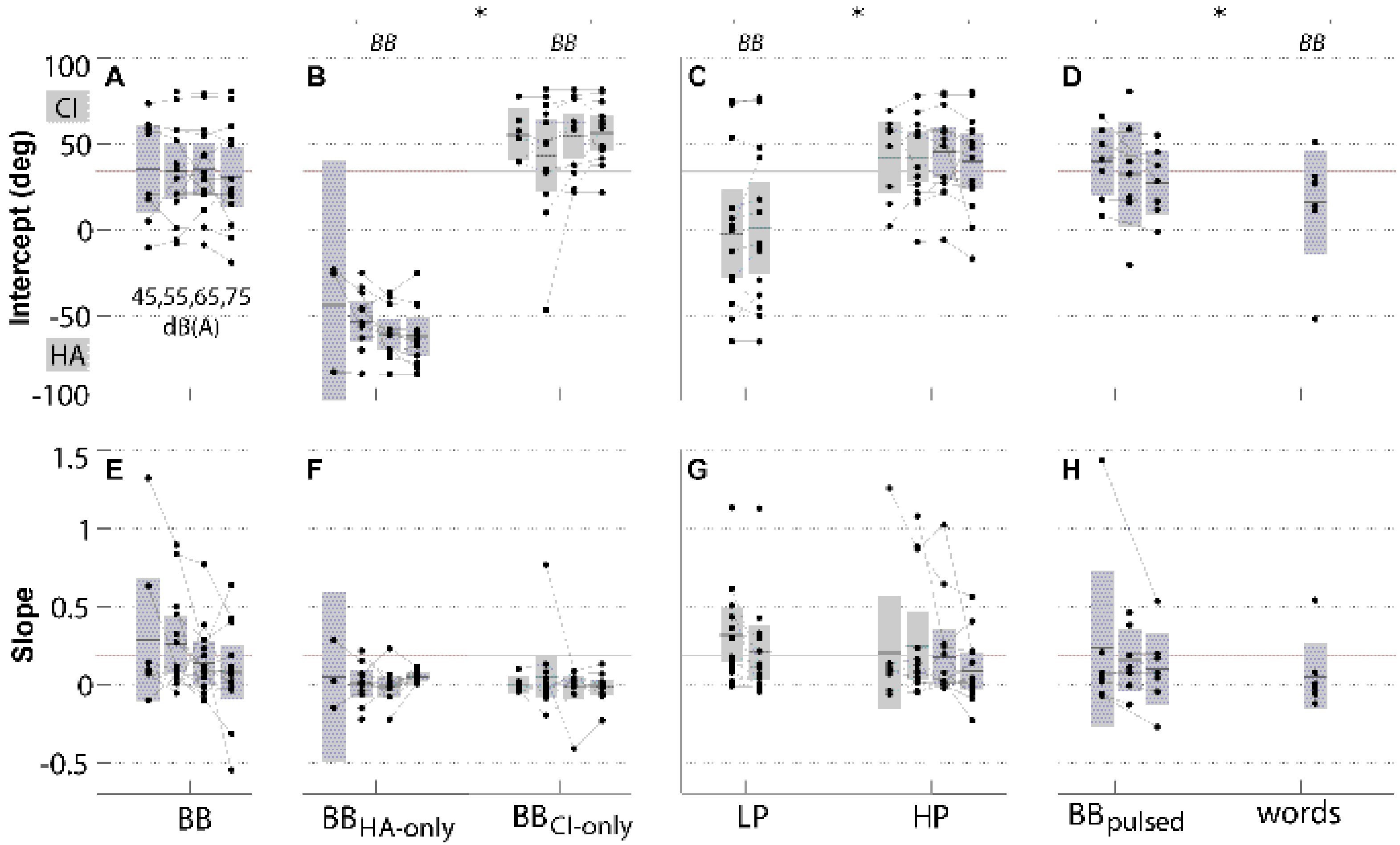

Figure 4. *Individual data points (dots) and the mean ± 95% confidence intervals (bar) per intensity (45, 55, 65 and 75 dB(A)) per stimulus for* ***A-D)*** *localization intercepts and E-H) localization slopes.* ***A,E)*** *Broadband (BB) stimuli in the bimodal condition.* ***B,F)*** *BB stimuli in the HA-only and CI-only condition.* ***C,G)*** *Low-pass (LP) and high-pass (HP) filtered stimuli.* ***D,H)*** *Pulsed broadband noise and monosyllabic words. Black lines connect data points from each individual. Mean values of the bimodal BB condition are visualized in all subplots by the horizontal gray line. Perfect localization responses have a slope of 1 and an intercept of 0 °.* *$*p < 0.01$. The letters 'BB' indicate a significant difference with BB stimuli at $p<0.01$.*

THE EFFECT OF STIMULUS BANDWIDTH

Response intercepts were on average -1 ± 44° for LP sounds and 44 ± 25° for HP sounds (Figure 4C). This means that on average HP sounds were perceived on the CI side and LP sounds towards the center in front of the subject, in most cases with low localization slopes. . The 33 ± 28° intercept for BB sounds was in between the LP and HP sounds, although more biased towards the CI side. Intercepts were significantly different between stimuli ($F(2,105)=46.12$, $p<0.001$): the intercept for LP sounds was significantly smaller compared to both BB and HP sounds (both $p<0.001$). No significant difference was found between the intercept of the BB and HP stimuli ($p=0.08$). No effect of intensity was found ($F(3,105)=0.41$, $p=0.8$).

Slopes of the BB, LP and HP sounds were on average 0.18 ± 0.26, 0.18 ± 0.29 and 0.26 ± 0.28 respectively (Figure 4G) and did not significantly differ from each other ($F(2,105)=0.80$, $p=0.45$). We obtained a significant effect of stimulus intensity ($F(3,105)=6.22$, $p=0.001$), showing that low-intensity sounds led to higher slopes than high-intensity sounds: the slope for 75 dB(A) stimuli was significantly lower compared to 45 ($p=0.005$) and 55 ($p=0.001$) dB(A) sounds.

In summary, the differences in intercept demonstrate a clear, albeit perturbed spatial percept for the stimuli as a function of bandwidth, but all with small localization slopes, lacking a clear stimulus-response relationship.

4

ENVELOPE CUES

We tested two stimuli with additional envelope cues, as compared to the standard BB stimuli, including monosyllabic words at 65 dB(A) and pulsed BB noise bursts at 45, 55 and 65 dB(A). The intercept for pulsed BB stimuli at 33 ± 24° was comparable to the value of 34 ± 27° for standard BB stimuli (at 45, 55 and 65 dB(A)). Monosyllabic words had an intercept of 16 ± 33° (Figure 4D). Testing across standard BB, pulsed BB and monosyllabic words, we obtained a significant difference between intercepts ($F(2,51)=6.00$, $p=0.005$). The intercept for monosyllabic words was smaller than both standard BB ($p=0.003$) and pulsed BB ($p=0.008$) sounds. There was no effect of stimulus intensity ($F(2,51)=1.70$, $p=0.19$).

The slope was on average 0.22 ± 0.32, 0.16 ± 0.35 and 0.05 ± 0.22 for respectively standard BB noise, pulsed BB noise and monosyllabic words (Figure 4H). No significant differences were found between stimuli ($F(1,52)=0.26$, $p=0.76$). No effect of stimulus intensity was found either ($F(2,52)=4.17$, $p=0.02$. This suggests that subjects did not use the additional time envelope cues present in the pulsed BB and word stimuli, as compared to the on- and offset cues in the standard BB stimulus.

NARROWBAND NOISE BLOCK

Because responses to BB stimuli were lateralized more towards the CI than expected, we tested additional stimuli with subjects S2, S7, S13 and S14. The goal of this experiment was to assess whether bimodal sound localization could be improved by adjusting the frequency content of the stimulus (see Methods 2.4.2). In all four subjects, adjustment of the frequency range of the stimulus (Figure 5) systematically influenced their bimodal sound localization behavior. The overall pattern across the listeners was again that the intercept changed from the HA side to the CI side with increasing high frequency content in the stimulus. In three out of four subjects, we observed a higher localization slope for sounds with a frequency band that fell between the LP and BB stimuli.

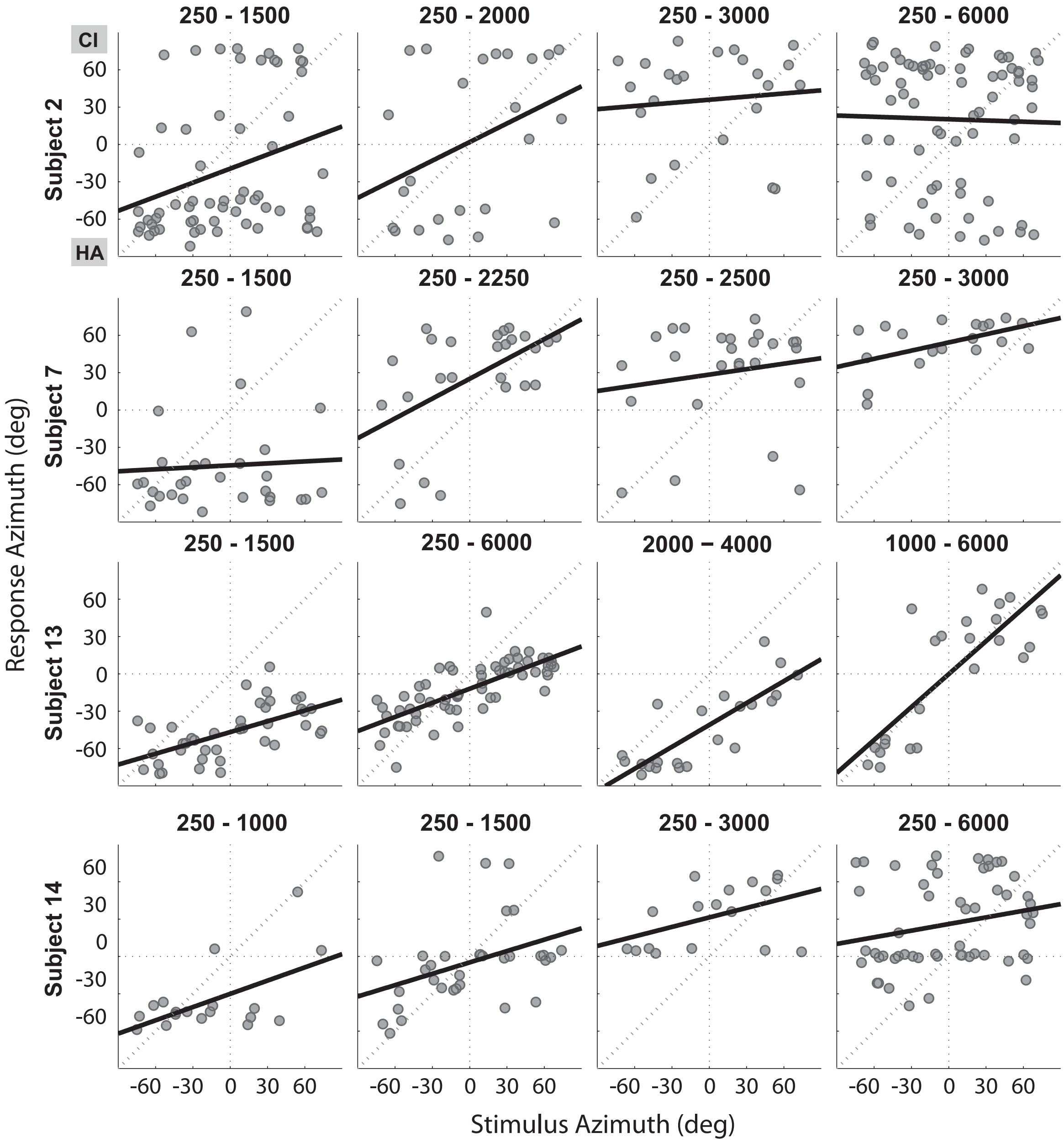

Figure 5. *Stimulus-response plots for four subjects for the band-pass filtered noise bursts. The frequency band of the stimulus in Hz is denoted above the plots. Bold black lines denote best-fit regression lines. Dotted lines denote perfect localization.*

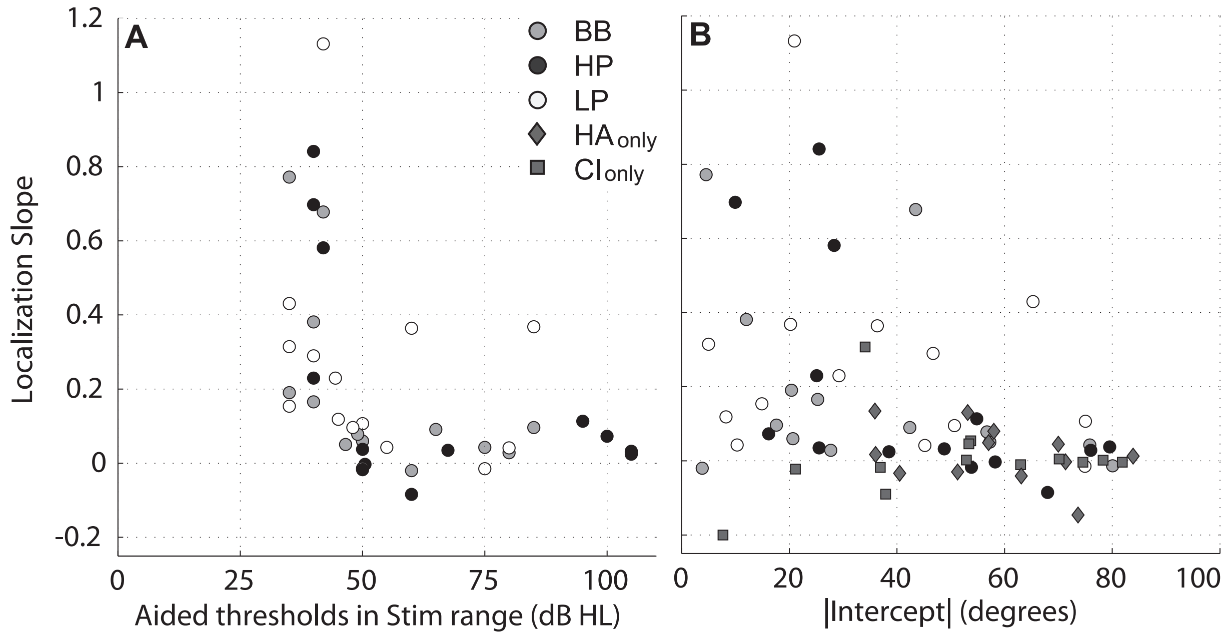

Figure 6. *Individual localization slopes (averaged over stimulus levels) versus* ***A)*** *the median HA aided hearing thresholds in the frequency band of the broadband (BB), high-pass (HP) and low-pass (LP) filtered stimuli; and* ***B)*** *the absolute localization intercept, also for HA- and CI-only.*

4

DISCUSSION

The main goal of the present study was to investigate horizontal sound localization behavior in bimodal listeners, fitted with a cochlear implant and a hearing aid in opposite ears. This is the first free-field sound localization study that thoroughly tested whether bimodal stimulation allows for the use of interaural level and/or time difference cues. We found a systematic and significant response bias towards the CI side that shifted gradually to the HA side with decreasing high frequency content (Figures 3-5). A minority of subjects, having HA-aided thresholds better than 45 dB HL, were able to indicate the sound-source direction.

BIMODAL LOCALIZATION

Our results show that the sound's frequency content can play a crucial role in horizontal sound localization behavior of bimodal listeners. The localization intercept for high-pass filtered sounds was biased towards the CI side, while low-pass filtered sounds were perceived towards the center. Even though responses often lacked an accurate stimulus-response relationship, the difference in intercept demonstrated a clear, yet perturbed, percept for stimuli with a different bandwidth. Responses were likely influenced by a loudness imbalance across the CI and the HA. If binaural input is out of balance, responses will shift to the louder side (Figure 4A-D), explaining the observed bias towards the CI for high frequency stimuli that are often beyond residual hearing of the non-implanted ear. In some cases, an adequate loudness balance seemed to have been achieved, resulting in small response intercepts, and higher response slopes (Figure 6B). The results of the four subjects for whom we could adjust localization behavior by fine-tuning the bandwidth of the stimulus (Figure 5) can be explained by a reduced loudness imbalance across frequencies.

Assuming that accurate sound localization can only arise from integrated binaural input, we argue that the sound localization behavior of bimodal subjects is best explained by the amount of audible spectral overlap between both ears. HA-aided sound detection thresholds below 45 dB HL seemed a necessary, but not a sufficient requirement for adequate horizontal sound localization in bimodal listeners (Figure 6). This was even the case when these lower thresholds only applied to a limited frequency range of the stimulus; for example, one of the subjects able to localize high-pass filtered stimuli (Figure 3I) had a threshold of 40 dB HL at 2 kHz and immeasurable thresholds at higher frequencies (>105 dB HL from 3 kHz upward). Aided thresholds through the CI approached normal hearing levels around 30 dB HL from 250 Hz to 8 kHz (Figure 2B). It thus seems that sufficient audibility of frequencies accessible to both ears is required for integration of the CI and HA input, to allow for high-behavior sound localization.

In the bimodal condition, broadband sounds (250-6000 Hz) were strongly lateralized towards the CI side for almost all subjects (Figure 4A), demonstrating an imbalance between HA and CI. On average, lateralization was absent for low-pass filtered stimuli (250-1500 Hz; Figure 4C). These sounds were fully transmitted by the CI and, for most subjects, also well audible through the HA over its entire frequency range (Figure 2). However, individually, many subjects still lateralized LP sounds, about half of the group to the HA side and half to the CI side (Figure 3D,G; Figure 4C). This finding may seem unexpected given the loudness matching applied to the devices (see Methods; (Veugen et al., 2016a)). Note, however, that the three-band balancing procedure was not intended to improve spatial hearing per se, but to enhance speech recognition. A possible explanation for the apparent discrepancy might simply be that three-band balancing between the CI and HA was performed with steady-state speech-shaped sounds that differ in spectral shape from the Gaussian white noise stimuli of the localization task. Moreover, a balanced loudness across both ears for improved speech perception does not necessary result in a merged auditory percept and binaural sensitivity, as has also been shown for bilateral CI users (Goupell et al., 2013a). An ideal bilateral processing scheme would achieve balanced loudness with a centered auditory percept for mid-sagittal sound locations and audibility in both ears for all possible combinations of frequencies to support accurate localization. In addition, synchronized timing could further optimize interaural correlations and possible cues. However, even if technically and clinically possible, the question remains how a fitting that is optimized for sound localization, would affect speech perception. Improved sound localization possibly yields better spatial segregation cues that might benefit speech perception in noise, but the required signal processing could at the same time degrade speech quality and intelligibility.

Idiosyncratic differences in the daily localization strategies of our listeners may also have contributed to inter-subject variability. Some subjects may predominantly live in quiet environments with stable ILD cues, whereas others may have learned to ignore these cues, having experienced inconsistencies in different noisy listening environments. In addition, a potential pitch mismatch between the ears, which depends on the insertion depth of the CI, is known to affect ILD and ITD perception in bilateral CI users (Kan et al., 2015). Pitch mismatch might possibly play a role in bimodal listeners as well, so that even in the presence of loudness balance and spectral overlap, bimodal sound localization could still be poor (Figure 6). The limited dynamic range of hearing and the compression applied by the CI and HA may also have affected localization behavior. Sounds above the compression knee point (63 dB SPL in both the CI and HA) become more similar in loudness, thus impairing veridical ILD perception. This would explain the higher (i.e., better) localization slopes found for low-intensity sounds (45 and 55 dB(A)) below the compression knee-point, as compared to high-intensity sounds (75 dB(A)).

Behavior of our best subject (root mean square [RMS] error = 33° pooled across intensities) corresponded well with comparable studies of Dunn et al. (2005) (2010) and Ching et al. (2004) who observed best bimodal behavior around 30° RMS error (Ching et al., 2004; Dunn et al., 2005; Dunn et al., 2010). Seeber et al. (2004) reported results from one bimodal subject who showed a 9° MAE. For comparison, normal hearing localization behavior yields an RMS error of only 2-10° (Dorman et al., 2015). Note that here we assessed localization behavior in a completely dark environment without any feedback on behavior. Bimodal listeners possibly have adopted new strategies to orient in daily life, e.g. by relying more on visual cues. Psychophysical experiments in naturalistic environments, with more visual and auditory background sources, potential distracters, and feedback (e.g. Corneil et al., 2002; Van Wanrooij et al., 2010; Van Wanrooij et al., 2009) may complement the current study, and test how audiovisual integration affects sound localization in bimodal listeners.

A limitation of the current study is the relatively small group size, especially for the experimental block with narrow-band stimuli that was so far only tested in four subjects. In addition, our subjects had rather poor residual hearing. Localization behavior might be better in the expanding group of bimodal listeners with better hearing thresholds (Gifford, 2011). This would be supported by our finding of higher localization slopes with HA-aided thresholds below 45 dB HL. Also, the best performer in the study of Seeber et al. (2004) had better residual hearing than any of our subjects (68 dB HL at 1 kHz, 9° MAE). Furthermore, even though all bimodal listeners in the present study had at least one year of experience with bimodal stimulation, some of them had possibly not yet reached their full potential regarding binaural integration. Full adaptation to bimodal listening may perhaps take up to two years (as in the case of bimodal pitch perception: (Reiss et al., 2015).

LOCALIZATION CUES

The question remains which sound localization cues were used by the bimodal listeners. Access to ITDs (normally considered vital for localizing low frequency sounds) is thought to be impossible, since the fine-structure of sounds is lost as a consequence of the

CI's signal processing algorithms (Rubinstein, 2004). Francart et al. (2009) showed that bimodal listeners can only detect ITDs above 1 kHz, which would imply that they rely on envelope ITDs instead of fine-structure cues. However, our subjects did not improve localization behavior for pulsed broadband stimuli, or for words, suggesting low sensitivity to additional envelope ITDs in these stimuli as compared to the on- and offset cues of single-burst stimuli. This finding is in line with a free-field localization study in bilateral CI users (Seeber et al., 2008). Note that words elicited a smaller response bias than the standard BB sound, but not a higher response slope. The latter is consistent with insensitivity to envelope ITDs, while the former hints at improvement due to ILD sensitivity that is less biased towards the CI side; words typically have lower-frequency content than the standard BB sounds (Table 2), so that perturbation of ILD cues are likely more biased towards the HA side similar to LP stimuli.

Even though fine-structure ITDs are the dominant cue in the low frequencies for normal hearing subjects, ILDs ranging from 1 to 10 dB are still present below 1.5 kHz (Shaw, 1974). Bimodal listeners may therefore adapt their localization strategy over time and become more sensitive to these low-frequency ILD cues. It has been shown that ILDs as small as 1.7 dB can be discriminated with bimodal stimulation (Francart et al., 2008), which would be high enough for low-frequency ILD detection. A recent study showed that bilateral CI users can lateralize sounds below 500 Hz with ILDs after CI signal processing of only 1-2 dB (Dorman et al., 2014b). Taken together, we consider it likely that ILDs were the most important, and possibly the only available, cue for the bimodal listeners in our study. This idea is further supported by the recent finding that unilateral CI users with normal hearing in the opposite ear use ILDs to determine sound source locations (Dorman et al., 2015).

MONAURAL LOCALIZATION

Localization with unilateral CI input is generally poor, with stimulus perception lateralized to the aided side (Grantham et al., 2007; Laszig et al., 2004; Litovsky et al., 2009; van Hoesel, 2012). In line with these findings, none of our subjects could successfully indicate the direction of sounds with either the CI or HA alone. This is not surprising, since monaural high-frequency spectral pinna cues (4-12 kHz) are thought to be absent in bimodal listeners because they are (1) largely beyond the audible/transmitted frequency range, (2) badly preserved by the microphones of both hearing devices and (3) possibly deteriorated by the substantial spread of excitation in the CI stimulated ear.

Previous studies showed that monaurally deaf subjects and subjects with unilateral conductive hearing loss strongly relied on the acoustic head shadow in a localization task comparable to our paradigm (Agterberg et al., 2011; Van Wanrooij et al., 2004). In our study, five subjects (two in the CI-only condition and three in the HA-only condition) lateralized the lowest-intensity stimuli towards the unaided side (or more towards the center), and the other stimuli to the hearing side (e.g. Fig 3A), which could result from using sound level as a (false) monaural localization cue. However, we observed no gradual shift in localization bias as function of sound intensity, and no seemingly accurate localization at any intermediate sound intensity, like generally expected from subjects that rely on a head shadow cue. If the bimodal subjects did attempt to use a head shadow localization cue, discrimination of the various presentation levels may have been impaired by the compression of sounds into their limited dynamic range. Psycho-acoustic measurements of loudness perception should be performed to test this idea. Note that the CI- and HA-only conditions were unfamiliar acute listening conditions for all bimodal subjects. They might therefore have ignored (or have been unable to use) any monaural cue in the monaural conditions, as they were more accustomed to bimodal or binaural cues.

CONCLUSION

Our results show that varying the sound's bandwidth had a strong effect on the localization bias of bimodal listeners, although overall localization performance often remained poor. Responses systematically changed as a function of the frequency band of sounds, or by simply switching the HA and CI on and off. Sound localization was impossible for subjects with HA-aided thresholds poorer than 45 dB HL in the frequency range of the stimulus. A minority of subjects having better HA-aided thresholds were able to indicate the sound-source direction. Their sound-localization behavior was likely based on the binaural integration of overlapping low-frequency ILD cues.

CHAPTER 5

EFFECT OF EXTREME ADAPTIVE FREQUENCY COMPRESSION IN BIMODAL LISTENERS ON SOUND LOCALIZATION AND SPEECH PERCEPTION

Published as *Veugen, L.C., Chalupper, J., Mens, L.H., Snik, A.F. and van Opstal, A.J., 2017. Effect of extreme adaptive frequency compression in bimodal listeners on sound localization and speech perception. Cochlear Implants International, pp.1-12.*

ABSTRACT

OBJECTIVES. This study aimed to improve access to high-frequency interaural level differences (ILD), by applying extreme frequency compression (FC) in the hearing aid (HA) of thirteen bimodal listeners, using a cochlear implant (CI) and conventional HA in opposite ears.

DESIGN. An experimental signal-adaptive frequency-lowering algorithm was tested, compressing frequencies above 160 Hz into the individual audible range of residual hearing, but only for consonants (adaptive FC), thus protecting vowel formants, with the aim to preserve speech perception. In a cross-over design with at least five weeks of acclimatization between sessions, bimodal performance with and without adaptive FC was compared for horizontal sound localization, speech understanding in quiet and in noise, and vowel, consonant and voice-pitch perception.

RESULTS. On average, adaptive FC did not significantly affect any of the test results. Yet, two subjects who were fitted with a relatively weak frequency compression ratio, showed improved horizontal sound localization. After the study, four subjects preferred adaptive FC, four preferred standard frequency mapping, and four had no preference. Noteworthy, the subjects preferring adaptive FC were those with best performance on all tasks, both with and without adaptive FC.

CONCLUSION. On a group level, extreme adaptive FC did not change sound localization and speech understanding in bimodal listeners. Possible reasons are too strong compression ratios, insufficient residual hearing or that the adaptive switching, although preserving vowel perception, may have been ineffective to produce consistent ILD cues. Individual results suggested that two subjects were able to integrate the frequency-compressed HA input with that of the CI, and benefitted from enhanced binaural cues for horizontal sound localization.

INTRODUCTION

The primary acoustic cues for sound localization in the horizontal plane in normal-hearing listeners are interaural time differences (ITDs, for frequencies <1.5 kHz), which arise as a result of the differences in traveling distance for the acoustic input to the two ears, and interaural level differences (ILDs, for frequencies >1.5 kHz), which result from the acoustic head shadow. As these binaural differences are identical for all locations on the so-called 'cone of confusion' and for all locations in the vertical plane (Blauert, 1997), high-frequency (>4 kHz) spectral-shape cues from the pinnae are essential to further disambiguate these cues to enable full directional hearing. Bimodal listeners, who use a cochlear implant (CI) and a conventional hearing aid (HA) in opposite ears, are thought to rely predominantly on ILDs for localization, since CI signal processing eliminates fine-structure ITDs (Francart et al., 2013). However, residual hearing in the non-implanted ear is typically restricted to the lower frequencies, where the head-shadow related ILD cues are poor (Rayleigh, 1907). In this study, we tried to enhance ILDs by compressing high-frequency information, containing potentially large ILDs, into the low-frequency audible range of residual binaural hearing. In this way, inaudible high-frequency ILD information would be made available as audible low-frequency ILD information. Although in natural environments the low frequencies do not produce significant ILDs, and therefore the intact auditory system does not need to process low-frequency ILD information, it might be conceivable that the impaired auditory system could learn to use consistently applied low-frequency ILD cues to localize sound sources, even though they are derived from intensity differences in the higher frequency bands. We therefore tested whether frequency-compressed ILDs could be helpful to establish horizontal sound localization in bimodal listeners, as for these listeners the fine-structure ITD cues are not available.

Several studies have reported an improvement in speech perception from frequency compression (FC) in HA users (Bohnert et al., 2010; Glista et al., 2009; McCreery et al., 2013; Simpson et al., 2005; Wolfe et al., 2011), most likely resulting from a better audibility of consonants (Alexander, 2013). Others, however, found no effect, neither beneficial, nor detrimental (O'Brien et al., 2010; Simpson et al., 2006). So far, only a few studies have investigated FC in bimodal listeners, but no significant benefits were found. Yet, speech and consonant perception remained unaffected in those studies (Davidson et al., 2015; Hua et al., 2012; McDermott et al., 2010; Park et al., 2012; Perreau et al., 2013). Perreau (2013) and Davidson et al. (2015) investigated bimodal sound localization of daily sounds or words, but found no improvement as a result of FC in the HA. None of these studies specifically addressed ILD perception after FC. The present study therefore set out to test the potential use of ILDs, by measuring sound-localization responses to high-pass filtered noise stimuli, which had been compressed into the audible range of the HA.

In previous studies and in commercial hearing aids, FC was typically applied to frequencies above 1.5 kHz, which does not fully match the poor residual hearing of many bimodal listeners (often < 1kHz). We therefore wondered whether more benefit could be obtained with stronger compression settings of the FC algorithm. Considering the severe hearing loss of most bimodal listeners, we applied an experimental hearing-aid algorithm, in which the compression knee-point was set as low as 160 Hz, hence further referred to as 'extreme frequency compression'. Clearly, such extreme FC potentially creates severe pitch distortions, and thus may degrade speech intelligibility. To preserve speech perception, as well as the perceived naturalness of sounds, the experimental algorithm employed signal-adaptive FC, by protecting vowel formants (i.e., leaving their harmonic structure undistorted), and only compressing the consonants into the low-frequency audible area. For normal- and hearing-impaired listeners, vowels contribute significantly more to sentence intelligibility than consonants (Kewley-Port et al., 2007), although the opposite has been reported for isolated words (Owren et al., 2006). The switching between consonants and vowels was based on the energy content of the input signal: fricative consonants are characterized by their non-harmonic broad-band frequency content, in which the high-frequency part contains potentially strong ILD information. We hypothesized that adaptive FC would also avoid the decrement in vowel perception of bimodal users equipped with a HA with FC (Perreau et al., 2013). Therefore, besides localization performance in the horizontal plane (which might improve with FC), we also performed additional psychophysical tests to determine the effects of adaptive FC on speech perception (which should not deteriorate). Finally, adaptive frequency compression, preserving low frequency temporal information in vowels, hopefully protects cues to voice pitch, which we tested by measuring the perception of pitch accents in words.

It is not clear to what extent sound localization is possible on the basis of ILDs that are spectrally compressed in one ear. The study of Goupell and co-workers (2013b) in normal-hearing listeners showed robust lateralization responses, by using stimuli with an interaural mismatch in center frequency up to 4 kHz in one ear, and 14 kHz in the other. Francart et al. (2007) demonstrated ILD sensitivity in normal-hearing listeners for signals with an interaural frequency mismatch of up to one octave. Even though just noticeable differences in ILD increased with increasing interaural frequency mismatch, the 3-5 dB thresholds for 1-octave shifts might still be usable in realistic listening environments (Shaw, 1974). Most likely, the auditory system will learn to use such artificially altered cues, only after prolonged exposure to many listening situations. The study of Dorman et al. (2015) describes another example of sound localization with interaural frequency mismatches. They showed that single-sided deaf listeners with a CI in the deaf ear had learned to adequately use ILD information for sound localization within 1-3 months after device activation.

We tested sound-localization performance and bimodal listening with extreme FC in the HA after a period of chronic use, also considering that extended use could result in a better match of pitch between the ears over time. The latter has been reported for CI users who gradually lowered their perceived pitch in the implanted ear over several octaves towards the pitch in the acoustically stimulated ear (Reiss et al., 2007).

METHODS

SUBJECTS

Thirteen experienced bimodal listeners participated in this study, all using on a daily basis a Naida S IX UP hearing aid (Phonak, Stäfa, Switzerland) in one ear, and a CI processor (Advanced Bionics, Sylmar, CA) on the contralateral ear. Note that these devices did not feature synchronized processing. Seven subjects used the Harmony processor, and six used the Naida Q70 (see Table 1 for demographic details). The devices were balanced in loudness and automatic gain control, as described before (Veugen et al., 2016a), at least two months prior to this study. This is further referred to as their 'standard HA program', which never used frequency compression or transposition. Briefly, loudness matching was performed using steady-state speech-shaped noise, at two loudness levels (45 and 80 dB SPL) and in three frequency bands (250 – 548 Hz, 548 - 1000 Hz and 1000 Hz up to the frequency where hearing loss in the non-implanted ear exceeded 120 dB HL). The hearing aids were altered to enable compression knee-points, as well as the attack and release times to be the same in both devices (Veugen et al., 2016b). We always checked the loudness balance between the CI and the HA during the fitting sessions using a running speech signal from straight ahead. In a few subjects, we increased the overall HA gain when using FC, to re-establish a loudness balance with the CI.

Adaptive features including noise reduction and directional microphones were turned off in both devices for the entire duration of the study, to avoid mismatches in signal processing between the CI and HA that could possibly disturb ILD perception. Twelve subjects were tested with both the adaptive FC algorithm, and the standard HA program (see below, study design). One subject (P8) could not complete the study because of personal problems, and was only measured with the adaptive FC algorithm. For this subject we could still compare sound localization performance with and without FC, using data of a previous study. The study was approved by the Local Ethics Committee of the Radboud University Nijmegen (protocol number 40327.091.12).

FREQUENCY COMPRESSION

Like contemporary FC algorithms, the experimental algorithm in this study compressed high frequencies above a certain knee point into the better audible low-frequency area. If the knee point is set below the upper boundary of usable hearing, part or all of the high frequencies previously inaudible become accessible to the listener. Such an approach is called 'nonlinear frequency compression' (Alexander, 2013). A particular novelty of our experimental HA algorithms over current FC techniques is that it only applies the compression to consonants, thus protecting vowels by preserving their harmonic structure. Switching between these two states is based on the spectral energy content of the input signal (i.e. output of the hearing aid's filter bank analysis; see Fig. 2B, below, for an example). We will further refer to this algorithm as 'adaptive' (applied to consonants only) FC.

Table 1. *Demographic details of the participating subjects. In all figures, subjects are indicated by the symbol displayed in the last column. The group column indicates the distribution of subjects over the cross-over design; group 1 started with the standard HA fitting and group 2 started with FC. *Subject 8 only participated in the first session*

Participant	Age (yr)	Gender	Ear	Etiology	CR	CI experience (yr)	Bimodal experience (yr)	CI processor	Group	Symbol
1	65	M	L	Unknown	4	8.3	8.3	Harmony	1	■
2	48	M	L	Genetic	4	5.7	5.7	Naida Q70	1	◀
3	81	F	L	Hereditary	4	10.9	1.9	Naida Q70	2	■
4	60	M	R	Ototoxicity	2.7	7.3	7.3	Harmony	2	★
5	77	M	R	Genetic	1.9	3.3	3.3	Harmony	2	◀
6	54	F	R	Hereditary	4	7.6	1.7	Harmony	2	▶
7	67	M	R	Unknown	1.9	4.8	1.9	Harmony	1	▲
8	80	F	R	Unknown	4	8.8	1.9	Harmony	2*	▼
9	66	F	R	Hereditary	4	7.6	7.6	Harmony	1	●
10	61	M	R	Genetic	4	6.8	2.0	Naida Q70	1	✶
11	72	M	R	Genetic	2.3	1.3	1.3	Naida Q70	2	◆
12	61	F	R	Genetic	1.9	1.9	1.9	Naida Q70	1	■
13	61	M	R	Meningitis	1.9	1.9	1.9	Naida Q70	2	●

Frequency compression was only applied above the compression knee point that was fixed at 160 Hz for all listeners. The compression ratio (CR) above 160 Hz was determined individually for each subject, aiming to map the input range between 160 Hz and 10 kHz onto the frequency range where pure-tone thresholds were below 90 dB HL. The relationship between input and output frequency for a detected consonant is described as follows:

$$\begin{aligned} f_{out} &= f_{in} && f_{in} < 160 \\ f_{out} &= f_{in}^{\frac{1}{CR}} \cdot 160^{\,1-\frac{1}{CR}} && f_{in} \geq 160 \end{aligned} \tag{1}$$

with f_{in} the original sound frequency (in Hz), f_{out} the compressed frequency (in Hz), and CR (dimensionless) the compression ratio of the algorithm.

For seven subjects the mapping procedure resulted in a frequency compression ratio that would exceed the maximum of 4; for these listeners, we fixed CR=4 (see Figure 1 for the applied input-output curves). Figure 2 visualizes a spectrogram of a speech fragment (Fig. 2A) after adaptive (Fig. 2B) and continuous static (Fig. 2C) FC, with a CR 4.0. HA-aided thresholds with and without FC were measured using warble tones with a calibrated Diagnostic Audiometer AD229e (Interacoustics, Denmark) in a sound field using a loudspeaker (JBL Control 1, Washington DC, USA) and an amplifier (MPA 4-80, Ecler, Spain) (Figure 3).

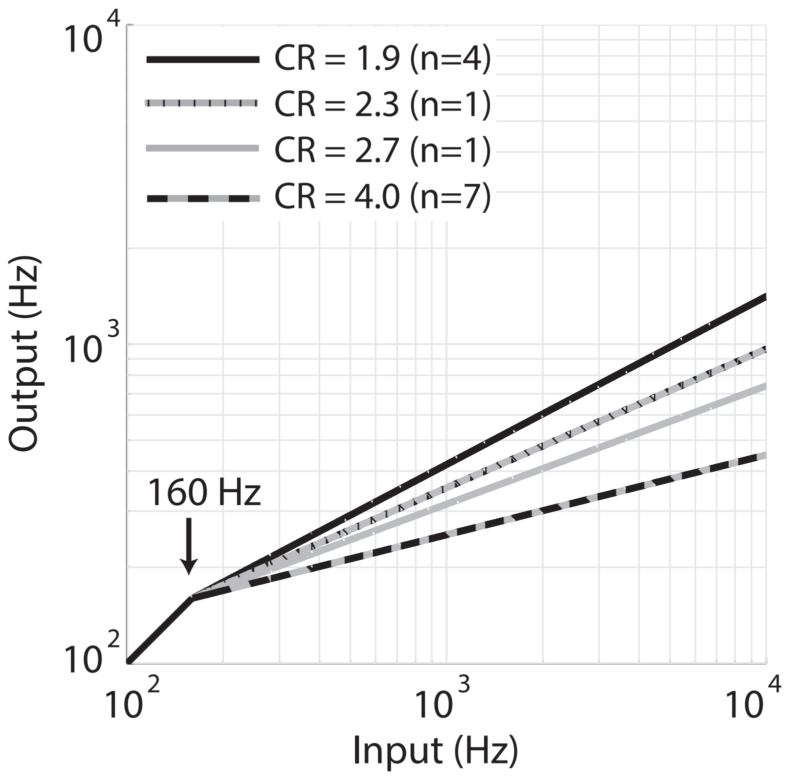

Figure 1. *Input-output characteristic of experimental frequency compression of consonants in the HA, with the four different compression ratios (CR) used in this study, according to Eqn. 1. The number of subjects tested with each CR under adaptive FC, is indicated.*

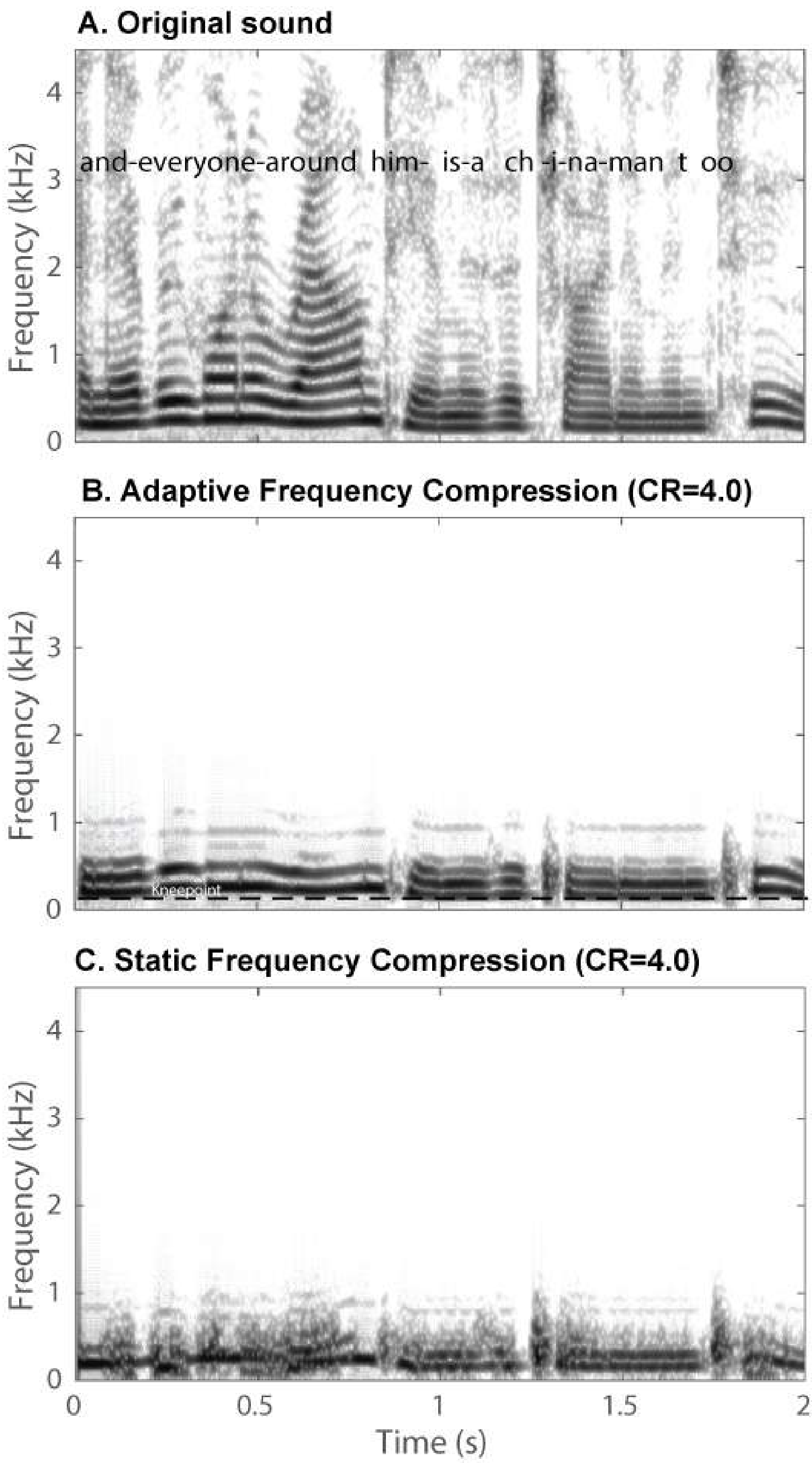

Figure 2. A) *Spectrogram of a female speech fragment, uttering the sentence "And everyone around him is a china man too".* ***B)*** *Spectrogram of the same fragment after HA processing with experimental adaptive frequency compression. The knee point of the frequency lowering was 160 Hz (dashed line) with a compression ratio of 4.0.* ***C)*** *Static frequency compression of the same fragment and compression settings.*

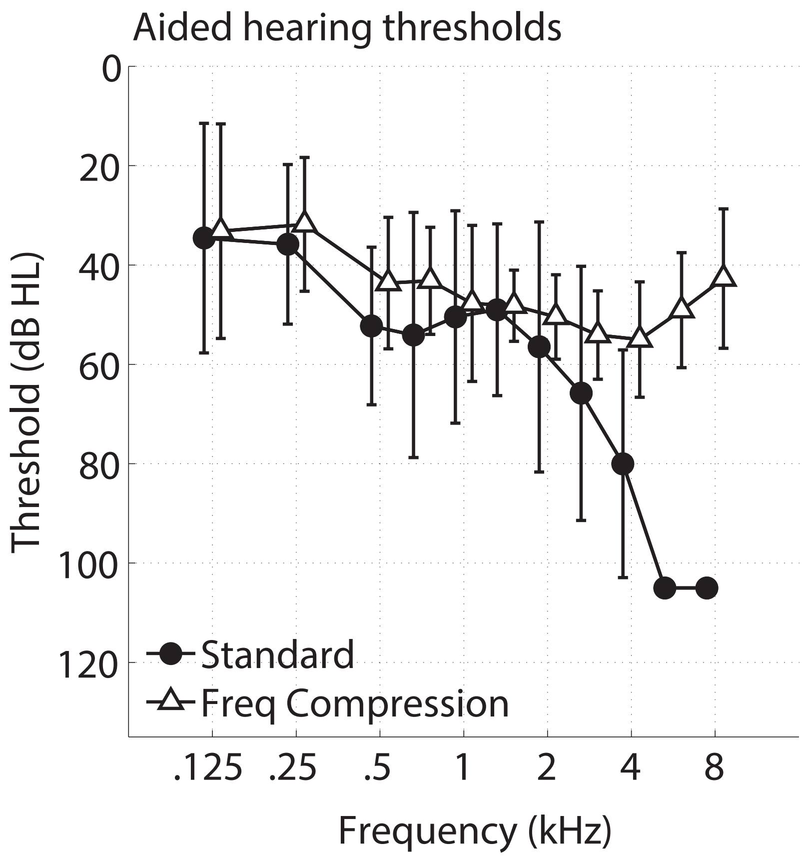

Figure 3. *Aided pure tone thresholds (mean ± 1 SD) in the non-implanted ear across listeners (n=12), with (open triangles) and without (closed circles) frequency compression. Thresholds beyond the audiometer limit (100 dB HL) were assigned a value of 105 dB HL.*

5

STUDY DESIGN

Using a three-visit cross-over design, we compared the subjects' bimodal performance using their standard HA fitting without FC to a fitting with adaptive FC. During the first session, six subjects were fitted with the standard HA program, and the other seven were fitted with adaptive FC. In the second session, bimodal listening performance was measured and thereafter subjects were fitted with the other HA program (excluding subject 8). In the third session, listening performance was measured again. The two groups of subjects were matched in age, pure-tone thresholds, gender and duration of CI and bimodal use (see Table 1). Visits were scheduled at least five weeks apart (on average 8.5 ± 2.1 weeks standard deviation) to allow for acclimatization to the bimodal hearing condition with the new HA program. The subjects had no access to another HA fitting program during their take-home period.

After the five-week take-home acclimatization period for each of the two HA fittings, we assessed performance and subjective preferences in a series of tests, described below. For all tests, performance was measured in the bimodal listening condition and sometimes also with only the HA. For some tests, CI-only performance was measured to be able to calculate the bimodal benefit.

LOCALIZATION

Horizontal sound-localization performance was tested in the bimodal listening condition as described by Bremen et al. (2010). Sounds were presented via a motorized hoop with 58 speakers that rotated around the subject's chair in a completely dark and sound-attenuated room. Subjects wore a custom-built spectacle frame with a small search coil attached to the nose bridge in order to record their head movements using the magnetic search-coil induction technique (Agterberg et al., 2011; Robinson, 1963).

Trials always started with the presentation of a green LED on the wall. Subjects were instructed to fixate the LED in front of them and then press a button that triggered the sound and turned the LED off. Subjects had to indicate the perceived direction of the sound by making a rapid goal-directed head movement. Stimulus intensities were set at comfortable and well-audible levels. We presented broadband noise bursts (250-6000 Hz) at 55 and 65 dB(A) and high-pass filtered noises (2000-6000 Hz) at 65 and 75 dB(A), both with a duration of 150 ms (including 5 ms sine squared on- and offset ramps); we applied the 10 dB sound-level roving to minimize the use of (ambiguous) monaural level cues (Van Wanrooij et al., 2004). We did not test all three stimulus intensities for both stimuli to reduce total testing time. Twenty trials per stimulus and intensity were presented in random order, and at a randomly selected location in the frontal hemifield limited to -75° to +75° in azimuth, and -30° to +30° in elevation (to avoid uncomfortable head-orienting movements). To calibrate the head coil, subjects also fixated 18 LEDs at known locations to map the corresponding head-position data. For eleven subjects, localization performance with the standard HA program had been collected already in our previous study (Veugen et al., 2016c); in that case we used their existing data for a comparative analysis. We could thus still analyze the localization data of subject 8, who was only tested on adaptive FC in this study. Unfortunately, localization data for subjects 2 and 3 could not be collected due to technical problems with the localization setup.

Head-movement signals were analyzed in MatLab (version 2014b) using custom-written software (Van Barneveld et al., 2013). For each stimulus and participant, we determined the best-fit linear regression line between stimulus and response location, respectively α_T and α_R:

$$\alpha_R = b + g \cdot \alpha_T \quad (2)$$

The intercept (b; in degrees) is the response location for stimuli presented at straight ahead; the slope (g, dimensionless) represents the change in response per degree change in stimulus location. Perfect localization would yield an intercept of b = 0° and a slope of g = 1.

SPEECH TESTS

APPARATUS

Target signals for all tests described below were presented from a speaker (JBL Control 1, Harmon International Industries, Washington DC, USA) at one meter in front of the subject, at 65 dB(A), as calibrated by a sound-level meter (Brüel & Kjaer 2260). Sound stimuli were delivered through an external sound card (RME Babyface, Audio AG, Germany) and a main amplifier (MPA 4-80, Ecler, Spain).

SPEECH UNDERSTANDING IN QUIET

We used the NVA (Nederlandse Vereniging voor Audiologie) Dutch monosyllabic word test to assess speech understanding in quiet (Bosman, 1992). Three lists of twelve words, spoken by a female speaker, were used in each listening condition: CI-only, HA-only and bimodal listening. The number of phonemes correct was calculated as a percentage from the last two lists per condition. Percentage scores were transformed to rationalized arcsine units (RAU) before statistical analysis to stabilize the error variance in the presence of floor and ceiling effects (Studebaker, 1985).

SPEECH UNDERSTANDING IN NOISE

Speech understanding in noise was measured in the CI-only and bimodal listening condition using the Matrix Sentence Test (Dutch version) (Houben et al., 2014; Theelen-van den Hoek et al., 2014). Target sentences with five pseudo-randomly selected words in a fixed grammatical order (example, translated to English: 'Mark gives five large flowers') were spoken by a female speaker with F0 at 175 Hz as determined by Praat software (Boersma et al., 2001). Sentences were presented from the front at 65 dB(A) and noise was either presented from the front (S0N0) or from the HA side (S0NHA) (to reduce testing time we did not test noise from the CI side). Subjects selected on a touch-screen which of the closed set of 50 words (5 columns of 10 words) they heard from every sentence. We used single-talker babble noise (International Speech Test Signal, (Holube et al., 2010), with shortened pauses between words and with F0 lowered to 127 Hz to make it more distinct from the target speaker's voice, possibly facilitating speech understanding. Noise started playing two seconds before the target sentence. The level of the noise was varied adaptively according to the Brand and Kollmeier procedure (Brand et al., 2002), but with a minimum step size of ± 1 dB for a correctly repeated sentence. The 50% speech reception threshold (SRT) was calculated as the average SNR level in dB of the last six sentences. Two lists of ten sentences were presented per noise configuration and listening condition; an additional list was measured if the difference in SRT between these two was larger than 2 dB SNR. Prior to testing, two training lists were used to familiarize the subjects with the task, the first one starting at an easy 20 dB SNR. The SRT after the second training list was used as a starting point for the actual measurements and this was kept the same over the two blocks of the cross-over design.

VOICE PITCH PERCEPTION

Just notable differences in fundamental frequency (F0) were determined for the bimodal listening condition as described by Straatman et al. (2010). Stimuli were based on the bi-syllabic nonsense word 'baba', produced by a female speaker with F0 at 200 Hz. Pitch accents ranging from 0.85 to 22.1 semitones (step size of 0.85 semitones) were artificially imposed on the first syllable (200 ms) of the stimulus. In an adaptive two interval, same-different procedure, subjects had to indicate whether they heard two identical stimuli (both without a pitch accent) or two stimuli that differed in word accent. The 70.7% threshold in semitones was determined from the last six out of ten reversals; the test was repeated if the standard deviation of these last six reversals was larger than two semitones. Visual feedback was provided during the test.

CONSONANT AND VOWEL PERCEPTION

A custom-written Matlab program was made to test the identification of consonants (C) and vowels (V) in a vCv and cVc format respectively. Three different vCv stimuli were used for each of the seventeen consonants /b, d, f, g, h, j, k, l, m, n, p, r, s, t, v, w, z/, which were composed like /ibi/, /aba/ and /obo/ for the consonant /b/. Four different cVc stimuli were used for each of the ten vowels /a, ɑu, e, ɛi, ø, i, u, o, œy, y/ *(aa, au, ee, ei, eu, ie, oe, oo, ui, uu) as follows: /hat/, /hag/, /pat/ and /paf/ for the vowel /a/*. All stimuli were spoken by a female speaker and presented twice in random order, resulting in 102 trials for the vCv task and 80 trials for the cVc task. Subjects were instructed to select the perceived letter on a touch-screen, which also initiated the next trial. No feedback was provided during the test.

STATISTICS

We used a Linear Mixed Model (LMM) with a random factor Subject to compare performance with the standard HA program versus adaptive FC. Localization slopes and intercepts were tested with fixed factors Condition (standard, frequency compression) and Stimulus (Broadband, High-pass noise). Speech in quiet was analyzed using the factor Device, containing six levels (CI-only, Bimodal and HA-only for the standard and frequency compression block). Speech in noise was tested using the factor Condition, separately for S0N0 and S0NHA. Paired t-tests were performed for the other tests: voice pitch perception thresholds in semitones, percentage consonant correct scores in the vCv task and for each of the five subtypes of consonants (stops, liquids, nasals, fricatives and glides) and for the percentage vowel correct scores in the cVc task. We calculated Pearson's correlation coefficient for the frequency compression ratio versus performance with the FC program for speech understanding in quiet (bimodal and HA-only), bimodal benefit in speech understanding in noise (S0N0 and S0NHA) and versus the localization slope and intercept, averaged over stimuli. Level of significance was set at $p < 0.05$.

RESULTS

LOCALIZATION

Figure 4 shows the localization slope and intercept for the broadband and high-pass filtered stimuli. Perfect localization would result in a slope of one and an intercept of zero degrees. Clearly, the listeners' responses were far from perfect under both listening conditions, and for both stimulus types. Expressed differently, the mean absolute group errors with and without frequency compression were 47 ± 13° and 47 ± 20°, respectively, for BB sounds, and 47 ± 13° and 50 ± 20° for the HP sounds. These differences were not significant (p>0.05).

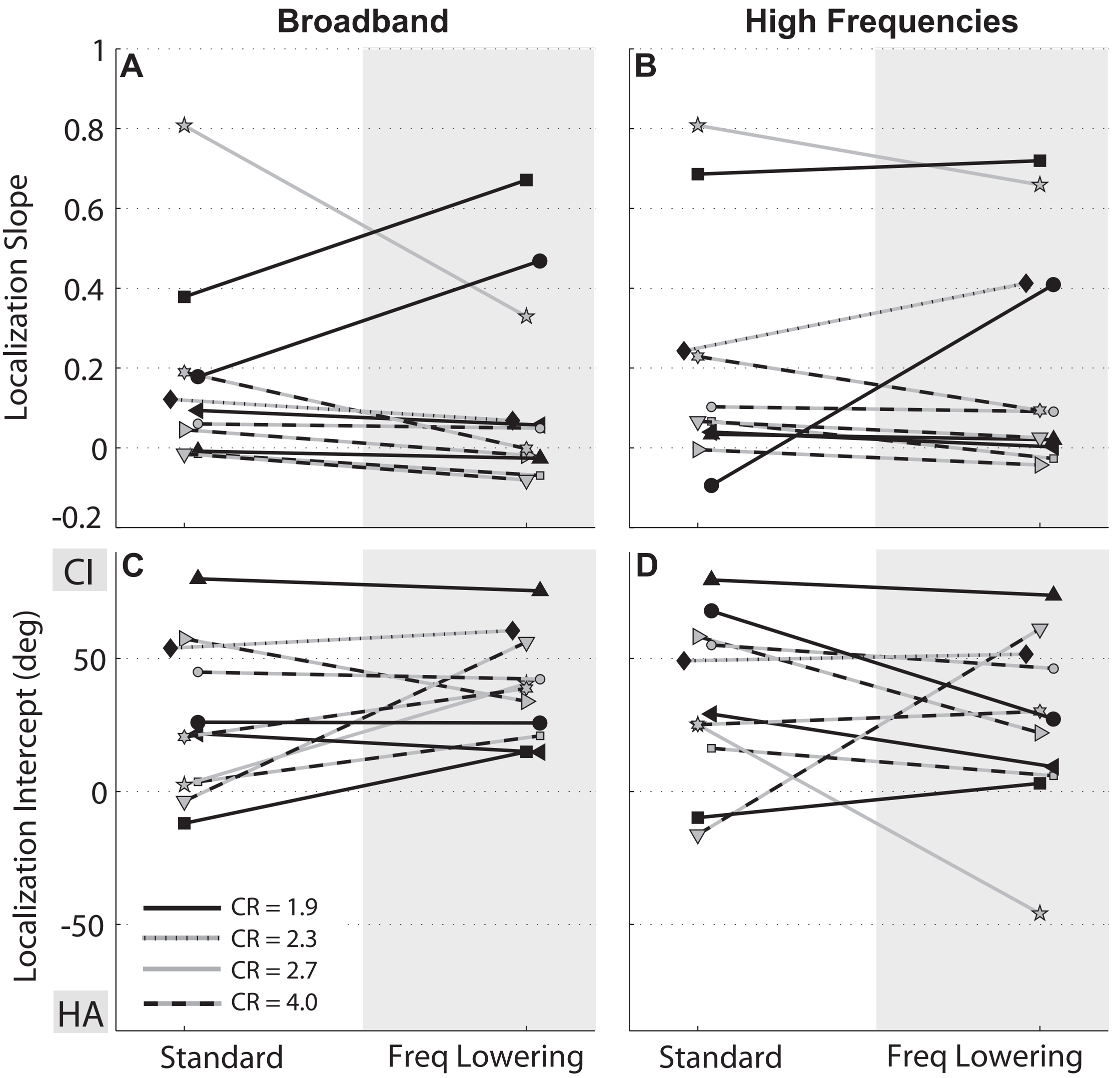

Figure 4. *Best-fit regression slopes (**A,B**) and intercepts (**C,D**) for listeners (n=11) in the bimodal condition (with and without frequency compression), for the (**A,C**) broadband and (**B,D**) high-pass filtered noise. Subjects are identified by the same marker in all panels. The line indicates the compression ratio. Perfect localization responses have a slope of 1 and an offset of 0 degrees.*

5

Frequency compression had no effect on the localization slope ($F(1,30)=0.06$, $p=0.82$), and we did not obtain a difference between the performance for BB and HP stimuli ($F(1,30)=2.09$, $p=0.16$). The localization slope with and without frequency compression was 0.13 ± 0.25 and 0.17 ± 0.24 for BB sounds and 0.21 ± 0.28 and 0.20 ± 0.29 for HP sounds.

The localization intercept (left-right bias) with and without frequency compression was directed towards the CI side for both stimuli: 39 ± 19 vs. 27 ± 29 degrees for BB sounds, and 26 ± 33 and 34 ± 31 degrees for HP sounds. No significant effect of frequency compression ($F(1,30)=0.07$, $p=0.79$), and type of stimulus ($F(1,30)=0.18$, $p=0.67$) was found on the intercept.

Interestingly, two listeners (12, 13) with the lowest CR improved their localization performance after frequency compression: these listeners had increased slopes, and a relatively small localization intercept of the stimulus-response relations with FC for both stimulus types. Note that both listeners had the lowest CR (1.9), because of better residual hearing at higher frequencies on the HA side (<95 dB HL up to 2 and 6 kHz) compared to the other subjects. In contrast, localization performance for subject 4 (CR=2.7) worsened.

SPEECH UNDERSTANDING IN QUIET

For speech understanding in quiet, we determined the percentage phonemes correct, transformed to RAU scores, in the HA-only, CI-only and bimodal listening condition. With and without adaptive FC, scores were respectively 17 ± 26 and 26 ± 25 RAU for HA-only, 76 ± 14 and 80 ± 16 RAU for CI-only and 81 ± 12 and 84 ± 9 RAU for bimodal listening. A significant difference was found between the different listening conditions ($F(1,5)=34.21$, $p<0.001$): both HA-only conditions differed significantly from all other listening conditions (all $p<0.001$). All other pairwise comparisons had p-values of 1.0, so no effect (neither negative, nor positive) of adaptive frequency compression was found.

SPEECH UNDERSTANDING IN NOISE

Speech reception thresholds at 50% correct for speech understanding in noise are displayed in Figure 5 for the different listening conditions (CI-only and bimodal) and noise configurations (S0N0 and S0NHA).

Bimodal benefit was calculated as the difference between the CI-only and the bimodal condition (Figure 5BD). On average, the bimodal benefit in S0N0 was 2.2 ± 3.7 dB with the standard HA program and 0.9 ± 4.1 dB with adaptive FC. For S0NHA these values were respectively 1.6 ± 3.0 dB and -0.3 ± 3.2 dB. However, statistical tests showed that the bimodal benefit was not significantly different from zero (S0N0: $F(1,11)=2.20$, $p=0.17$; S0NHA: $F(1,11)=2.3$, $p=0.16$), although we obtained a large variability across subjects (Figure 5BD). No significant effect of adaptive FC was found on the bimodal benefit in S0N0 ($F(1,11)=2.22$, $p=0.16$) or S0NHA ($F(1,11)=0.97$, $p=0.34$).

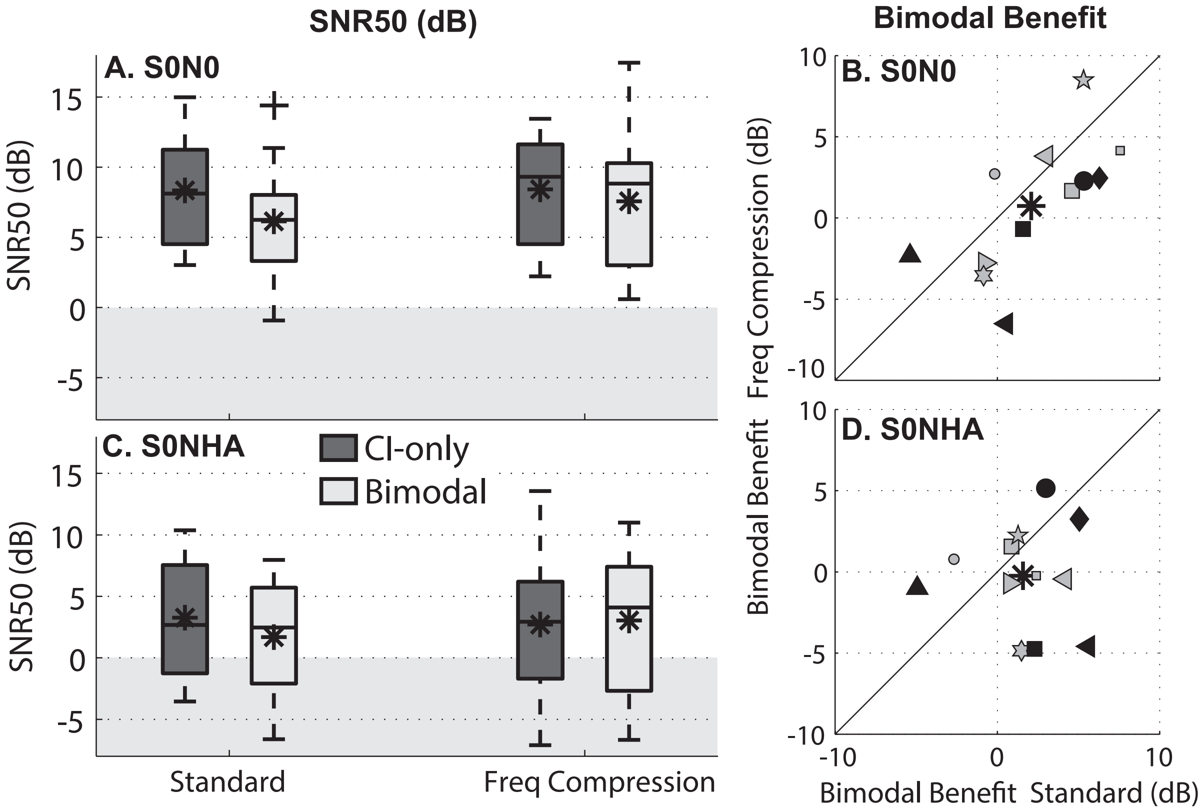

***Figure 5**. **A and C)** Box-whisker plots of the speech reception threshold in noise in the two noise configurations S0N0 and S0NHA, for the CI-only and bimodal listening conditions with and without frequency compression. SNR50 represents the signal to noise ratio at 50% correct responses (n=12). Lower values indicate better performance. Outliers beyond 1.5 interquartile range are represented by plus symbols. **B and D)** Individual scores of bimodal benefit with the standard HA program versus frequency compression. Stars indicate mean values. Subjects are identified by the same marker as in Figure 4. Points above the diagonal indicate more bimodal benefit with FC and points below the diagonal indicate more benefit with the standard HA fitting.*

VOICE PITCH PERCEPTION

On average, F0 differences at threshold in the bimodal listening condition were 6.0 ± 4.5 and 6.3 ± 5.2 semitones with and without adaptive FC in the HA, respectively, which were not significantly different from each other ($t(11)=0.27$, $p=0.79$).

CONSONANT AND VOWEL PERCEPTION

The percentage consonants correct in the vCv-task for bimodal listening was 62 ± 13 % with adaptive FC in the HA and 65 ± 12 % with the standard HA program. We did not obtain a significant difference in performance between the two HA programs for the percentage of correct consonants ($t(11)=1.55$, $p=0.15$), nor for any of its subtypes (stops, liquids, nasals, fricatives and glides; all $p>0.10$).

Vowels in the cVc-task were correctly identified in 76 ± 17 % of the trials with adaptive FC and in 75 ± 17 % with the standard HA program. No difference was found between the two bimodal listening conditions ($t(11)=-0.86$, $p=0.41$). This was expected since adaptive FC did not affect the vowels.

CORRELATIONS

The lower the compression ratio of frequency compression, the higher RAU scores for speech understanding in quiet in the HA-only condition ($r=-0.9$, $p<0.001$), which was most likely also related to the degree of residual hearing. The compression ratio did not correlate with bimodal benefit for speech understanding in noise, or with the localization intercept and slope.

SUBJECTIVE PREFERENCE

We asked subjects which HA program they preferred to use after the study. Four subjects opted for the standard HA program (1, 7, 10, 11), four subjects preferred adaptive FC (2, 4, 6, 12), and four subjects had no preference (3, 5, 9, 13). Only one of the two subjects (12, 13) that seemed to improve localization performance with adaptive FC had a preference for FC. Interestingly, the four subjects that preferred adaptive FC performed better in the various tests compared to the other two groups, when using FC, as well as with the standard HA program. They performed above average (reported as Standard score and FC score) for sound localization (slope 0.41 ± 0.38 and 0.33 ± 0.35, intercept 16 ± 37 ° and 30 ± 13 °), showed a favorable SNR50 (2.0 ± 3.4 dB and 2.2 ± 4.4 dB), consonant (77 ± 6 % and 75 ± 10 %) and vowel (88 ± 17 % and 85 ± 11 %) perception and pitch threshold (3.9 ± 3.1 smt and 2.8 ± 1.4 smt). In contrast, the HA-only thresholds of this group were not better than average and the mean compression ratio was 3 ± 1 for all three groups (also see supplementary material Table S1). Because each group consisted only of four subjects, we performed no further statistics on these data.

DISCUSSION

In the present study we tested a novel experimental frequency compression algorithm with an extremely broad operating range in the HA of thirteen bimodal listeners, with the aim to provide ILD cues for localization in the horizontal plane, while at the same time preserving the harmonic structure of vowels in speech, by applying the FC exclusively to the consonants (Fig. 2B). Even though FC was effective in producing markedly better audiometric thresholds at high frequencies (Figure 3), localization performance remained poor, when averaged across subjects. Our results therefore suggest that frequency compression does not lead to a veridical localization related ILD percept. Two of the thirteen listeners, however, seemed to improve their localization performance with adaptive FC (Figure 4). In addition, despite the absence of an overall effect on localization and on the speech tests, four subjects preferred to use adaptive frequency compression in the HA, another four preferred the standard HA program, whereas the remaining four listeners had no preference.

Our results extend previous studies in bimodal listeners, in which FC was limited to high frequencies above 1.5 kHz. These studies reported no benefit of FC for speech perception or sound localization either (Davidson et al., 2015; McDermott et al., 2010; Park et al., 2012; Perreau et al., 2013). Perreau et al. (2013) reported impaired vowel perception, in contrast to our results, even though we used FC across a much broader frequency range. This discrepancy most likely is a consequence of our signal-adaptive variant of FC, specifically aiming at preserving the vowels (Figure 2B).

ILD CUES

A possible explanation for the lack of a benefit in localization and speech understanding in noise is that our subjects were not able to use the frequency shifted ILDs. Francart et al. (2007) showed reasonable ILD thresholds for noise bands that differed up to one octave in center frequency across the ears in normal-hearing listeners, even without acclimatization. Note, however, that reasonable ILD thresholds do not necessarily imply a percept of an auditory object with a distinct location in space. Moreover, in our study, the frequency mismatches were even larger than one octave. Most subjects used a compression ratio of 4:1, meaning that an input of 1 kHz was presented to the cochlea as 250 Hz (Figure 1). Even the weakest compression ratio used in this study (1.9:1) mapped a 1 kHz input to 420 Hz (a mismatch of more than 1 octave).

Still, it is noteworthy that two subjects considerably improved their localization performance with adaptive FC, compared to the standard HA program (Figure 4). A third subject showed a small improvement in the localization slope for high-pass filtered noise sounds. All three listeners had a relatively low frequency compression ratio (1.9:1 and 2.3:1). The subjects provided with a compression ratio of 4:1 generally showed very low localization slopes (< 0.2) for both stimulus types and HA programs. The one subject with a clear decrement in localization slope for broadband sounds (0.8 without FC) used a compression ratio of 2.7:1. This could hint at the importance to limit interaural frequency mismatches in promoting potentially usable binaural difference cues. Alternatively, the subjects with low compression ratios all had better residual hearing and therefore possibly a better spectral and temporal resolution.

Apart from the frequency mismatch possibly preventing subjects to integrate the CI and the HA signal, a shorter processing delay in the CI device compared to the HA may have caused the strong localization bias observed in most subjects towards the CI side. This effect is evidenced by the large intercepts in Fig. 4 (Chalupper et al., 2014). Such a consistent timing difference would induce a strong precedence effect (Blauert, 1997), in which the leading ear fully dominates the localization percept. If true, such an effect can hardly ever be overcome by competing (unnatural) ILD cues, and additional changes should then be considered to match the potential processing delays as good as possible.

Unfortunately, the improvement in sound localization with adaptive FC, observed in two subjects, was not reflected in the other listening tasks. The bimodal benefit for speech understanding in noise for these subjects decreased on average from 2.9 dB with the standard HA program, to only 0.5 dB with adaptive FC in these subjects in contrast to the idea that improved ILD cues could help speech understanding in challenging listening situations. Perhaps, there might have been a trade-off with the distortions in the speech signal, due to the extreme FC applied in this study. No noteworthy differences were found in the results of these two subjects in the other tasks, but they did perform below average in the CI-only listening condition for speech understanding in quiet and in noise, which suggests that they relied more on their (normal) HA input, and possibly had adapted better to the bimodal listening situation. Besides the low compression ratio for these listeners, this could be an additional explanation for their better adaptation to FC. Still, the question remains why these subjects could not use their improved localization skills for speech understanding in noise.

We tested a signal-adaptive algorithm that was designed to apply extreme FC only to consonants and not to vowels. As a result, the binaural cues differed across input signals. This could provide a third reason for the lack of improvement in the binaural tasks: the brain may just be unable to adapt to such inconsistencies (Hofman et al., 1998), resulting in ineffective binaural cues.

Four out of thirteen subjects preferred to continue using extreme FC after the study. These four subjects also performed better than the other subjects in almost all tests, with and without adaptive FC in the HA. They reported that the adaptive FC resulted in a more fused sound percept between the CI and HA. Possibly these subjects had a better capacity to adapt to new auditory inputs, or they were more motivated participants, trying harder to acclimatize. An alternative explanation would be that these subjects were already acclimatized to gross pitch mismatches between their ears, as their high-frequency residual hearing was poorer than for the subjects who did not prefer adaptive FC (70 ± 25 versus 63 ± 18 dB HL).

The present study was based on the assumption that bimodal listeners have to rely on ILD cues for sound localization. A speculative explanation for the lack of an overall improvement in our results is that bimodal subjects may instead use temporal cues in the envelope of sounds (envelope ITDs), or yet unknown cues. A final explanation could be that our subjects had too limited residual hearing in the non-implanted ear. As a result, even with extreme FC the high-frequency information may not have been sufficiently audible to be used as binaural cues. It has been argued that a bimodal benefit for speech understanding is largely the result of a better audible fundamental frequency through the HA (Zhang et al., 2010). This could explain the similarity in performance for our speech and voice-pitch perception tasks, since adaptive FC was only applied to consonants, leaving voicing cues in vowels unaffected.

BIMODAL BENEFIT

Results of the voice-pitch perception task (both with and without adaptive FC) corresponded well with an earlier study in bimodal subjects (Straatman et al., 2010), which also reported F0 differences at thresholds of about 6 – 7 semitones. Also the bimodal benefit for speech understanding in noise was in line with the 1-3 (S0N0) and 1 (S0NHA) dB reported in the literature (Ching et al., 2007; van Hoesel, 2012). Still, the benefit in S0NHA was rather small compared to our previous study with largely the same subjects (1.6 versus 3.0 dB) (Veugen et al., 2016b). This can possibly be explained by the different sentence materials that we used. The fundamental frequency of the target speaker in the Matrix test was 175 Hz, and in our previous study we used the Leuven Intelligibility Sentence Test (van Wieringen et al., 2008) with sentences that had an F0 of 200 Hz. Since the single-talker babble noise used in both studies had an F0 of 127 Hz, target and noise may have been more difficult to separate in the present study. Maybe this has reduced the potential benefit of the complementary low-frequency information provided by the HA that helps to segregate different voices.

CONCLUSION

We aimed to improve ILD cues and preserve speech perception by an experimental adaptive frequency compression algorithm in the HA of thirteen bimodal listeners. Frequency compression was applied above 160 Hz, but only to consonant signals, therefore preserving vowel formants. Adaptive frequency compression did not significantly affect bimodal localization or listening performance, most likely because (1) subjects could not use frequency compressed ILDs, or (2) the adaptive switching between vowels and consonants prevented adaptation to consistent ILD cues.

CHAPTER 6

REACTION TIMES TO MONAURAL AND BINAURAL SPECTROTEMPORAL MODULATIONS: NORMAL-HEARING AND SIMULATED IMPAIRED-HEARING

Manuscript in preparation

Veugen L.C., van Opstal A.J., Louvet D., van Wanrooij M.M.

ABSTRACT

We tested whether joint spectrotemporal sensitivity follows from spectrotemporal separability for normal-hearing listeners, both for normal-hearing conditions and for impaired-hearing simulations. In a manual reaction-time task, normal-hearing listeners had to detect the onset of a dynamic ripple that encompassed all combinations of 8 spectral ([0-8] cycles/octave) and 17 temporal modulations (±[0-64] Hz) at a fixed modulation depth of 50%.

Spectral and temporal modulations elicited band-pass filtered sensitivity characteristics, with fastest detection rates around 1 cyc/oct and 32 Hz for normal-hearing conditions. These results closely resemble data from other studies that typically used the modulation depth detection threshold as a sensitivity measure for spectral-temporal modulations. Binaural performance was always slower than predicted by the benchmark race model of statistical facilitation of independent monaural channels, suggesting binaural integration.

To simulate hearing-impairment, stimuli were processed with a 6-channel cochlear implant (CI) vocoder, and a hearing aid (HA) simulation that introduced spectral smearing and low-pass filtering. Reaction times were always much slower compared to normal hearing, especially for the highest spectral densities.

Although singular-value decomposition indicated that the joint spectrotemporal sensitivity matrix could be largely reconstructed from independent temporal and spectral sensitivity functions, in line with time-spectral separability, we obtained a significant inseparable spectral-temporal contribution to the responses for all hearing conditions.

These results affirm that the reaction-time task yields a solid and effective objective measure of acoustic ripple sensitivity, which may also be applicable to hearing-impaired patients.

INTRODUCTION

Human speech and other complex sounds in the natural environment are typically dynamic signals that rapidly change in amplitude over both time and frequency (Elliott et al., 2009). Fluctuations in the temporal domain provide information about the rhythm of speech, such as syllable and word boundaries, whereas variations in the spectral domain are essential for formant and voice pitch perception (Liberman, 1997). Sensitivity to these joint spectral and temporal modulations is deemed crucial for the identification of complex sound features, and for speech comprehension.

Spectrotemporal dynamic ripples have been introduced in psychoacoustics to investigate the spectrotemporal modulation sensitivity of auditory perception. Ripples are broadband noise stimuli that are modulated sinusoidally in amplitude over time and/or frequency (Bernstein et al., 1987; Supin et al., 1994). Ripples are ideal to assess hearing performance as they represent, but are not recognizable as, naturalistic sounds. Sensitivity of the healthy human auditory system has been studied thoroughly with ripples and generally shows a band- or low-pass response to spectral and temporal modulations, reflecting the limits of auditory sensitivity at higher modulation rates (Chi et al., 1999; Viemeister, 1979). Speech understanding is thought to relate mostly to joint spectrotemporal sensitivity. Chi et al. (1999) reported that in normal-hearing subjects the modulation transfer function of combined spectrotemporal ripples is highly separable, as it can be well approximated by the product of a single temporal and spectral filter. Separability implies that the joint spectrotemporal sensitivity can be directly obtained from pure temporal and spectral sensitivity measurements.

In the present study we used manual reaction times to construct the spectrotemporal modulation transfer function (stMTF), rather than the conventionally used modulation detection or discrimination thresholds. Research in monkeys recently showed that reaction times systematically depend on the acoustic modulation rates (Massoudi et al., 2014). Several models have been proposed to explain the underlying process of response latency in reaction-time tasks (Ratcliff et al., 2004). It is commonly assumed that a decision signal rises with accumulating evidence of the stimulus, until a certain threshold is reached that triggers the response (Reddi et al., 2003). As such, reaction times are directly related to the difficulty of a task, and could thus provide more detailed information on the audibility of spectrotemporal ripples.

Furthermore, reaction times allow for testing the presence or absence of binaural integration on the basis of monaural responses, by comparing binaural reaction times against the prediction of a so-called 'race model'. In such a model, the signals from either ear compete independently to reach the detection threshold, so that the response latency is determined by the winner of an independent parallel race between the two ears (Raab, 1962). Due to statistical facilitation, this race to threshold leads to faster reaction times for binaural stimulation than for monaural stimuli, as the distribution of minimum monaural reaction times yields faster responses than those produced by either ear (Gielen et al., 1983a; Hershenson, 1962). However, when this so-called redundant stimulus effect differs from the race-model prediction, it could imply true integration in an underlying neural interaction process (Gielen et al., 1983a; Miller, 1982; Schroter et al., 2007).

We tested whether reaction times are a solid objective measure of auditory sensitivity to ripples with various spectrotemporal modulations for six normal-hearing listeners. We assessed (in)separability of joint spectrotemporal sensitivity, and investigated how binaural listening affected modulation sensitivity compared to monaural listening conditions by comparing the data with the race-model prediction. As a validation of our reaction-time paradigm, we also collected data under more challenging impaired-hearing simulations that are known to affect temporal and spectral sensitivity (Bacon et al., 1985; Golub et al., 2012; Henry et al., 2005; Moore et al., 2001).

METHODS

SUBJECTS

Six listeners participated in this study (3 male, ages 20-25 years), none of whom reported a history of auditory deficits. All subjects had normal hearing (< 20 dB HL) in both ears from 125 to 8000 Hz. Except for two of the authors, subjects were naïve to the purpose of the experiments. The study was approved by the Local Ethics Committee of the Radboud University Nijmegen (protocol number 40327.091.12).

APPARATUS

Subjects were seated in an acoustically shielded sound chamber. Stimuli were presented through TDH 39 headphones (Telephonics Corporation, Farmingdale, NY, USA). For sound processing and data acquisition we used Tucker Davis Technologies System 3 (Alachua, FL, USA). Stored sounds were sent via the PC to a real-time processor (RP2.1) at a sampling rate of 48828.125 Hz, and passed through a programmable attenuator (PA5). Stimuli were set at a comfortable, well-audible loudness of 65 dB(A) (calibrated using a KEMAR head calibration set, connected to a Brüel & Kjaer measuring amplifier type 2610 [Nærum, Denmark]).

STIMULI

Dynamic ripples were created in MATLAB [version R2012a; Mathworks Inc., Natick, MA, USA] as described by Depireux et al. (2001). The carrier of these stimuli consisted of a broadband spectrum of multiple harmonic tones, each described by:

$$c_i(t) = \sin(2\pi f_i t + \varphi_i) \tag{1}$$

where t is time (s), f_i frequency (Hz) of the i-th harmonic, and is its phase (rad). In this experiment, the broadband carrier consisted of 128 harmonic tones, equally spaced (20 tones/octave) over 6.4 octaves (250 Hz-20.4 kHz). All components had random phase except for the first (=0). The i-th frequency was determined by with f_0 = 250 Hz the lowest frequency, and i = 0 - 127.

The spectrotemporal envelope determined the ripple fluctuations in amplitude over time and/or frequency:

$$s_i(t, x_i) = 1 + \Delta M \cdot \sin(2\pi\omega t + 2\pi\,\Omega\,x_i) \tag{2}$$

with t is time (s), x_i is the position on the frequency axis (in octaves above the lowest frequency), ΔM is the modulation depth (here set to 0.5 for all components), ω is the ripple velocity (Hz) and Ω is ripple density (cycles/octave).

Together the carrier and the modulator formed the ripple in our experiments as follows:

$$R(t) = \begin{cases} t < t_{onset}, & \sum_i c_i(t) \\ t > t_{onset}, & \sum_i s_i(t - t_{onset}, x_i) \cdot c_i(t) \end{cases} \tag{3}$$

The modulated sounds were thus preceded by a non-modulated harmonic complex () with a randomized duration (t_{onset}) between 700 and 1200 ms with a step size of 100 ms. Moving ripples were presented with velocities of 0 Hz and ±[0.5; 1; 2; 4; 8; 16; 32; 64] Hz and densities [0; 0.25; 0.5; 0.75; 1; 2; 4; 8] cycles/octave, yielding a total of (17 velocities * 8 densities =) 136 different stimuli.

CI SIMULATION

CI vocoder simulations were created using a previously described method by Litvak et al. (Litvak et al., 2007) that models the Advanced Bionics Harmony CI processor. Briefly, the vocoder algorithm works as follows. After resampling the input signal to 17.4 kHz, the vocoder applied a high-pass pre-emphasis filter (cut-off at 1.5 kHz). Then, the signal was band-pass filtered by a short-time Fourier transform with 256 bins and 75% temporal overlap (192 bins). Bins were grouped into 6 non-overlapping logarithmically spaced channels (Fig. 1B; at center frequencies: 452, 715, 1132, 1792, 2836 and 4488 Hz). Random-phase noise bands with similar center frequencies were modulated with amplitudes equal to the square root of the total energy in the channel. The channels were summed and inverse short-time Fourier transformed to reproduce a temporal waveform for presentation to the listeners.

We used 6 vocoder channels to simulate hearing via a CI, as CI users are typically unable to effectively utilize information from all available CI channels (Henry et al., 2003). Normal-hearing subjects have shown similar performance as CI users for speech understanding scores in quiet with 4-6 channels (Loizou et al., 1999). This is in line with pilot experiments in our lab that revealed that five normal-hearing subjects achieved a performance level of ~80% in a consonant-vowel-consonant recognition test, when the words were vocoded with only 6 channels.

HA SIMULATION

HA simulations were generated by using a fourth-order Butterworth low-pass filter with a cut-off at 500 Hz, mimicking residual hearing in the lower frequencies present in the bimodal users of our previous study (Veugen et al., 2016b). Additionally, the loss of frequency selectivity (spectral smearing) was simulated as previously described by Baer and Moore (Baer et al., 1994). Asymmetrically broadened auditory filters were used with broadening factors 6 and 3 for the lower and upper branch respectively, as these are representative for moderate-to-severe hearing impairment (Glasberg et al., 1986).

The CI and HA simulated stimuli were normalized to the same root-mean-squared value as the original non-vocoded sounds. Figure 1 visualizes the effect of the CI and HA simulation on ripples.

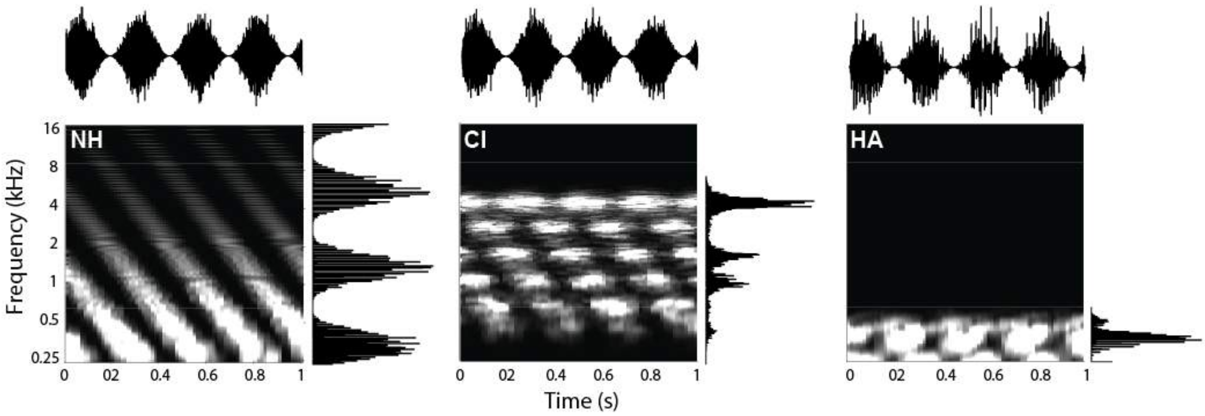

Figure 1. Moving ripple spectra. *Ripple with velocity 4 Hz and density 0.5 cycles/octave) for the normal-hearing condition (NH), cochlear implant simulation (CI) and hearing-aid simulation (HA). The signals on the top row represent the temporal waveforms for a purely amplitude-modulated sound (4 Hz, 0 c/o). The signal on the right of each figure visualizes a pure stationary spectral ripple modulation (0 Hz, 0.5 c/o). For clarity, the sound is shown after* t_{onset} *(at t=0).*

PARADIGM

Subjects were instructed to press a button as quickly as possible when they heard the sound change from static noise to modulated ripple. Modulated ripples lasted for 3000 ms, unless the button was pressed, in which case the sound was ended prematurely and the next trial was initiated after a brief (0.5 - 1 s) period of silence between each trial. If the button was pressed before ripple onset, the trial was reiterated, but no more than 4 times. The outcome measure of the experiment was the subject's manual reaction time, which was defined as the time between the onset of the ripple and the moment the button was pressed.

We tested five different listening conditions; acoustic stimuli were presented 1) monaurally (monaural NH), 2) monaurally via CI vocoder (unimodal CI), 3) monaurally via HA simulation (unimodal HA), 4) binaurally (binaural NH), 5) binaurally via bimodal CI and HA simulation (bimodal). In monaural conditions, both ears were tested separately. In the bimodal condition, CI and HA were tested in both the right and left ear in different sessions. We did not test the binaural unimodal listening conditions (CI-CI or HA-HA). Each stimulus was presented 5 times in each listening condition. A complete data set thus contained a total of 6120 stimuli, which were split in 12 sessions of 30-40 minutes, each containing 510 trials. Sessions were distributed over 6 days of two sessions each. Ripples and conditions were presented in pseudo-randomized order.

Because of time constraints, data collection was not fully completed in the four naïve subjects. Two naïve subjects completed 11 out of 12 sessions, with all ripples measured at least twice. The other two subjects completed 9 out of 12 sessions (20 and 23 ripples not measured, respectively; 83 and 141 ripples were presented only once in these subjects).

The four naïve subjects performed one training session under normal hearing conditions prior to the recording sessions, in order to become familiarized with the ripple stimuli and experimental procedures. We observed no systematic change in the average reaction times during the training session for these four subjects, or over the time course of the experimental sessions for all six subjects. This observation indicates that procedural learning effects did not confound the reaction-time data.

ANALYSIS

Data analysis was performed with custom-written MATLAB software. Reaction times generally show a skewed distribution with an extended tail towards longer reaction times. To obtain normally distributed data (Carpenter et al., 2009), the reaction-time data were transformed to their reciprocal (1/reaction time), referred to here as 'promptness' (1/s). This also allows the measurements to be more readily interpreted as sensitivity measures to the different spectrotemporal modulations, as a higher/lower promptness (as opposed to a shorter/longer reaction time) indicates a higher/lower sensitivity. Responses were pooled across subjects and ears for grand average analyses. Reaction times below 150 ms (clear anticipatory responses) were removed from the analysis. If a response was not initiated within 3 s (considered a sign for inattentiveness, or of an inability to detect the ripple), we set the response time (promptness) to 3 s (1/3 s^{-1}). Non-responses were found in 10% of the trials under NH conditions and in 46% of the trials of the hearing-impaired conditions (especially at the high spectral modulations). We do not explicitly take into account the percentage of non-responses, but note that in our analyses a higher number of non-responses would yield a median promptness (reaction time) closer to 1/3 s^{-1} (3 s). The non-modulated sound (velocity 0 Hz and density 0 cycles/octave) served as catch stimulus, in order to determine the false alarm (or guess) rate of the participant. The guess rate varied from 7% for binaural NH, to 21% for monaural HA listening, with the average guess rate across conditions at 12%.

6

SPECTROTEMPORAL TRANSFER FUNCTION

For each of the five listening conditions (monaural and binaural normal-hearing, and monaural and bimodal CI and HA simulation), we calculated the mean promptness per ripple to construct a two-dimensional spectrotemporal modulation transfer function as a joint function of ripple density, , and ripple velocity, . Similarly, we determined the temporal modulation transfer function tMTF, F(ω), and the spectral modulation transfer function sMTF, G(Ω), for the 0-density and 0-velocity stimuli, respectively.

SEPARABILITY

To analyze the degree of separability of the stMTF, we applied singular value decomposition (SVD) for all listening conditions. SVD transforms the stMTF into two unitary matrices containing temporal and spectral singular vectors, respectively, and a rectangular diagonal matrix that contains the singular values: stMTF(ω,Ω)=F(ω)·Σ·G(Ω). In case of a fully separable stMTF, the spectral and temporal components are independent of each other and the total of all 136 spectrotemporal responses can be expressed by the vectorial outer product of a single temporal $F_1(\omega)$ (17 components) and spectral $G_1(\Omega)$ (8 components) modulation transfer function, as follows:

$$stMTF_1(\omega, \Omega) = \sigma_1 \cdot F_1(\omega) \times G_1(\Omega) \tag{4}$$

with σ_1 the largest singular value. We calculated the separability index, which ranges between zero (totally inseparable) to one (fully separable), and is based on the relative dominance of the first singular value:

$$\alpha_{SVD} = \frac{\sigma_1^2}{\sum_i^n \sigma_i^2} \tag{5}$$

The separable stMTF estimate was reconstructed according to Eq. 4, for which we also determined Pearson's linear correlation coefficient with the original data. We also reconstructed the stMTF on the basis of the first two singular values, according to

$$stMTF_{12}(\omega, \Omega) = \sigma_1 \cdot F_1(\omega) \times G_1(\Omega) + \sigma_2 \cdot F_2(\omega) \times G_2(\Omega) \tag{6}$$

and tested whether this second-order estimate was significantly better than the separable, first-order estimate, by comparing the correlation coefficients with the original data.

RACE MODEL

We compared the observed reaction times for binaural hearing with the quantitative predictions of performance on the basis of the monaural reaction times, using the race model of statistical facilitation. This model assumes independence of the two monaural processes (Gielen et al., 1983a; Raab, 1962). Any violation to the race model suggests neural interactions when processing the input from both ears:

$$P(\tau_{race} \leq t) = P(\tau_{M1} \leq t) + P(\tau_{M2} \leq t) - P(\tau_{M1} \leq t)\times P(\tau_{M2} \leq t) \qquad (7)$$

with P($\boldsymbol{\tau} \leq t$) the cumulative probability function (CDF) of an observed reaction time g at time t; M1 and M2 represent monaurally-presented stimuli (NH, CI and HA). We estimated the cumulative probability density functions (CDF) from the promptness values. The race model CDF was constructed from the two monaural CDFs according to Eqn. 7. For comparative purposes, we calculated the difference in the medians (at the 50% cumulative probability level) between actual performance and race-model predictions. Ripples for which fewer than 10 responses were collected were discarded from this analysis because no reliable CDF could be constructed. Non-responses (reaction times > 2500 ms) were also discarded from the race-model analysis.

STATISTICS

Data were always reported as mean values ± 1 standard deviation. We also calculated 95% confidence intervals of promptness for the pure temporal and spectral ripples. As a criterion of significance for a statistical difference we took $p<0.05$.

RESULTS

REACTION-TIME TASK

To illustrate the systematic dependence of the reaction times on the acoustic conditions, Figure 2 shows the reaction-time data for one listener (P4). Each panel shows a different situation: Figs. 2A and 2C show the binaural hearing condition for pure amplitude-modulated noises (Fig. 2A; density zero, three different velocities) and for three different spectral modulations (Fig. 2C; fixed velocity at 1 Hz, at three different densities), whereas Fig. 2B shows the reaction times for the same ripple (1Hz amplitude modulation) for three different hearing conditions: binaural, bimodal, and HA-only. The 1Hz, 0 cyc/oct data are common in each panel for comparison. The data are plotted in so-called recip-robit format, in which the ordinate represents the promptness (inverted scale), while the abscissa is plotted on probit scale. In this format, the data can be described by a straight line when the promptness responses form a normal distribution. The examples show that for nearly all conditions this was indeed the case. One may also note that the slopes of the lines, and the median of the distributions (promptness at 50% cumulative probability) systematically change with the stimulus parameters and hearing condition. An interesting case is shown in Fig. 2C for the 8 cyc/oct ripple (yellow dots). This ripple is poorly detectable for the subject, which immediately shows as a strong deviation of the slope towards short reaction times (mainly guesses), and many responses that are found at the maximum response time of 3 s.

6

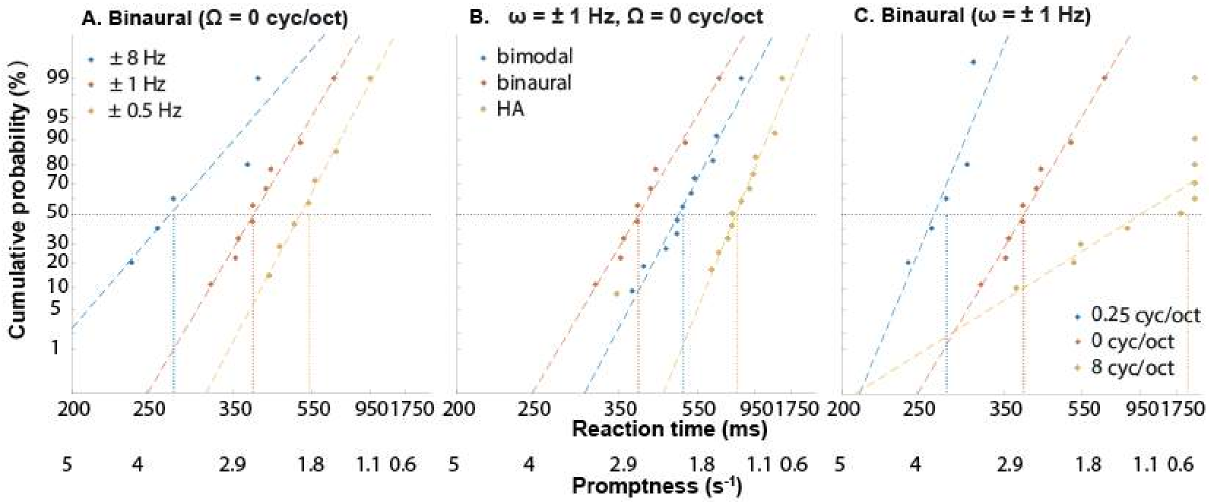

Figure 2. Example reciprobit plots for participant P4. A) *Reaction times to amplitude modulations (i.e. density = 0 cyc/oct) of frequencies 0.5, 1 and 8 Hz (yellow, red, and blue, respectively) in a binaural listening condition are plotted in reciprobit format. Note that most points lie along a straight line; best-fit regression lines are indicated by dashed lines. The median promptness is indicated by the vertical dotted lines between 0 and 50%.* ***B)*** *Reciprobit plots of reaction times to an amplitude modulation of 1 Hz for the binaural, bimodal and hearing-aid listening conditions (red, blue and yellow, respectively).* ***C)*** *Reciprobit plots of reaction times to spectrotemporal modulations with densities 0, 0.25 and 8 cyc/oct (blue, red, and yellow, respectively) for the binaural listening condition. Note that the binaural [1 Hz, 0 c/o] condition is shown in each panel (red).*

TEMPORAL- AND SPECTRAL-ONLY MODULATIONS

NORMAL HEARING

The promptness data exhibited predominantly a high-pass characteristic in the temporal dimension under both binaural and monaural normal-hearing conditions (Fig. 3A, black and red curves, respectively) for the pure temporal amplitude modulations (density = 0 cycles/octave). Responses were faster (higher promptness) for larger absolute velocities, until around ±64 Hz, where the relationship between ripple velocity and promptness saturated or started to decline. The fastest responses, with an average promptness of 3.3 (monaural) and 3.5 (binaural) s^{-1}, were observed at ±32 Hz. Responses to upward (< 0 Hz) and downward (> 0 Hz) ripples were very similar: correlation coefficients between the responses to up- and downward ripples were 0.97 (monaural) and 0.95 (binaural), respectively. Binaural responses (Fig. 3A, black triangles) were on average about 20 ms faster than monaural responses (Fig. 3A, red dots). Differences in response times and speeds will be compared in more detail below with the race model predictions.

For the static ripples (purely spectral modulations at velocity = 0 Hz), the promptness data can be characterized by a low-pass filter for both the binaural and monaural normal-hearing conditions (Fig. 3B, black triangles and red dots, respectively). A small peak in promptness is observed between 0.75 - 2 cycles/octave, while detection is very poor (low promptness) for the highest density of 8 cycles/octave. This property presumably reflects the limits of resolvable power of the human auditory filters, leading to a poorer discrimination of spectral patterns with finer spectral detail.

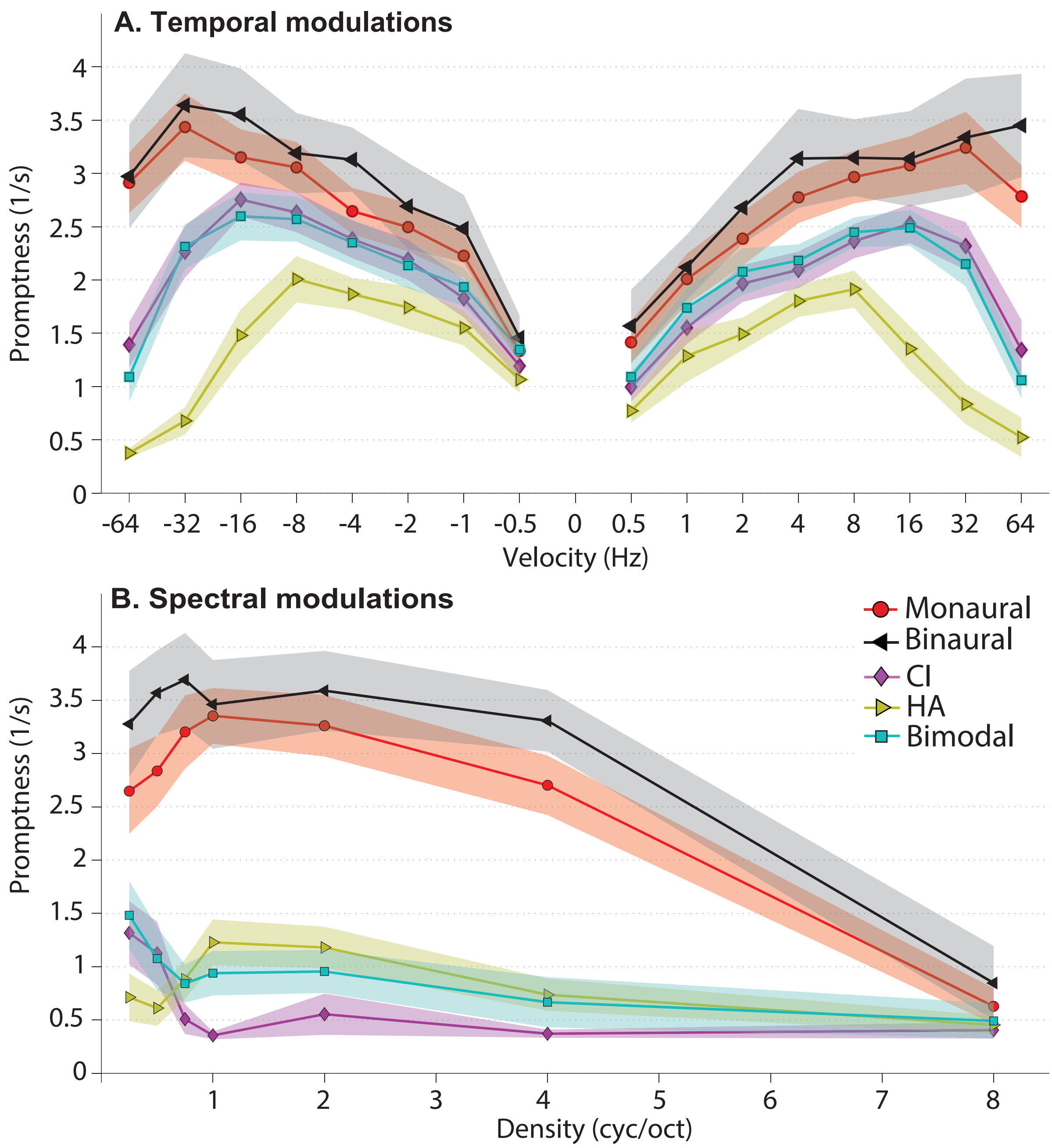

Figure 3. Pure temporal (A) and spectral (B) modulation transfer functions. A) *Mean promptness (line and markers) ± 95% confidence interval (patch) as a function of velocity (density = 0 c/o) and* ***B)*** *as a function of density (velocity = 0 Hz), for the normal-hearing monaural and binaural conditions, and the impaired-hearing CI, HA and bimodal simulations.*

IMPAIRED-HEARING SIMULATIONS

Responses to pure amplitude modulations under bimodal and monaural impaired-hearing conditions were on average 182 ms slower than under normal-hearing conditions, clearly implicating increased difficulty in temporal-modulation detection (Figure 3A, blue squares, purple diamonds and green triangles for bimodal and monaural CI and HA simulations, respectively). Furthermore, the decrease of promptness systematically varied across velocities such that the CI and HA simulation data still resembled a band-pass pattern. Response promptness peaked at 2.5 s^{-1} for velocities of ±16 Hz for the bimodal and CI conditions, and at 2 s^{-1} around ±8 Hz for the HA condition. The fastest responses were obtained for the bimodal hearing condition; they were on average about 30 ms faster than for the CI-only condition. Correlation coefficients between the responses to up- and downward ripples were 0.97-0.99 for all three conditions.

The static ripples (spectral modulations only) elicited fewer responses within the response time window of 3 s in the CI, HA and bimodal conditions, especially for the higher ripple densities. Responses made for the CI simulation resembled a low-pass filter characteristic with a cutoff around 0.75 c/o, although only 83 responses were made (in total for all subjects) within 2500 ms of ripple onset for ripple densities above 0.75 cycles/octave. Responses in the HA condition followed a band-pass characteristic with its highest power around 1-2 cycles/octave. Bimodal responses fell in between those of the CI and HA condition. For the HA and bimodal condition, we obtained 12-40 responses < 2500 ms for all densities, except for 8 c/o (8 responses).

JOINT SPECTROTEMPORAL MODULATIONS

Figure 4A-E shows the spectro-temporal modulation transfer functions for the two NH conditions, and for the three impaired-hearing simulations, for all joint spectrotemporal ripples, as mean promptness (averaged across listeners) per ripple density (abscissa) and velocity (ordinate).

Deep yellow colours correspond to high spectral-temporal sensitivity, blue colours to low sensitivity (low promptness values). The results for pure AM stimuli (Fig. 3A) are encountered at the bottom horizontal line of the stMTF matrix, at Ω=0 c/o; the results for pure spectral modulations (Fig. 3B) are found along the central vertical line, at ω=0 Hz.

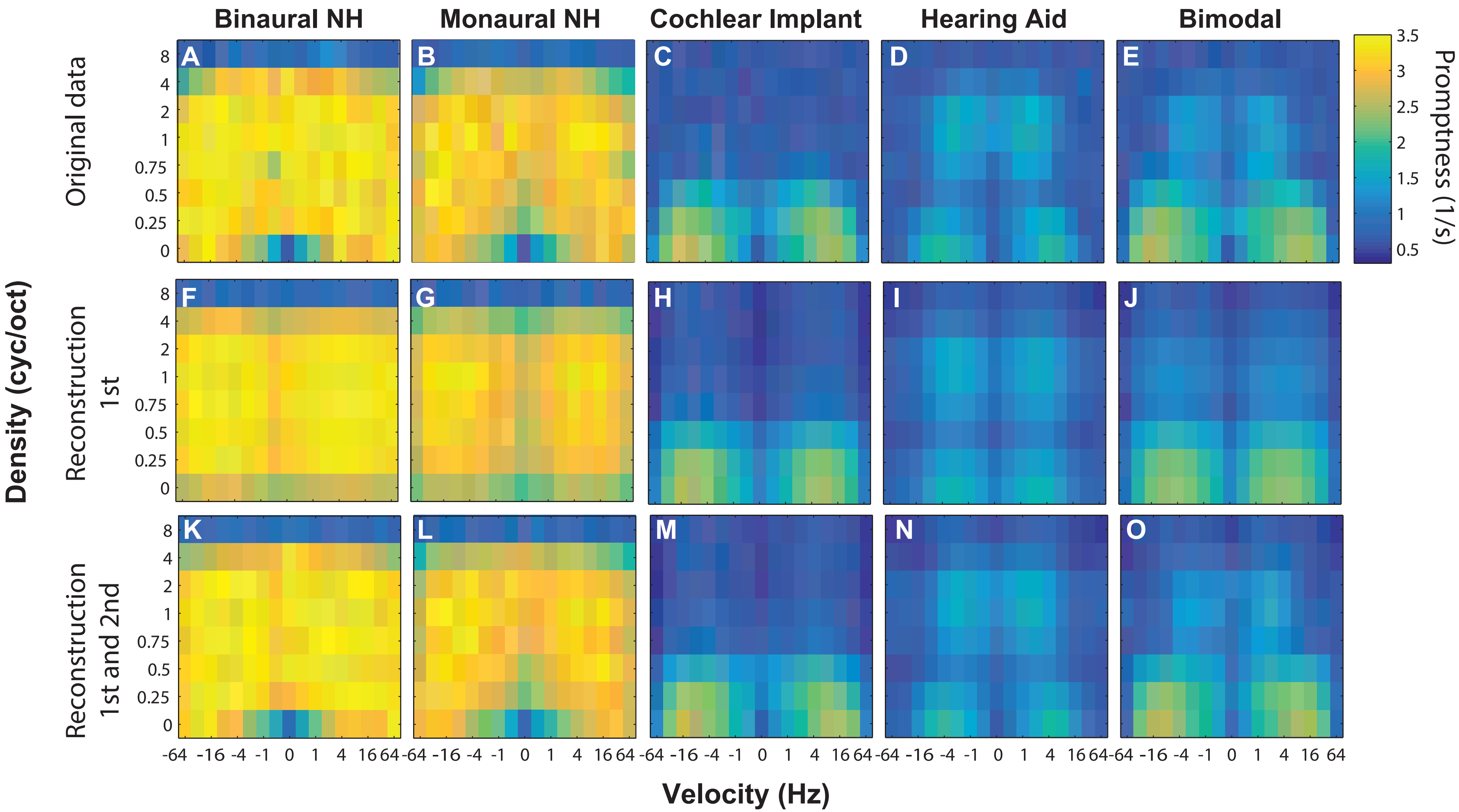

Figure 4. *Mean promptness (1/reaction time) as a function of velocity and density, representing spectro-temporal sensitivity in the normal hearing (NH) and simulated hearing-impaired conditions. The higher promptness, the faster ripple detection.* ***F-J)*** *Reconstructions of the spectrotemporal modulation transfer functions using only the first singular value after singular value decomposition (SVD; Eq. 4).* ***K-O)*** *Reconstructions of the spectrotemporal modulation transfer functions using the first and second SVD components (Eq. 6).*

6

NORMAL HEARING

It is immediately clear that the binaural NH condition (Fig. 4A) led to the fastest responses for all ripples, followed by the monaural NH condition (Fig. 4B; see also below). Whereas responses to pure temporal modulations depended on the ripple's velocity (Fig. 3A), this effect vanished for combined spectrotemporal modulations, resulting in faster responses for low velocities.

Spectro-temporal modulation sensitivity showed a band-pass filter characteristic for both density and velocity, with fastest detection rates around ripple velocities of 16-32 Hz and densities around 1 cycle/octave. Ripples were well detectable up to 4 cycles/octave, where promptness started to decrease when combined with high temporal modulation rates.

IMPAIRED-HEARING SIMULATIONS

The spectro-temporal transfer functions for the CI, HA and bimodal conditions (Figure 4C-E) were distinctly slower (more blue) when compared to normal hearing. Temporal modulation sensitivity again showed a band-pass filter characteristic with fastest detection rates around 16 Hz. CI simulations mainly affected the detection of spectral modulations, which is consistent with the modus operandi of a CI, whereby its band-pass filtering mechanism reduces spectral modulation sensitivity. Ripple detection with the CI became impossible for densities exceeding 0.75 cycles/octave. The monaural HA simulation resulted in the slowest responses of all hearing conditions. However, in contrast to the CI, higher densities of up to 4 cycles/octave could still be detected with the HA, albeit not when combined with fast temporal modulation rates like under normal hearing conditions. The result of bimodal stimulation resembled a conjunction of the CI and HA simulation results, seemingly exhibiting a 'best of both worlds' principle (Corneil et al., 2002) with reaction times as fast as for the best unimodal condition (Figure 4E). For high spectral modulation frequencies, the bimodal condition was comparable to the HA condition; for low spectral modulations it followed the CI-only condition.

SEPARABILITY

We assessed the degree of separability of the stMTF into a pure temporal and spectral component through singular value decomposition (SVD) using the separability index α_{SVD} (Eq. 5) and the correlation coefficient between the original data and reconstructions from the first (Eq. 4) as well as from the first two singular vectors (Eq. 6). If the two reconstructions result to be statistically indistinguishable from the original data, the MTF is considered to be separable.

NORMAL HEARING

The central row of Figure 4 shows the reconstructed $stMTF_1$ for the different hearing conditions. Note that the first-order reconstruction of the stMTF yields purely orthogonal patterns in the matrix, resulting from the full separability assumption. For both normal hearing conditions the α_{SVD} was 0.99, which suggests a high degree of separability. Similar results have been reported for other response methods (e.g. Chi et al., 1999). In line with this, the correlations between the original data and reconstructed stMTFs using the first SVD components were r_1=0.93 for the monaural, and r_1=0.94 for the binaural data, indicating that the reconstructions closely resembled the original data for both conditions (panels 4F-G). The mean squared errors between the original and reconstructed data were 7% and 9%.

However, by adding the second singular value with its spectral and temporal components (bottom row of Figure 4), the stMTF reconstructions yielded significantly ($p<0.005$) higher correlation coefficients: r_{12}=0.98 (monaural) and 0.99 (binaural), respectively, with the mean-squared errors reduced to 2% and 4% (panels 4K-L). The size of the second singular value was about 10% of the first component, which indicates a significant inseparable spectrotemporal contribution to hearing. The third and fourth singular components amounted to only 4% and 3% of the first, respectively, which did not further improve the stMTF estimates ($p>0.05$).

IMPAIRED-HEARING SIMULATIONS

For the hearing-impaired conditions, the separability index was best for the CI simulation (α_{SVD} = 0.98) and comparable, but worse, for the HA and bimodal conditions (α_{SVD} = 0.96). Correlation coefficients between the reconstructions and original data ranged from r_1=0.91 (HA), 0.93 (bimodal) to 0.98 (CI), respectively (Figure 4H-J). The mean squared errors ranged between 2-5%. stMTF reconstructions using the first two SVD components resulted in significantly higher correlation coefficients for all three hearing conditions: r_{12}=0.99 (CI and bimodal; $p<0.01$) and 0.98 (HA) with mean-squared errors of only 1% for all three conditions (Figure 4M-O). The second singular value was as high as 11% (CI) or even 18% (HA and bimodal) of the first singular value (the third and fourth were 6-7% and 4-5% of the first), again suggesting a considerable inseparable spectrotemporal component to the responses

RACE MODELS

To investigate to what extent monaural reaction times can predict binaural performance, we used the race model of statistical facilitation, which postulates independence between ears. As an example, Figure 5 displays the cumulative probability for a typical ripple (8Hz, 0.25 c/o), for the monaural and binaural NH conditions, as well as for the promptness that would be reached based on the race model of statistical facilitation. For this ripple, binaural performance was faster than monaural promptness, but it did not reach the level of statistical facilitation.

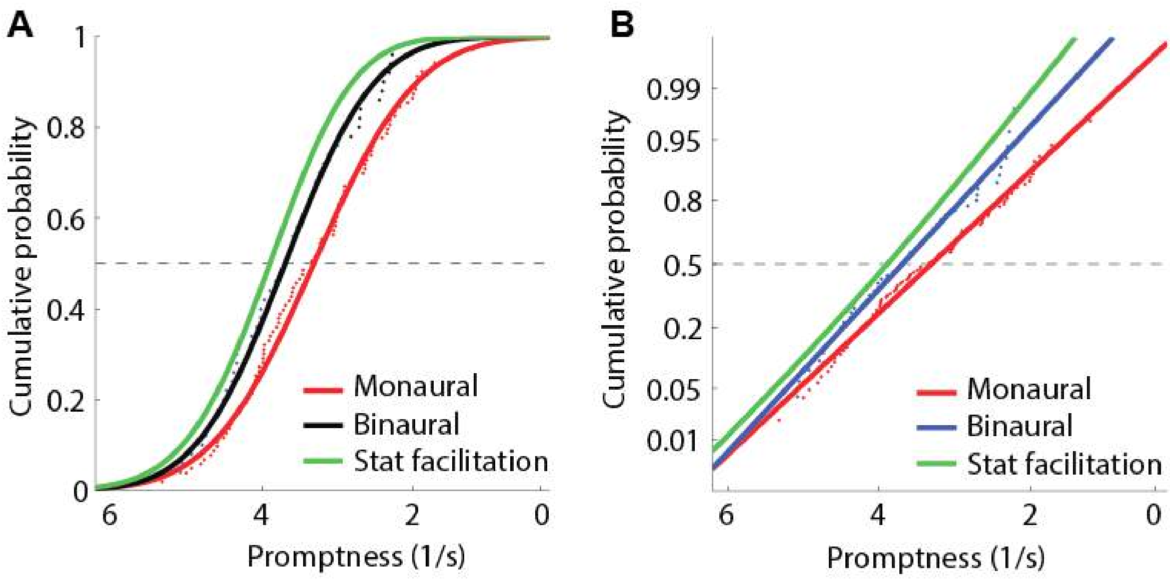

***Figure 5 A-B**. Cumulative distribution functions fitted through the promptness data for a typical ripple (8 Hz, 0.25 c/o) in the monaural (red) and binaural (black) listening conditions, and for the race model of statistical facilitation (green curve; Eqn. 7),* **B)** *on a probit scale). The horizontal dotted line is drawn at a cumulative probability of 0.5 (the mode of the distribution), used to calculate the difference between actual performance and race model predictions. For convenience, the horizontal axis was mirrored.*

Likewise, we compared the predictions from the race model (at the level of the median) for promptness to the median promptness of actually measured binaural or bimodal performance for all ripples (Figure 6). We also plotted binaural performance versus the fastest median monaural condition in this figure (red dots).

NORMAL HEARING

Overall, binaural performance was on average 32 ± 59 ms slower (dots above the diagonal in Figure 6A) than the benchmark race model prediction, but faster (on average 39 ± 71 ms) than the fastest monaural condition for the majority of ripples. These results show that, overall, binaural normal hearing performance fell between the fastest monaural condition and statistical facilitation (Eqn. 7).

IMPAIRED-HEARING SIMULATIONS

Bimodal hearing was also slower than predicted by the benchmark race model of statistical facilitation. Compared to the fastest monaural condition, we obtained a varying pattern of faster and slower bimodal responses (data around the diagonal in Figure 6B). Bimodal responses were sometimes slower than the fastest monaural responses, especially for higher ripple densities (>0.75 cycles/octave), for which the HA typically was the best monaural condition. For the lower modulation rates, where the CI dominated, bimodal performance seemed to fall between the race model of statistical facilitation and the fastest monaural response, like in the normal-hearing binaural condition. Taken together, these results suggest the existence of neural integration of the information that is available at both ears before eliciting a response.

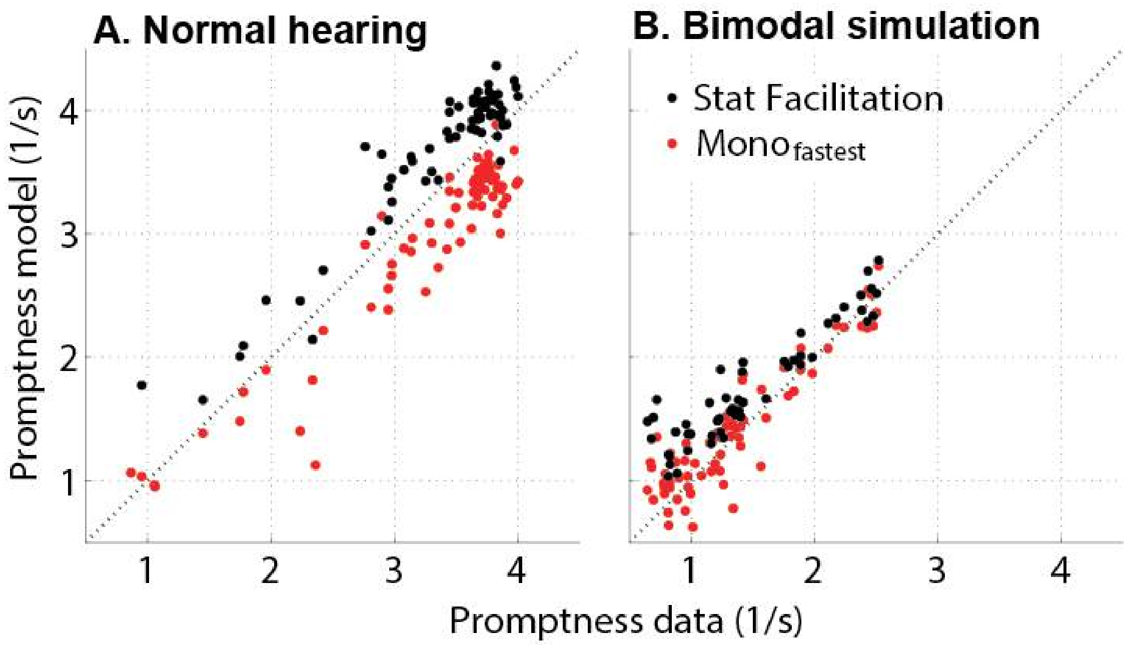

Figure 6. *Median recorded promptness values for all ripples in the binaural condition versus the median of the fastest monaural ear and versus the promptness at 50% cumulative probability of the race model of statistical facilitation, for* ***A)*** *the normal hearing condition and* ***B)*** *the hearing-impaired simulations.*

DISCUSSION

This study is the first to use reaction times to determine the auditory spectrotemporal modulation transfer function in human listeners. Reaction times appeared to be a solid and effective objective measure for ripple sensitivity, given its systematic relationship with the parameters that determine both temporal and spectral modulation rates. Sensitivity was highest for ripples with modulations around 16 Hz and 1 cycle/octave, and decreased for higher and lower modulation rates. Using simulations of a CI, HA and bimodal restorative hearing, spectrotemporal sensitivity worsened (reaction times increased) compared to normal hearing, in line with the impaired signal processing of the simulations. Although the separability of the stMTF into a spectral and temporal component was high for both the normal-hearing data, and for the simulated impaired-hearing responses, the responses also suggested a significant inseparable second-order spectrotemporal component, with a value between 10-20% of the first component, for all hearing conditions.

For a majority of ripples, binaural and bimodal reaction-time performance was slower than the prediction of the benchmark race model of statistical facilitation, suggesting binaural integration of monaural signals.

Constructing the stMTF based on reaction times is a fairly new approach that has been introduced so far only in monkey research (Massoudi et al., 2013; Massoudi et al., 2014). In those studies, responses to temporal ripples systematically depended on the ripple's period, and we therefore analyzed the pure temporal and spectral modulation sensitivities separately (Figure 3A). All other ripples contained spectral modulations, and therefore had a more abrupt change from stationary noise to the modulated ripple in our stimuli. Like in conventional modulation detection threshold procedures, responses were then driven by the subject's auditory spectrotemporal sensitivity, rather than by the speed of the temporal modulation (Figure 3B). An advantage over a modulation detection threshold paradigm is that stimuli can be presented at supra-threshold levels, allowing for a relatively easy task, suitable for clinical assessments, and for studies with children, or with experimental animals. This is supported by the fact that we did not observe procedural learning effects during the course of the experiments.

NORMAL HEARING

Our stMTFs (Figure 3 and Figure 4A-E) correspond well with the results of other studies, which usually measured ripple modulation detection thresholds. Chi et al. (1999) measured the full stMTF for normal-hearing subjects using an adaptive modulation detection threshold paradigm, and found band-pass functions for both the spectral and temporal dimensions. They found best ripple detection-thresholds at spectral modulations of 1 cycle/octave, and temporal modulations around 4-8 Hz. The spectrotemporal sensitivity decreased for both lower, but especially at higher modulation rates. Two other studies found best spectral ripple detection thresholds at slightly higher spectral modulations, i.e., around 3 cycles/octave (Anderson et al., 2012; Eddins et al., 2007). Also the maximum temporal modulation sensitivity has been reported to reach 16 Hz in the work of Bacon and Viemeister (1985). Despite these small quantitative differences between studies, the general patterns were similar, and in line with our results, suggesting that reaction times are indeed a solid and quick objective measure to determine the spectrotemporal sensitivity of (naïve) listeners.

In Figure 7 we compare our promptness data with the modulation thresholds collected in the study of Chi et al. (1999), clearly visualizing the similar band-pass relationships. Note that the axes are reversed in order, as best responses correspond to high promptness values, but low modulation indices. Chi et al. (1999) proposed a computational model to explain their data, in which the spectrotemporal modulation sensitivity is based on cortical responses to the ripple's spectrogram. The modulation transfer functions generated by their model closely resembled their data, and thus will resemble our data as well. They concluded that 'the upper limits of the spectral and temporal modulation rates are related through the effective bandwidths of the cochlear filters'.

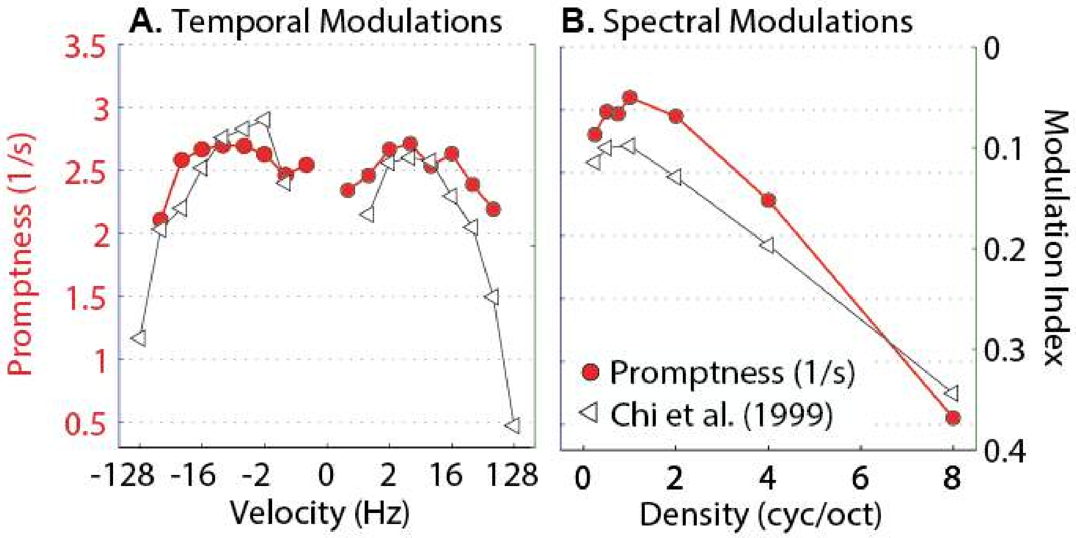

Figure 7. *Comparison between the one-dimensional MTFs obtained from the first component in the singular value decomposition procedure in our NH reaction-time study (transformed to promptness, left axis, red) and from the study of Chi et. al (1999) using a threshold searching paradigm (right axis, black).* ***A)*** *Temporal modulations).* ***B)*** *Spectral modulations.*

IMPAIRED-HEARING SIMULATIONS

To evaluate our reaction-time test under more challenging listening situations, we manipulated the ripple stimuli using HA and CI simulations. Both simulations made it substantially harder to detect ripple modulations and even impossible for certain parameters, eliciting much longer reaction times compared to monaural NH for all ripples. Bimodal hearing exhibited a 'best-of-both-worlds' effect, following the fastest unimodal condition, which was the CI for spectral modulations below 0.75 c/o, and the HA for higher spectral modulations (Fig. 4). An improvement in spectral ripple discrimination for bimodal hearing over the CI alone has also been found in users with combined electro-acoustic stimulation in the same ear (Golub et al., 2012).

Other studies have shown a 5-10 dB reduction in the temporal modulation detection threshold for CI users compared to normal hearing (Bacon et al., 1985; Golub et al., 2012; Won et al., 2011), whereas we found a decrease in promptness of 0.4 ± 0.2 1/s for the well-detectable rates below 16 Hz. In these studies, hearing impaired subjects performed in between normal hearing and CI users. Our HA simulation, however, showed longer reaction times for temporal modulations (at 0 cycles/octave) compared to the CI. It should be emphasized that our HA condition was based on a worst-case-scenario, simulating very little residual hearing, whereas hearing thresholds of the hearing impaired subjects in the study of Bacon and Viemeister (1985) still reached up to 10-20 dB HL at 1 kHz. Their study also showed a link between degraded temporal sensitivity and reduced listening bandwidth.

Impaired spectral modulation sensitivity with a CI is a likely result of its band-pass filtering mechanism that limits the spectral information to a set number of spectral bands. Henry et al. (2005) and Berenstein et al. (2008) both found lower spectral ripple modulation thresholds for CI users compared to NH listeners, roughly corresponding to the increased reaction times in our study. Spectral modulation thresholds in hearing-impaired subjects have been reported to be 5-10 dB worse than for normal hearing (Davies-Venn et al., 2015; Summers et al., 1994), which may be in agreement with the longer reaction times of our HA simulation compared to NH. Only a few studies investigated combined spectro-temporal modulation detection thresholds in hearing-impaired listeners, which were often worse compared to normal hearing listeners, especially for low temporal modulation rates (Bernstein et al., 2013; Mehraei et al., 2014).

RACE MODEL

To get insights in the mechanism of combining input at both ears, we used race models to test whether monaural responses could predict binaural performance. For NH conditions, faster reaction times were elicited when stimuli were presented binaurally compared to monaural presentation. Binaural responses fell in between the fastest monaural response and the race model of statistical facilitation (Figure 6A). This suggests that ripple detection was not determined by a parallel race between the two ears, but results, at least partially, from neural integration that causes additional time delays. This apparent slowing of reaction times under bimodal hearing could thus correspond to a *decreased* ripple sensitivity.

As NH responses were near ceiling performance (fastest possible reaction times, as stimuli were readily detectable), and given that in our simulations there was little spectral overlap between the HA and CI sides, a lack of positive effects in our data on bilateral or bimodal integration, by exceeding race-model performance, may perhaps not be too surprising.

It may be interesting to compare these findings to audiovisual gaze-orienting experiments that aim to study neural integration of visual and auditory signals. Strongest benefits of multisensory interactions (i.e., increased speed, accuracy, and precision of responses) are obtained for stimuli that overlap both in space and time, and thus provide multisensory evidence for a single object. Moreover, these interactions are strongest when the uni-sensory evoked responses are variable and slow (i.e., away from ceiling performance). This phenomenon is known as the ‘principle of inverse effectiveness’ (e.g., Stein and Meredith, 1993; Corneil et al. 2002; Van Wanrooij et al., 2009).

We here propose that beneficial effects of bimodal (CI-HA) integration will depend on whether the auditory system has sufficient evidence that left vs. right acoustic inputs arose from the same auditory object, rather than from unrelated sounds. The strongest bimodal benefits (i.e., enhanced sensitivity) will thus be found: (i) when spectral ranges of CI and HA overlap sufficiently (for within-spectral comparisons), and (ii) when monaural reaction-time distributions have sufficient variability, and overlap considerably.

A further remarkable finding was that bimodal performance for ripples with high modulation rates was even slower than the fastest monaural condition, suggesting even stronger inhibitory interactions between CI and HA inputs, but a mechanism for such an interaction cannot be proposed on the basis of the current data. Possibly, these results follow from hearing-impaired simulations with otherwise normal-hearing subjects. The question remains how long-term hearing impairment and adaptation to CI/HA inputs affect bimodal responses and spectrotemporal sensitivity. In a follow-up study, we report on the ripple-evoked responses of bimodal patients, and investigate the potential relationship between ripple-detection sensitivity and more relevant auditory tasks, like speech perception.

In contrast to our results, several studies have shown reaction times to stimuli that exceeded the predictions based on statistical facilitation. However, these studies typically involved responses to multisensory stimuli, or to the dichotic presentation of two spectrally distinct sounds (Gielen et al., 1983a; Miller, 1982; Schroter et al., 2007; Townsend et al., 1995). Similar to the findings of Schroter et al. (2007) for auditory stimuli that fused into a single percept, we did not obtain faster responses than expected from statistical facilitation in the bimodal conditions. The finding that both binaural and bimodal hearing were slower than the race model of statistical facilitation in our study suggests that binaural fusion of the monaural signals occurs for ripple detection, but it is a costly, time-consuming process. The benefits of an integrative process have to be found outside this study, and likely include the ability to localize sounds and enhancement of speech perception in noisy environments.

SEPARABILITY

Measuring the stMTF is typically a time-consuming process, for which it would be valuable to know whether the two-dimensional function is simply the product of a temporal and spectral component. Separability of auditory ripple sensitivity has been suggested for NH subjects (Chi et al., 1999). In contrast, Bernstein et al. (2013) found an interaction effect between spectral and temporal modulations, suggesting non-separability for normal hearing and impaired-hearing listeners. Our data support the notion of non-complete separability, as the contribution of the second singular value was close to 10% in the NH condition, and could be as large as 20% for bimodal impaired hearing. Mean-squared errors between the original and reconstructed data for NH ranged in our study from 2-9 %, which is slightly higher than the 3.4% found by Chi et. al (1999), who strongly argued for 'full separability'. In our study, the correlation coefficients between the measured data and the reconstructions from the resulting SVD components were on average across listening conditions 0.94 ± 0.003 for the first SVD component, and 0.99 ± 0.05 when using the first two SVD components. These differences were highly significant.

As can be appreciated from the observable differences between the single and two-component reconstructions in Figure 4, the second component adds a diagonal interaction component to the two-dimensional MTF. Most of this interaction yields changes to the MTF at 0-density modulations: reactions become slower for low-velocity modulations, and faster for higher velocities. Interestingly, the normal-hearing one-dimensional temporal MTF derived from the 2nd component is nearly a linearly increasing function of velocity (not shown). This suggests that the 2nd component reflects an amplitude modulation detection rate that can operate faster, simply because the stimulus modulation period is shorter (in line with the linear dependence of monkey reaction times on the amplitude-modulation period, found by Massoudi et al. 2014), and that is independent of how well a modulation frequency can be detected by the auditory system as reflected by the first-component MTF.

Taken together, we conclude that a full assessment of spectrotemporal performance requires testing the complete spectrotemporal field of all ripples, by including the second SVD component (Figure 4).

CONCLUSION

Reaction times are a solid objective measure for ripple sensitivity. The joint spectro-temporal transfer function closely resembled data from earlier studies that used modulation detection thresholds. Responses to spectrotemporal modulated ripples could be reconstructed by using the first two components of singular value decomposition, suggesting significant spectrotemporal inseparability. We further found that binaural and simulated bimodal reaction times could not be predicted from the monaural performance data, suggesting neural integration of binaural information.

CHAPTER 7

SPECTROTEMPORAL MODULATION SENSITIVITY OF BIMODAL COCHLEAR IMPLANT USERS

Manuscript in preparation

Veugen, L.C., van Opstal, A.J., Louvet D., van Wanrooij M.M.

ABSTRACT

We tested whether joint spectrotemporal (ST) sensitivity follows from ST separability for hearing-impaired patients, equipped with bimodal restorative hearing devices: a cochlear implant on one side, and a hearing aid on the contralateral side. We presented broad-band dynamic ripples, which varied over all combinations of 7 spectral ([0-4] cycles/octave) and 13 temporal (±[0-64] Hz) frequencies, in an ongoing harmonic complex sound of randomized duration. Listeners had to react to the unpredictable ripple onset, by pressing a button as fast as possible. We determined their spectral-temporal sensitivity (Modulation Transfer Function) on the basis of their reaction times to the ripples. The joint MTF was partially inseparable for frequency and time modulations, as the MTF could not be fully described by a pure spectral and temporal component. A correlation analysis showed that ripple sensitivity is a strong performance indicator of speech understanding in quiet and in noise. We also found that binaural performance was always slower than predicted by the benchmark race model of statistical facilitation, which assumes independent monaural processing channels.

INTRODUCTION

Users of a cochlear implant (CI) often benefit from a contralateral hearing aid (HA) in challenging listening environments. Advantages are thought to arise from either binaural processing in the auditory system across overlapping frequency bands, or to result from the complementary information provided by the low-frequency acoustic HA input to the CI's electrical stimulation at higher frequency sites in the cochlea. Many studies have demonstrated improved speech perception, especially in noise, as a result from 'bimodal hearing' (Ching et al., 2007; Veugen et al., 2016b).

In current clinical and research practice, hearing performance is typically determined using traditional measures of speech understanding, including percentage words correct or speech reception thresholds. We aimed to investigate the benefit of bimodal stimulation more systematically and objectively, by using stimuli with co-varying spectral and temporal modulations, called ripples. Ripple stimuli represent naturalistic sounds, like vocalizations, in their spectral-temporal (ST) modulations, but cannot be recognized as speech, and could therefore be used to create an efficient parametric test for measuring sensitivity to speech-relevant signals, without an influence of the subject's linguistic and cognitive skills. We expected bimodal hearing to follow a 'best of both worlds' principle, which assumes that the information available at both ears could be successfully integrated in the brain.

Several studies using ripple stimuli have reported degraded temporal sensitivity in hearing impaired listeners, which seemed simply the consequence of their restricted hearing bandwidth (Bacon et al., 1985; Bacon et al., 1992; Moore et al., 2001). Besides reduced temporal sensitivity, sensorineural hearing loss also results in decreased spectral resolution (Henry et al., 2005; Summers et al., 1994). A CI typically reduces spectral resolution, as a result of its band-pass filtering mechanism, cross talk to neighboring electrodes, and the limited electrical dynamic range that can be applied (Henry et al., 2003; Henry et al., 2005). Sensitivity to temporal modulations is also worse in CI users as compared to normal hearing (Golub et al., 2012; Won et al., 2011). Compared to normal hearing, the sensitivity to combined spectro-temporal modulations is degraded in hearing-impaired subjects and CI users, especially for the higher spectral modulations (from 2 cycles/octave upward; (Bernstein et al., 2013; Won et al., 2015). Golub et al. (2012) evaluated performance with the bimodal combination of a CI and HA, but only for combined electroacoustic (hybrid) stimulation (EAS) in the same ear. It was found that electroacoustic spectral sensitivity was superior when compared to the CI alone, but the ability to detect temporal modulations was the same for both types of stimulation.

Multiple studies have found a correlation between spectral-ripple resolution and speech, consonant and vowel perception in quiet and in noise for CI users (Anderson et al., 2012; Berenstein et al., 2008; Fu, 2002; Henry et al., 2003; Litvak et al., 2007; Won et al., 2007). Also high temporal modulation detection thresholds (> 100 Hz) correlate well with word scores for CI users (Won et al., 2011). In hearing-impaired listeners, speech scores could be predicted by spectral ripple detection (at 2 cycles/octave) (Davies-Venn et al., 2015), as well as by spectro-temporal ripple detection (for ripples at 2 cycles/octave (spectral), and 4 Hz (temporal) modulations; (Bernstein et al., 2013). For bimodal listeners, the HA-aided spectral modulation detection threshold for a ripple with 1 cycle/octave could well predict a bimodal benefit for speech understanding (Zhang, 2013).

The majority of studies in hearing-impaired subjects have so far used pure temporal-only or spectral-only stimulus modulations, to quantify temporal modulation sensitivity, or spectral modulation sensitivity, respectively. However, speech typically contains joint ST modulations. Joint ST sensitivity is often implicitly assumed to arise from ST separability, which was proposed by Chi et al. (1999) for normal-hearing subjects (see, however, Veugen et al., 2016d). However, hearing loss could potentially induce changes in the mechanisms underlying spectral and temporal separability. This would follow from significant interaction effects between spectral and temporal modulations, as reported by Bernstein in hearing-impaired subjects (Bernstein et al., 2013). In addition, we recently demonstrated that S-T separability of a HA and CI simulation in normal-hearing listeners was less separable than under normal-hearing conditions (Veugen et al., 2016d). Thus, the question remains whether spectral- and temporal-only sensitivity functions would indeed capture all information about the full auditory ST sensitivity of CI and HA users. Using the same reaction-time paradigm as in our recent study (Veugen et al., 2016d), we here aimed to investigate: i) the benefit of bimodal stimulation for various S-T modulations, ii) ST-separability in bimodal hearing-impaired listeners, and iii) the relationship between ripple sensitivity and speech reception thresholds. Our findings demonstrate that reaction times provide an efficient, solid, and objective measure for S-T sensitivity of hearing-impaired listeners, and that the results can be used to predict speech-in-noise intelligibility.

METHODS

SUBJECTS

Fourteen post-lingual deaf listeners participated in this study (9 male; mean age 64 ± 12 years). All used the same CI processor (Advanced Bionics Harmony) and HA (Phonak Naida S IX UP) in opposite ears, that were fitted in all subjects using a loudness balancing procedure described in an earlier study (Veugen et al., 2016a). Compression characteristics in the HA were adapted for research purposes to match the compression of the CI (Veugen et al., 2016b). Audiometric thresholds and demographic details of the listeners are reported in Table 1. The study was approved by the Local Ethics Committee of the Radboud University Nijmegen (METC, protocol number CMO 40327.091.12).

The methods are largely similar to our previous study (Veugen et al., 2016d), and will be summarized here briefly.

STIMULI

We generated dynamic ripple stimuli, which are broadband sounds with sinusoidally modulated spectral and temporal envelopes (Depireux et al., 2001). The broadband carrier of these sounds consisted of 128 harmonic tones, $c_i(t)$ (i=0-127), equally spaced (20 tones/octave) over 6.4 octaves (here, from 100 Hz to 8.2 kHz):

$$c_i(t) = \sin(2\pi f_i t + \varphi_i) \tag{1}$$

All components had a randomly selected phase except for the first (=0).

Stimuli start out as the non-modulated harmonic complex sound, and after a randomized duration, t_{onset} (between 700 and 1200 ms, in steps of 100 ms), the modulation was initiated, as follows:

$$R(t) = \begin{cases} t < t_{onset}, & \sum_i c_i(t) \\ t > t_{onset}, & \sum_i (1 + \Delta M \cdot \sin(2\pi\omega(t - t_{onset}) + 2\pi\Omega x_i) \cdot c_i(t) \end{cases} \tag{2}$$

with t time (s), x_i the position on the frequency axis (in octaves above the lowest frequency), ω the ripple velocity (Hz), and Ω ripple density (cycles/octave). We fixed the modulation depth, ΔM, at 0.5 for all ripples. The ripples lasted 3000 ms, and were presented at 13 logarithmically spaced (base 2) velocities between 0 and ±64: ±[0; 0.5; 1.3; 3.5; 9.2; 24.3; 64] Hz and at 7 densities [0; 0.25; 0.5; 0.75; 1; 2; 4] cyc/oct, yielding a total of 91 different stimuli. Stimuli were generated using a sampling rate of 48828.125 Hz.

APPARATUS

The experiments took place in an acoustically shielded sound chamber with its walls, ceiling and floor covered by black foam (50 mm thick with 30 mm pyramids, AX2250, Uxem, Lelystad, NL) and with dimmed lights. Subjects were seated in a chair at 1 meter from a loudspeaker (JBL Control 1, Harmon International Industries, Washington DC, USA), that presented the ripple stimuli at a comfortable, well-audible 65 dB(A) (measured with a Brüel & Kjaer 2260 Investigator). Sounds were sent via the PC to a Tucker Davis Technologies System 3 (Alachua, FL, USA) consisting of a real-time processor (RP2.1).

PARADIGM

Listeners were instructed to press a button as quickly as possible, as soon as they heard the sound stimulus change from a static into a dynamic ripple at t_{onset}. Modulated ripples lasted for 3000 ms, unless the button was pressed, in which case the trial ended prematurely and the next trial was initiated. The trial was reiterated if the button was pressed before t_{onset}, but no more than 4 times. There was a short break period of 0.5-1 seconds of silence between each trial. Outcome measure of the experiment was the listener's manual reaction time, which was defined as the time between the t_{onset} and the moment of the button press.

We tested three hearing conditions: CI-only, HA-only and bimodal stimulation (CI+HA).

***Table 1**. Demographic details of the participating bimodal listeners. Hearing thresholds beyond the audiometer limit (120 dB HL) were assigned a value of 125 dB HL.*

Participant	Age (yr)	Gender	CI ear	Etiology	Experience (yr)	
					CI	Bimodal
1	65	M	L	Unknown	7.4	7.4
2	48	M	L	Genetic	5.1	5.1
3	81	F	L	Hereditary	10.3	1.4
4	60	M	R	Ototoxicity	6.7	6.7
5	76	M	R	Genetic	2.7	2.7
6	53	F	R	Hereditary	6.8	1.2
7	45	M	R	Congenital	5.2	5.2
8	77	M	L	Unknown	2.5	1.4
9	66	M	R	Unknown	4.0	1.2
10	80	F	R	Unknown	8.2	1.3

For CI- and HA-only, we presented 43 out of 91 ripple stimuli. Ripples were presented in pseudo-randomized order and were each repeated 8 times. Four measurement blocks, lasting 20-25 minutes each, were performed, always starting with the bimodal condition (364 trials), thereafter CI- or HA-only (counterbalanced across subjects, 344 trials each), and ending again with bimodal stimulation (364 trials). Prior to the actual experiment, subjects underwent a brief training session until they were comfortable with the experimental task (< 5 minutes for all subjects). During the experiment, we provided encouraging visual feedback on a screen in front of the subject, displaying their reaction time immediately after a trial, and a progress bar that indicated percentage completion of the block.

ANALYSIS

We analysed data in MATLAB [version R2012a; Mathworks Inc., Natick, MA, USA]. Reaction times shorter than 150 ms (premature response) were removed from the data analysis. Non-responses were assigned the maximum possible reaction time of 3000 ms (moving ripple duration). Reaction times generally show a skewed distribution with a long tail towards slower reaction times. Therefore, the data was transformed to its reciprocal (1/reaction time), further referred to as promptness, to obtain normally distributed data for further analyses (Carpenter et al., 2009). For each subject and condition, we calculated the mean promptness per ripple to construct a two-dimensional modulation transfer function with density along the ordinate, and velocity along the abscissa.

CI Strategy	Hearing threshold non-implanted ear (dB HL)				
	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz
HiRes-P	65	75	85	120	125
HiRes-S F120	65	80	115	120	125
HiRes-S F120	50	65	105	120	125
HiRes-P	60	70	80	95	100
HiRes-S F120	80	80	90	80	80
HiRes-S F120	60	85	110	105	120
HiRes-S F120	65	95	105	120	120
HiRes-S	30	60	105	110	125
HiRes-S	65	90	85	95	80
HiRes-S	60	85	110	115	125

SEPARABILITY

We applied singular value decomposition (SVD) to investigate separability of the spectro-temporal modulation transfer function (stMTF) into a pure temporal (tMTF), and a pure spectral (sTMF) component. Using only the first singular value (σ_1) and its associated temporal ($G_1(\omega)$) and spectral ($H_1(\Omega)$) singular vectors of the SVD, we reconstructed the separable estimate of the listener's stMTF, by

$$\widehat{stMTF}_1(\omega, \Omega) = \sigma_1 G_1(\omega) H_1(\Omega) \quad (3)$$

and calculated Pearson's linear correlation coefficient with the original data as a measure of separability (r=0 (inseparable) to r=1 one (full separability)). We compared this prediction with a second reconstruction that also included the second singular value and its associated singular vectors:

$$\widehat{stMTF}_{12}(\omega, \Omega) = \sigma_1 G_1(\omega) H_1(\Omega) + \sigma_2 G_2(\omega) H_2(\Omega) \quad (4)$$

As an additional measure, we calculated the so-called separability index, representing the relative dominance of the first singular value of the SVD (Veugen et al., 2016d):

$$\alpha_{svd} = \sigma_1^2 \Big/ \sum_i \sigma_i^2 \quad (5)$$

RACE MODEL

We compared the promptness values obtained in the bimodal condition to the monaural CI and HA conditions, using the race model of statistical facilitation. This benchmark model assumes that binaural responses result from the winner of an independent race between the processing channels of both ears, and predicts faster bimodal reaction times from the distribution of minimum reaction times obtained from the CI-only and HA-only conditions (Gielen et al., 1983b; Raab, 1962). Significant violations to this model (either faster, or slower than race responses) suggest binaural interactions between central neural processing stages of the inputs from the CI and HA. We calculated the cumulative probability of promptness for the three listening conditions, which were fitted with a cumulative distribution function (CDF). The cumulative race-model prediction for bimodal reaction times was constructed from the monaural CDFs, as described in our previous study (Veugen et al., 2016d):

$$P\left(\tau_{\mathbf{Race,BIM}} \leq \mathrm{t}\right) = P(\tau_{\mathrm{CI}} \leq \mathrm{t}) + P\ (\tau_{\mathrm{HA}} \leq \mathrm{t}) - P(\tau_{\mathrm{CI}} \leq \mathrm{t}) \times P(\tau_{\mathrm{HA}} \leq \mathrm{t}) \qquad (6)$$

We compared the median of the measured bimodal promptness CDF with the median predicted by the race model (Eqn. 6), as well as with the median promptness of the fastest monaural condition. Non-responses (reaction times > 2500 ms) were discarded from this analysis. Data for upward and downward ripples were pooled. Ripples with fewer than six responses were also discarded from the race model analysis; this was the case on average in 57 ± 28 % of the trials per subject (range: 12-100%, see also Figure 5).

CORRELATION WITH SPEECH UNDERSTANDING

Ripple responses were correlated with different measures of speech comprehension, which we collected for the same group of participants in a previous study (Veugen et al., 2016b). In brief, speech understanding in quiet was assessed in the CI-only, HA-only and bimodal listening condition with the NVA (Nederlandse Vereniging voor Audiologie) Dutch monosyllabic word test (Bosman 1992), resulting in an average percentage of phonemes correct across three lists of twelve words. Speech understanding in noise was measured in the CI-only and bimodal condition using the Leuven Intelligibility Sentence Test (LIST; van Wieringen et al. (2008)). We determined the signal-to-noise ratio (SNR) at 50% performance (SNR50) with target speech presented from the front. The SNR50 was determined for speech-shaped steady-state noise, simultaneously presented from both the left and the right side, as well as for single-talker babble noise, averaged over three speaker configurations (noise from the left, from the right and from both sides). For each ripple, we calculated Pearson's correlation coefficient between promptness versus the percentage phonemes correct (speech in quiet), and versus the SNR50 (speech in noise) in the corresponding listening condition. We also calculated correlation coefficients between ripple promptness and hearing thresholds (averaged over 250, 500, 1000, 2000 and 4000 Hz). Data for upward and downward ripples were pooled to reduce the number of variables.

RESULTS

To reduce the impact of outliers in the analysis, the reaction times were transformed to their reciprocal, promptness, which yields a more symmetric, Gaussian-like distribution of the data. Moreover, the interpretation of promptness as a sensitivity measure is more intuitive, as a high/low sensitivity corresponds to a high/low promptness value. A peak in reaction time around 0.5 s thus corresponds to a promptness value of 2 s^{-1}. For a Gaussian process, the cumulative distribution of promptness is expected to follow a straight line on a probit scale.

To illustrate this aspect of the data, Figure 1 shows examples of the recorded promptness distributions in a reciprobit format for four different patients during bimodal listening at three different ripples (0 cyc/oct and 9.2, 1.4 and 0.5 Hz). Note that the data typically follow a straight line, in which the slopes (a measure for the width of the distribution) and offsets (together with the slope related to the median reaction time) depend systematically on the ripple velocity, which is an acoustic parameter in the experiments. The (too early) outlier reaction times, indicative of predictive responses, guessing, or inattentiveness of the listener, are clearly identifiable in these plots. Figure 1C,D show different behavior in promptness data from two other listeners. Participant 7 (Fig. 1C) had some very short reaction times to the ripple of 9.2 Hz, even leading to some premature responses. Patient 2 (Fig. 1D) was unable to detect the ripple of 4 c/o and 0.5 Hz, which resulted in responses that were either too early (data scatter along a line with a low slope, despite the fact that these responses fell within the normal response-time window), or (the majority) far too late (and therefore all placed at the maximum of 3000 ms, see Methods). These cumulative distributions thus contain a wealth of information about the difficulty of the task for the listener, even on a trial-by-trial basis. In what follows, we will solely focus on the mode (or mean) (at 50% cumulative probability) of the promptness distributions, as a robust measure for ripple sensitivity.

AMPLITUDE MODULATIONS

Figure 2A shows the mean promptness to ripples with pure temporal modulations (Ω=0 cyc/oct) for bimodal, CI-only and HA-only hearing. The promptness values for the CI-only and bimodal hearing conditions were comparable, but responses for the HA-only condition were much slower (lower promptness values). Fastest responses were found at ±24 Hz. At higher velocities (±64 Hz) the relationship between ripple velocity and promptness declined.

The stimulus without S-T modulations (at ω=0 Hz and Ω=0 cyc/oct) served as catch stimulus, to determine the guess rate. Subjects pressed the button in 20 ± 27 % of these trials (note: below 17 % for 10/14 subjects).

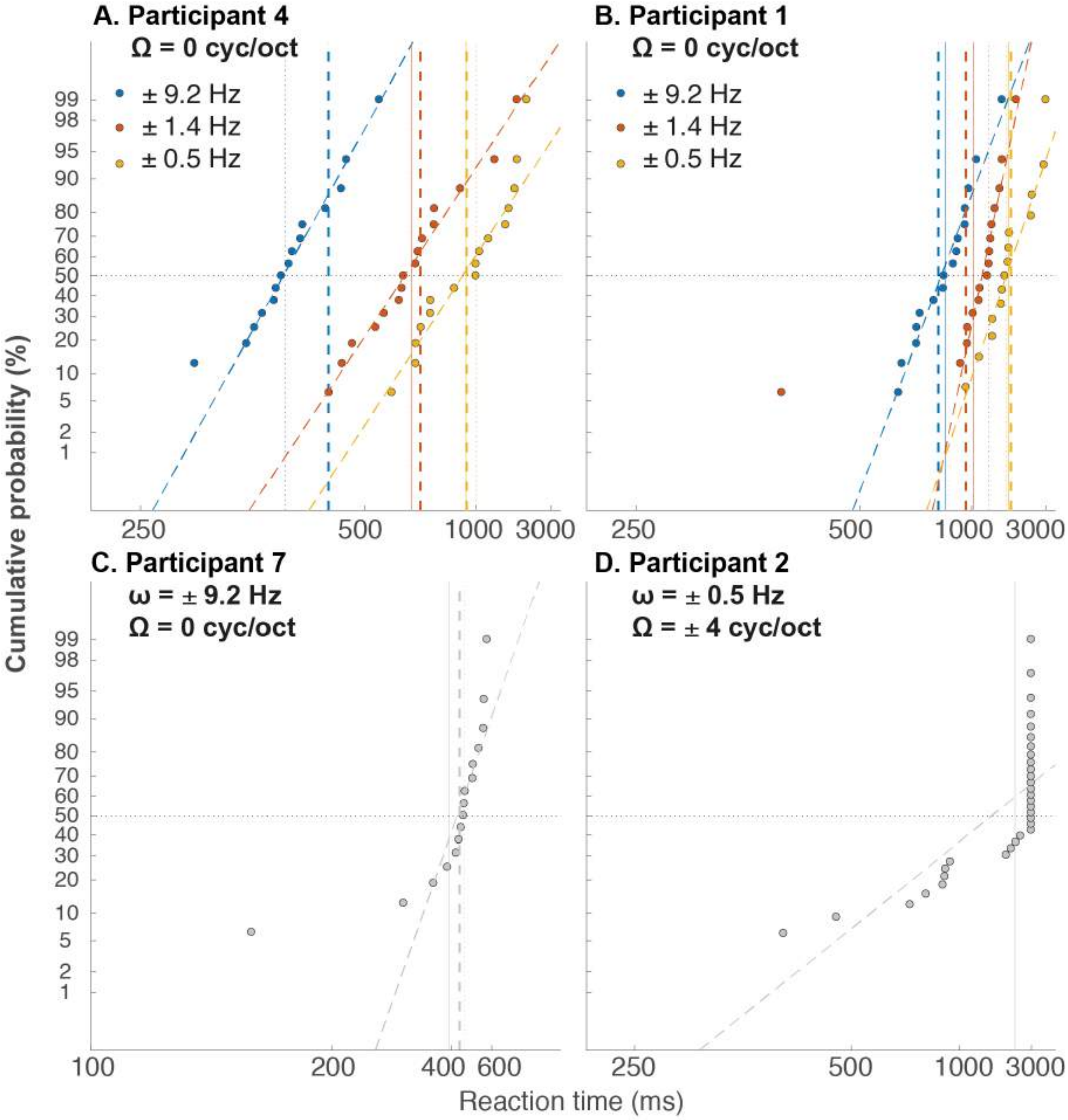

Figure 1. Examples of reciprobit plots for four different patients during bimodal listening. ***A)*** *Cumulative reaction times of patient S4 to ω = ±0.5, 1.4 and 9.2 Hz velocity modulations (and a density of Ω = 0 cyc/oct). Note that data points nicely align along a straight line, and that these lines systematically shift with ripple velocity. Points that strongly deviate from the line (e.g., the blue data point at ~270 ms) are outliers, due to inattentiveness, predictive responses, or mere guesses.* ***B)*** *Data for patient S10 for the same stimuli; qualitatively similar results as for S4. The red outlier is a predictive response.* ***C)*** *Data from patient S7 to the 9.2 Hz ripple; note several predictive responses.* ***D)*** *Data for patient S2 to an imperceptible ripple (ω = 0.5 Hz and Ω = 4 c/o), evidenced by many predictive responses/guesses, and maximum reaction times at 3000 ms.*

SPECTRAL MODULATIONS

Promptness to ripples with pure spectral modulations (ω=0 Hz) rapidly declined with increasing modulation rate (Figure 2B), suggesting that ripple detection became harder for increasing spectral detail. Again, reaction times for the bimodal and CI-only conditions were indistinguishable. Ripple detection in the HA-only condition was poorest and on average 713 ± 449 ms slower than for bimodal hearing. The number of non-responses (resulting in a promptness of 0.33) in the bimodal and CI-only condition increased with higher spectral modulation rates from about 20% at 0.25 c/o to 75% at 4 c/o. For HA-only these numbers were even higher: 59% at 0.25 c/o to 82% at 4 c/o.

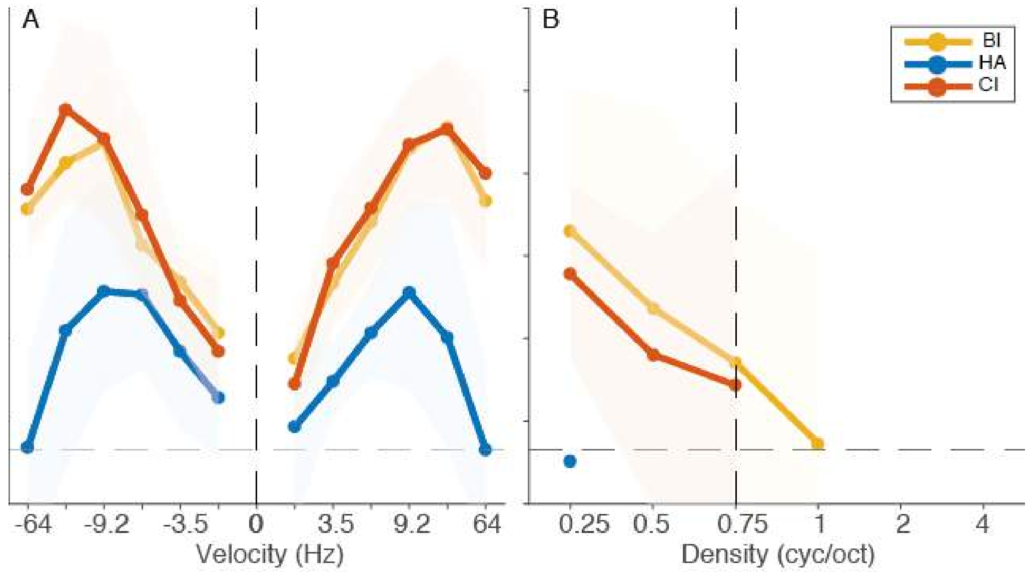

Figure 2. *Promptness in the CI, HA and bimodal conditions for ripples with **A)** pure temporal modulations (note logarithmic scale; density = 0 cyc/oct), and **B)** pure spectral modulations (velocity = 0 Hz).*

SPECTRO-TEMPORAL MODULATIONS

Mean promptness per ripple is visualized in Figure 3 for the bimodal and monaural CI-only and HA-only listening conditions. Note that we only measured part of the ripples for the monaural conditions (see Methods). For bimodal stimulation, spectro-temporal resolution showed a band-pass filter pattern for both spectral and temporal modulations. Detection was fastest for ripples with modulation rates around ω=9-24 Hz and Ω=0.25 cycles/octave. Like for the pure temporal (bottom row in the matrix) and pure spectral ripples (central vertical line in the matrix), ripple responses to joint S-T modulations for CI-only hearing were roughly comparable to, and perhaps even a bit faster than, bimodal stimulation. In contrast, S-T responses in the HA-only condition were much slower. Mean reaction times across the ST ripples that were measured in all three conditions were 1132 ± 920, 1125 ± 946 and 2091 ± 1032 ms [excluding the pure T and pure S responses] for respectively the bimodal, CI-only and HA-only condition.

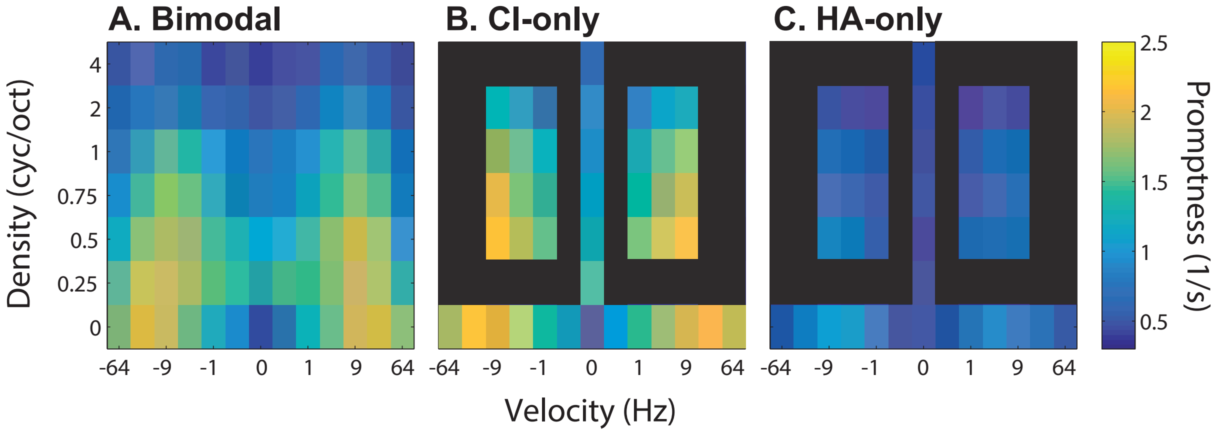

__Figure 3. A-C)__ Mean promptness (1/reaction time) as a function of velocity and density, representing spectro-temporal sensitivity in the __A)__ bimodal, __B)__ CI-only and __C)__ HA-only condition. The higher promptness, the faster ripple detection. Non-measured ripples are indicated by the black pixels.

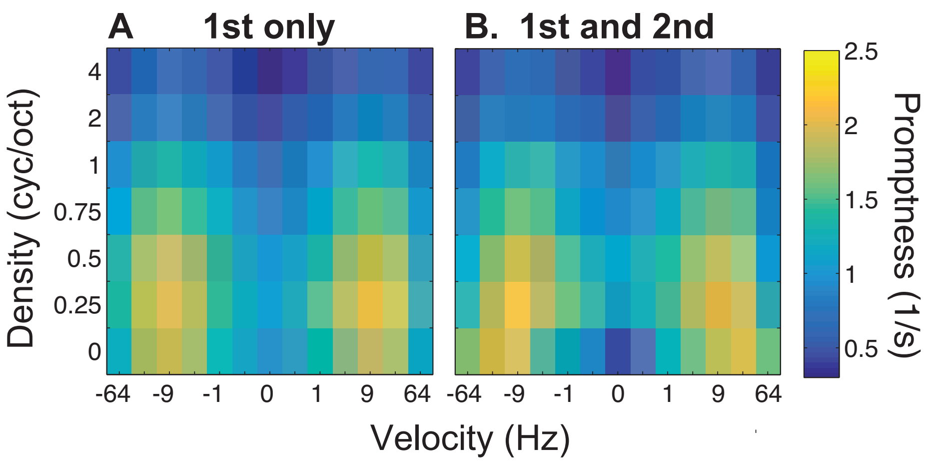

__Figure 4. A-B)__ Reconstructions of the spectrotemporal transfer function of bimodal hearing using only the first (Eqn. 3; __A)__ and first two (Eqn. 4; __B)__ components after singular value decomposition (SVD). Comparison with Figure 3A shows that the reconstruction of panel B better describes the original data.

SEPARABILITY

Using singular value decomposition (SVD) we investigated whether the spectro-temporal modulation transfer function could be obtained from a single pure temporal and spectral component (Eqn. 3). The separability index, α_{SVD}, was 0.99 for bimodal stimulation, suggesting a large amount of spectral-temporal separability. Because we only measured the full range of ripples in the bimodal condition, we restricted the separability analysis to this condition only. A reconstruction of the stMTF based on the first spectral and temporal SVD component (Eq. 3) is shown in Figure 4A. The correlation coefficient between the original and reconstructed data was high, r_1=0.95 (bimodal). However, by including the second singular value in the reconstruction (Eq. 4), the correlation coefficient increased to r_{12}=0.98 (bimodal), which was significantly higher than r_1 (z=-3.09; p<0.001; Fig. 4B). The value of the second singular value, σ_2, was approximately 10% (bimodal) of σ_1. Checking for the third (2%) and fourth (2%) singular values indicated no further improvement in the stMTF estimates.

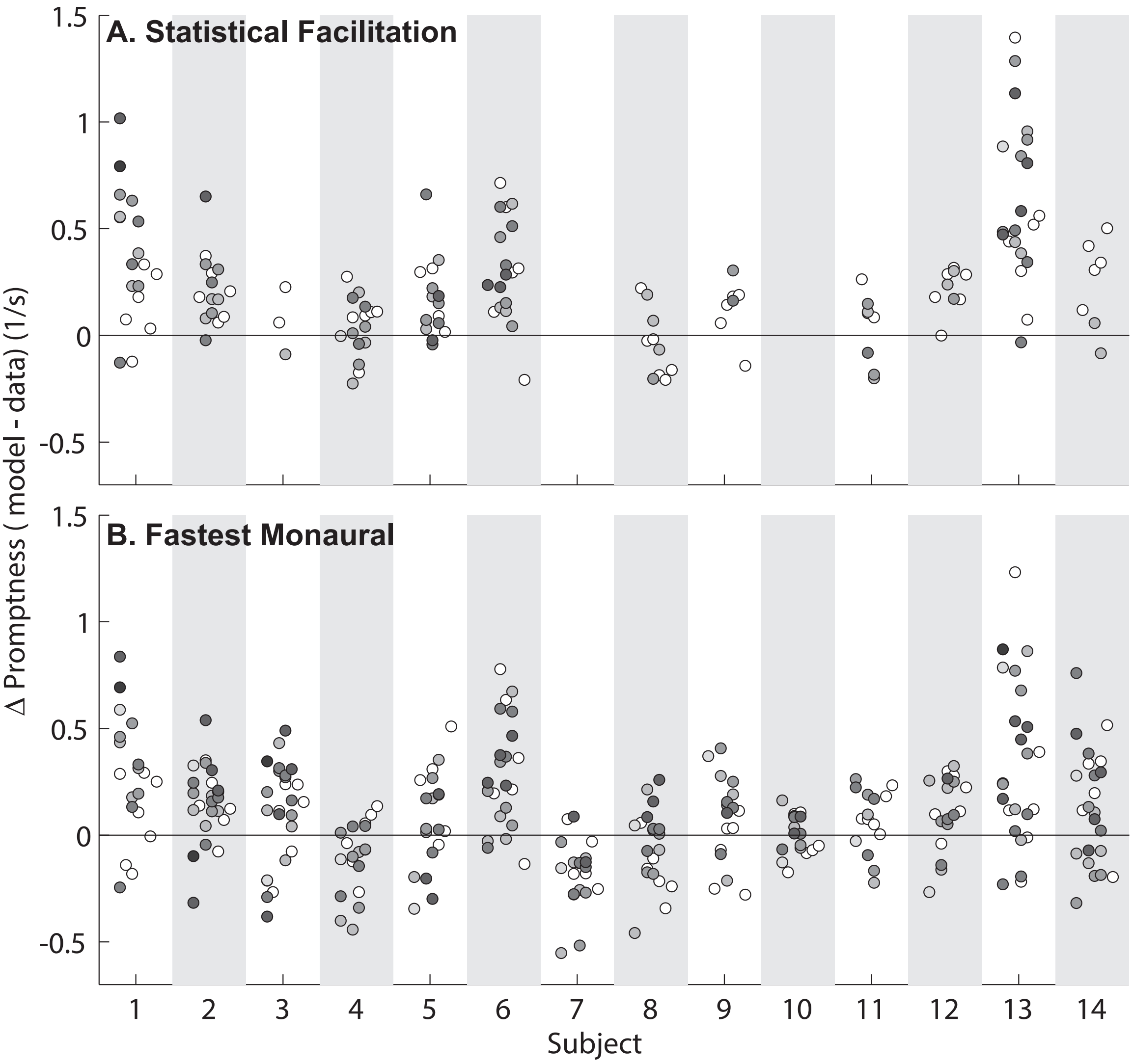

Figure 5. *Individual differences between median bimodal promptness and* ***A)*** *the race model at 50% cumulative probability, and* ***B)*** *the median promptness of the fastest monaural condition. Markers above the line at zero difference in panel* ***A*** *indicate that actual performance was slower than the race model prediction. This is the case for the far majority of ripple responses, despite the fact that many responses could be faster than the fastest monaural listening condition (negative differences in panel* ***B****). Marker gray scale represents spectral modulation rate, ranging from 0 cyc/oct (white) to 4 cyc/oct (black). The horizontal position of the marker for each participant represents the ripple's temporal modulation rate, ranging from 0 Hz (most left) to 64 Hz (most right).*

RACE MODELS

The race model of statistical facilitation (Gielen et al., 1983b; Raab, 1962) served as a benchmark to test whether bimodal hearing would outperform the performance of adding independent monaural responses. We calculated the promptness at the level of 50% cumulative probability of the race model, which was compared to the median reaction time for bimodal hearing. The median bimodal promptness was also compared to the median promptness of the fastest monaural condition, determined per ripple. Individual results are visualized in Figure 5. Responses to the majority of ripples were slower than the predictions from the race model of statistical facilitation (data above the line in Figure 5A), and often also slower than the fastest monaural response (data above the line in Figure 5B). These results suggest that the input from both the CI and HA interacted before eliciting a bimodal response, for the majority of subjects. Only a few subjects (especially subjects 4 and 7, but also 8 and 14) had, for most or half of the ripples, bimodal responses that were faster than the fastest monaural condition. The fastest monaural condition was always the CI (Figure 3), except for subjects 4 and 13 that showed fastest responses with the HA for 7 and 12 ripples, respectively. In subjects 4 and 8, bimodal promptness also exceeded the race model predictions for about half of the ripples. Note that race model predictions could not be calculated for two subjects that could not detect any ripple in the HA-only condition (7 and 10).

CORRELATION WITH SPEECH PERCEPTION

Clear patterns of correlation coefficients may be observed between the promptness for certain ripples and the scores for speech understanding in quiet and in noise (Figure 6). To illustrate this, the correlation result for one of the ripples for promptness vs. the signal-to-noise ratio at 50% speech perception (SNR50) is shown in Figure 6F, indicating that faster ripple responses are indicative for better speech perception. For speech understanding in quiet, we found that higher promptness (fast responses) correlated with a higher percentage phonemes correct scores (Figure 6A-C). For speech understanding in single-talker noise, a high promptness correlated with lower (i.e., better) SNR50 values (Figure 6DE). Similar results were found for the SNR50s in steady-state speech-shaped noise (not shown). Correlations were strongest for speech understanding in noise in the bimodal listening condition (which subjects used on a daily basis), resulting in p-values < 0.05 for half of the ripples. For both measures of speech understanding, we obtained a similar pattern of correlation coefficients, with strongest correlations around ω=9 Hz and below Ω=2 cycles/octave, the same S-T region that elicited the fastest ripple responses (Figure 3A-C). In this range, p-values could be as low as 0.002 for bimodal hearing, while ripple sensitivity could explain up to 57% (r^2 at 9 Hz, 0 c/o) of the performance for speech understanding in single-talker noise. No strong correlations were obtained between ripple promptness and the average hearing threshold in the non-implanted ear (mean p-value across all ripples and conditions was 0.58 ± 0.26). Multiple linear regression showed that both the average hearing threshold and SNR50 in the bimodal condition accounted for on average 42 ± 15 % (max 73% at 9 Hz and 1 c/o, $p < 0.0001$) of the variance in ripple promptness, compared to 30 ± 15 % (52% at 9 Hz and 1 c/o, $p = 0.003$) for the speech data alone. We did not obtain strong correlations between the bimodal benefit for speech understanding, i.e. the difference between CI-only and bimodal hearing, versus ripples responses in the CI-only, HA-only or bimodal condition, or the difference between CI-only and bimodal hearing (the mean correlation coefficient across all ripples and conditions was r = 0.05 ± 0.2, p = 0.62 ± 0.25).

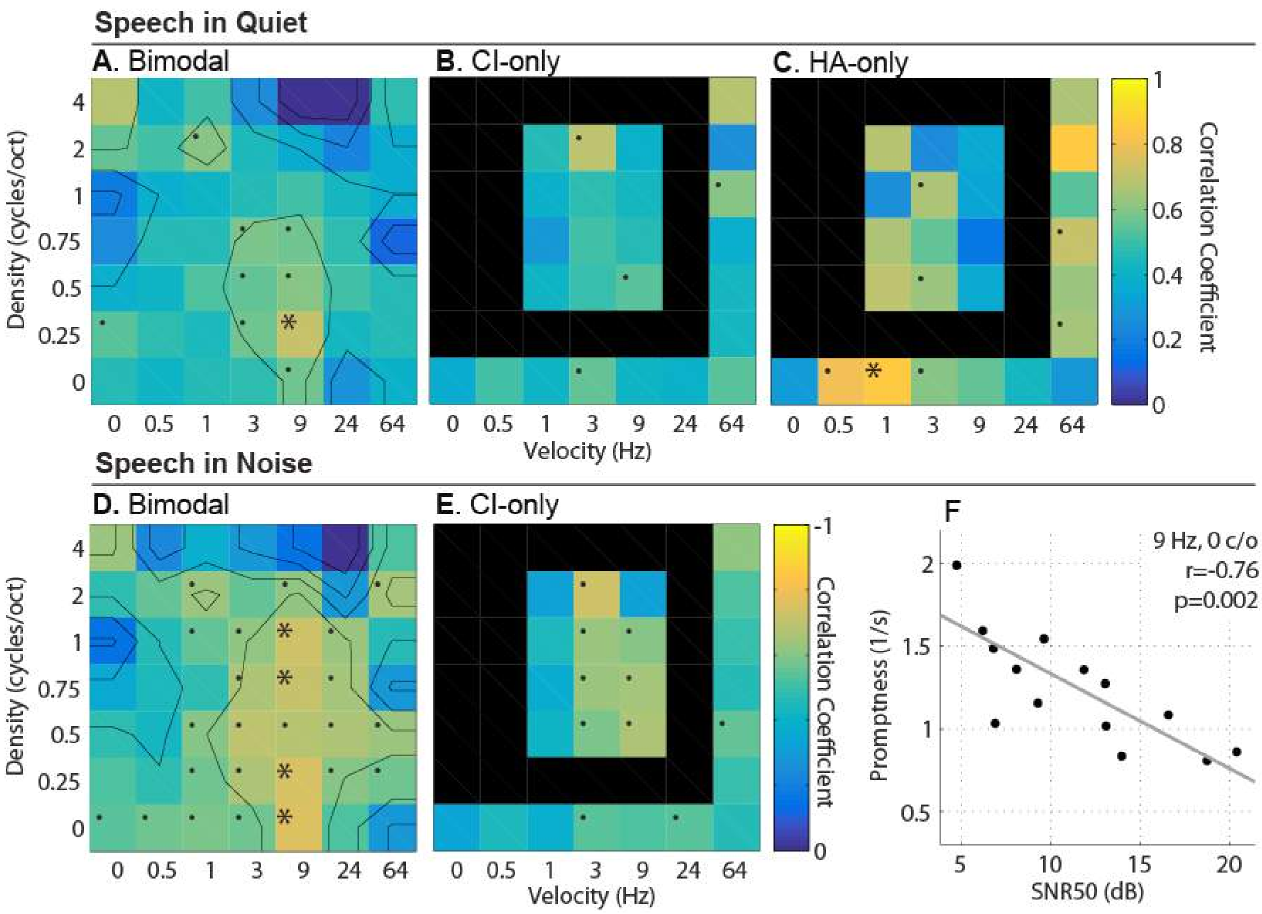

Figure 6*. Linear correlation coefficients between ripple promptness and measures of speech understanding, including* ***A-C)*** *percentage phonemes correct in quiet in the bimodal, CI-only and HA-only condition, and* ***DE)*** *the signal-to-noise ratio at 50% performance (SNR50) for speech understanding in single-talker noise in the bimodal and CI-only condition.* ***F)*** *Example for ripple (ω,Ω)=(9 Hz, 0 c/o) showing a strong (negative) correlation between promptness and SNR50 in the bimodal condition. * p < 0.005, ▪ p < 0.05*

7

DISCUSSION

In this paper, we investigated sensitivity to ripples with joint spectrotemporal modulations in fourteen hearing-impaired listeners, all equipped with bimodal restorative hearing, with a CI and HA in contralateral ears. Listeners were instructed to press a button as soon as they heard the sound change from a static noise into a modulated ripple. This is an easy task to perform, which does not require specific training for the listener. In addition, the task hardly imposes a cognitive load on the listener, as stimuli were typically presented well above the hearing threshold, and did neither require their active recognition, nor an indirect comparison to other stimuli. The task can easily be applied to nearly all categories of listeners, including children and experimental animals (e.g., Massoudi et al., 2014), and hundreds of trials can be obtained in a relatively short time to allow for a solid statistical evaluation. The presence of a response within a relatively narrow time window can ensure the experimenter in an objective way that the listener indeed perceived the ST modulation. Outlier responses, or inattentiveness of the participant, are readily recognized as such, and can be discarded from further analysis, if needed (Fig. 1).

Reaction times could be used as an objective measure of spectral-temporal sensitivity, as they depend in a systematic way on the ripple velocity (temporal AM) and density (spectral AM) parameters [Veugen et al., 2016d]. Outliers in the data are easily identified from the reciprobit plots, as acoustically-elicited responses are well described by a straight-line relationship, for which the slope and intercept depend systematically on the acoustic manipulations (Figs. 1 and 3). Similar acoustic sensitivity measures from reaction-time data have been reported for goal-directed saccadic eye movements of normal-hearing listeners to the locations of brief noise bursts (Hofman and Van Opstal, 1998), for manual reaction times to AM-noises in monkeys (Massoudi et al., 2014), and for ripple detection under normal-hearing, and simulated hearing-impaired conditions (Veugen et al., 2016d).

On average, the reaction times to bimodal hearing were comparable to responses evoked under the monaural CI condition (panels 3A and 3B), which elicited much faster reaction times than the monaural HA condition (panels 3B vs. 3C). We found that the joint spectral-temporal sensitivity of bimodal listeners could be best described by including two singular values (Fig. 4B), suggesting that ST hearing in these patients is not fully separable. As a consequence, to estimate the full joint ST sensitivity of hearing-impaired listeners, measurements of pure temporal and pure spectral modulations only will not suffice. We also found that the bimodal reaction times to ripple onsets could not be predicted from the CI/HA-only performance data, suggesting the involvement of an inhibitory neural integration stage of binaural information (see below). Finally, the relatively high correlations for specific ripples in our data (Fig. 6) showed that ripple sensitivity obtained with manual reaction times may be a good performance indicator of speech understanding in quiet and in noise.

BIMODAL VERSUS MONAURAL PERFORMANCE

For the majority of ripples, detection with bimodal stimulation was roughly similar to the CI-only condition in our study (Figs. 2, and 3A vs. 3B). Golub et al. (2012) tested a comparable group of listeners, using electro-acoustic stimulation (EAS) in the same ear. Their results also showed that temporal modulation detection was largely similar between EAS and CI-only hearing. However, spectral ripple detection with EAS was superior to CI-only. Possibly, residual hearing in their EAS users (hearing thresholds were not given) may have been better than in our subject group, given the fact that performance with only acoustic stimulation in the Golub et al. (2012) study was comparable to combined EAS hearing. The study of Won et al. (2011) compared temporal ripple detection thresholds of CI users with the data from hearing-impaired and normal-hearing listeners collected by Bacon and Viemeister (1985). Note that those hearing-impaired listeners still had 10-20 dB HL at 1 kHz, and that the remaining listening bandwidth in the low-frequency range plays a key role in temporal sensitivity (Bacon et al., 1985). It is conceivable that for this reason, the CI-only thresholds were inferior compared to both groups in the study of Won et al. (2011). More recently, Won et al. (2015) obtained similar results for ripples with joint spectro-temporal modulations. These observations contrast markedly with the superior performance of CI-only hearing, when compared to the poor performance for HA-only hearing in the group of patients of the present study.

S-T SEPARABILITY

Separability of ST sensitivity was reported earlier for normal-hearing subjects (Chi et al., 1999), although Bernstein et al. (2013) found a joint interaction between spectral and temporal components, which hinted at non-separability for both normal hearing and impaired-hearing listeners. Although the SVD analysis on our S-T sensitivity data (Fig. 4) suggested highly separable spectral-temporal processing in the auditory system, explaining about 90% (r_1^2) of the variance in the stMTF, we found a highly significant contribution of the second singular vectors describing the temporal and spectral modulations, which explains 96% (r_{12}^2) of the total variance. We recently obtained similar results for normal-hearing listening comparable ripples, as well as for hearing under simulated hearing-impaired conditions (CI-only, HA-only and bimodal, simulated by vocoders; Veugen et al., 2016d) by these same listeners. Despite the fact that the second singular value amounted to only 10% of the first component, it nearly fully accounted for the diagonal (i.e. joint) ST modulations that are clearly visible in the stMTF (cf. Figures 3A and 4B).

As suggested in our previous study (Veugen et al., 2016d), this second component mainly affects the stMTF at 0-density modulations (i.e., pure temporal amplitude modulations), making the reaction times to become slower at low-velocity modulations, and faster at higher velocities. This suggests that the second component reflects an AM detection rate that can operate faster at higher velocities, simply because the stimulus modulation period is shorter. This result is in line with the linear dependence recently reported for monkey reaction times on the AM period (Massoudi et al., 2014). This additional component is independent of how well a modulation frequency can be detected by the auditory system. This property of the system, which can be described by a band-pass filtering characteristic for temporal modulations (Fig. 2A), and a low-pass characteristic to spectral modulations (Fig. 2B), is reflected in the first-component of the stMTF.

In summary, a full assessment of ST performance in hearing-impaired patients requires testing of the complete two-dimensional ST field of relevant ripples, to include the second SVD component (Fig. 4B).

RACE MODEL AND BIMODAL INTEGRATION

The race-model analysis of the data in Fig. 5 indicates that none of our hearing-impaired listeners produced systematically faster responses than predicted on the basis of statistical facilitation. Instead, most responses were *slower* than predicted by the race model (Fig. 5A), even though many ripples elicited faster bimodal responses than the fastest monaural hearing condition (typically the CI; Fig. 5B). We obtained a similar result for simulated hearing-impaired listening with normal-hearing participants (Veugen et al., 2016d). It should be noted, however, that also when responses are slower than predicted by independent processing channels, a neural binaural interaction may be held responsible for this effect. This apparent slowing of ripple-evoked reaction times under bimodal hearing may then be interpreted as a slight *decrease* in ripple sensitivity.

As CI responses resulted to be close to ceiling performance (fastest possible reaction times), and given that in our group of listeners there was very little spectral overlap between the HA and CI sides, it may not be too surprising that a positive bimodal integration in the form of exceeding race-model performance was absent in our data. Indeed, the inferred negative integration effect could reveal an inhibitory neural interaction, which might result just because of the lack of sufficient spectral overlap between the two ears, and/or the impossibility for further speeding of responses.

For comparison, in audiovisual experiments, which probe the neural integration of visual and auditory channels, the strongest benefits of multisensory interactions are typically obtained when the stimuli overlap both in space and time (thus providing multisensory evidence for a single object), and, in addition, when both modalities in isolation yield variable and slow responses (i.e., away from ceiling performance. This phenomenon has become known as the principle of inverse effectiveness, and has been demonstrated both at the neurophysiological single-unit level in the midbrain orienting system (superior colliculus; e.g. Stein and Meredith, 1993), as well as in the speed and accuracy behavior of saccadic eye movements (Corneil et al. 2002; Van Wanrooij et al., 2009).

In line with the principles underlying multisensory integration, we argue that the effects of bimodal integration will be strongest when the auditory system has sufficient sensory evidence that the left vs. right acoustic inputs arose from the same acoustic object, rather than from different, unrelated sounds. We thus hypothesize that stronger bimodal benefits (i.e., enhanced sensitivity) may be expected: (i) when the effective spectral ranges of CI and HA overlap sufficiently (to allow for frequency-specific binaural comparisons), (ii) when technical time delays between the devices are negligible (so that interaural time delays remain within the physiological range), and (iii) when the monaural reaction-time distributions have sufficient variability, and overlap considerably. In the current experiments, there was little overlap in the reaction-time distributions of CI and HA (cf. panels 3B and 3C), leading to race-model predictions that virtually coincided with the CI-only data, and therefore provided little room for true excitatory binaural integration. Furthermore, the absence of sufficient spectral (and possibly temporal) overlap may have prevented the auditory system to assign the bilateral acoustic inputs to a single auditory object altogether.

RELATION TO SPEECH INTELLIGIBILITY

In line with several other studies, we obtained strong correlations between the promptness measures for a subset of spectrotemporal ripples and speech understanding in quiet and in noise. The higher the promptness for a particular ripple, the better the measure for speech-in-noise perception (lower thresholds). Sentence and word recognition in quiet and in noise has been demonstrated before to correlate with spectral ripple discrimination thresholds in HA users (Davies-Venn et al., 2015; Shim et al., 2014) and CI users (Anderson et al., 2012; Berenstein et al., 2008; Drennan et al., 2014; Won et al., 2007). Temporal modulation detection thresholds also correlated with word scores of CI users (Won et al., 2011), but not in all studies (Drennan et al., 2015). Unlike Zhang et al. (2013) we did not find a relationship between bimodal benefit for speech understanding versus spectral ripple detection with the HA, possibly because our subjects had on average about 20 dB HL worse residual hearing. For hearing-impaired listeners also sentence recognition at 0 dB SNR has been shown to correlate significantly with ST modulated ripples, for 4 Hz with 1 and 2 c/o, and for 32 Hz with 0.5 and 1 c/o (Bernstein et al., 2013). We found the strongest correlations for ripples around 9 Hz and below 2 c/o, most likely an important region for speech understanding. For Dutch listeners, modulations of 8-10 Hz were indeed demonstrated to be critical for speech understanding (Drullman et al., 1994). For American English, Elliott and Theunissen (2009) showed that temporal modulations below 12 Hz are crucial for speech understanding and that most spectral modulations in speech are below 4 c/o.

CHAPTER 8

GENERAL DISCUSSION & SUMMARIES

GENERAL DISCUSSION

Several studies were performed in this thesis to investigate the benefit of a hearing aid (HA) in conjunction with a cochlear implant (CI) in contralateral ears. Below, the main findings of these studies and possible future implications are discussed.

Because validated fitting guidelines do not yet exist for bimodal devices, we investigated a loudness balancing procedure in three frequency bands (0 - 548, 548 - 1000 and >1000 Hz) in *Chapter 2*. Ideally, but impossible to achieve within clinically acceptable time, a balanced loudness would be needed for all frequencies. Furthermore, the question is how reliable and consistent subjective loudness judgments are. An alternative would be to use a more objective procedure, e.g. based on loudness growth models, aiming at normalized loudness perception, as proposed by Francart et al. (2012b). Either way, loudness balancing remains challenging, because of the extremely restricted dynamic range of hearing in both ears of typical bimodal users (Blamey et al., 2000). Still, the brain may be plastic enough to extract valid binaural difference cues, given the accurate sound localization results in *Chapter 4,* described for several subjects equipped with a CI and HA that were balanced in loudness in the three frequency bands.

In *Chapter 3* we succeeded to improve the bimodal benefit for speech understanding in noise by matching the characteristics of the automatic gain control (AGC) from the HA to the CI. Compared to the standard HA, matched AGC characteristics significantly improved speech understanding in single-talker noise up to 2 dB. This is a promising first step in the development of bimodal devices that are optimized for combined use. Listening comfort, and possibly binaural hearing, may be further improved by real-time synchronization of the compression and timing in the CI and HA. Ensuring equal gain at both ears at all times could avoid unnatural ILD distortions as a result of the head shadow effect (Van Wanrooij and Van Opstal 2004). In bilateral HA users, wireless ear-to-ear communication has already been shown to improve sound localization and questionnaire scores (Kreisman et al. 2010; LB 2008; Smith P 2008; Sockalingam R 2009).

We concluded that AGC matching possibly improved binaural processing, since we found improved bimodal benefit for speech understanding with noise presented from the (least favorable SNR) HA side. However, our sound localization study, described in *Chapter 4,* did not yield a clear correlation between bimodal benefit for speech understanding in noise and localization performance. Although several subjects with high bimodal benefit for speech understanding in noise also had a high localization sensitivity (evidenced by larger slopes in their stimulus-response relations), others performed poor in the localization task. This suggests that, besides binaural cues, other mechanisms might have contributed to the bimodal benefit for speech understanding in noise. These could, for example, include a better audibility of the fundamental frequency through the HA (Zhang et al., 2010), or the listener's ability to better fuse CI and HA input.

Accurate sound localization is only possible in the case of audible spectral overlap between both ears, as demonstrated in *Chapter 4*. We concluded that bimodal listeners must rely on weak low-frequency interaural level differences (ILDs) for horizontal sound localization, since signal processing in the CI destroys temporal fine-structure cues. However, low frequencies in natural listening conditions produce only small (often negligible) ILD cues. In an attempt to further optimize the audibility of high frequency ILD cues, we tested an experimental frequency-lowering algorithm in the HA, that continuously switched between the compression of consonants and preservation of vowels (*Chapter 5*). Although overall group effects were absent, two subjects did improve their sound localization performance, possibly because of the relatively low compression ratios in their device settings. Although this could also be related to the quality of their residual hearing, frequency lowering deserves further investigation. Less extreme compression settings, or other smart ways to increase the audibility of high frequencies (like in (Brown, 2014), or ILDs, could possibly enhance hearing performance in larger groups of bimodal users. A second argument that could support this expectation is that frequency lowering resulted in improved horizontal sound localization in four of our subjects. This suggests that the brain could, in principle, be endowed with sufficient plasticity to learn to integrate the different inputs from the CI and HA in order to process binaural cues.

In *Chapters 6 and 7* we used broadband dynamic spectral-temporal ripple stimuli to assess hearing performance in normal hearing- and bimodal listeners. Because of the reported strong correlations between ripple responses and speech perception, ripples may have great potential to provide a reliable objective measure of spectral-temporal hearing ability. In comparison with traditional measures of speech understanding, like percentage correct word scores, or signal-to-noise ratios at 50% performance, we think that ripple sensitivity could give a less time-consuming objective scores, that are less susceptible to procedural experience, or to cognitive and linguistic factors. Unlike adaptive modulation detection threshold procedures, the reaction time paradigm described in *Chapter 7* is a quick, sensitive, and easy test that could be readily implemented in a clinical setting as a diagnostic tool.

After participating in the studies described in this thesis, the majority of bimodal listeners kept using the HA that was provided for this research, contralateral to their CI. Still, we found large variations in individual benefit, ranging from -7 to +15 dB for speech understanding in noise (*Chapter 3*). Also, sound localization performance was generally poor for most subjects (*Chapter 4*). We found, however, that having an HA-aided hearing thresholds better than 45 dB HL was a prerequisite for accurate sound localization, although not all subjects with such good residual hearing could localize sounds. Also in speech tests, we did not find significant correlations between hearing performance and thresholds in the non-implanted ear, in line with other studies (Ching et al., 2005; Luntz et al., 2005). Taken together, the degree of residual hearing plays a role in the success rate of bimodal hearing, but other factors must be involved as well. These could possibly include, among others, the interaural pitch match performance with the CI alone, subjective motivation, or the adaptive capacity of the brain. A clear predictor for performance benefit is still not available.

In summary, the findings in this thesis encourage CI users with residual hearing in the opposite ear to use a HA, ideally with matched compression characteristics. Individual fine-tuning of the bimodal fitting, regarding loudness balancing and frequency compression, may further improve the bimodal benefit for speech understanding, or sound localization. When hearing loss in the non-implanted ear becomes too profound, bilateral cochlear implantation might be an option, but this topic falls beyond the scope of this thesis.

ENGLISH SUMMARY

Since inclusion criteria for cochlear implantation are expanding, more and more people with considerable residual hearing are implanted who can potentially benefit from a conventional hearing aid in the non-implanted ear. Benefits from such 'bimodal stimulation' can include improved speech understanding in noise, music and voice pitch perception, and sound-source localization. However, individual differences are large, CI and HA devices are often not technologically prepared for their combined use, and validated bimodal fitting guidelines do not yet exist. This thesis aimed to improve bimodal benefit, and possibly binaural cues, by adjusting the settings and algorithms in the HA, to create a better match with the CI. To this end, a group of 15-20 bimodal listeners participated in a large set of experiments, stretching out over two years. All participants used an Advanced Bionics CI processor in one ear, and a Phonak Naida S IX UP hearing aid in the other ear.

Suboptimal fittings and mismatches in loudness are possible explanations for the large individual differences seen in listeners using a CI in one ear and a HA in the other ear. In *Chapter 2* we investigated the effect of loudness balancing in separate frequency bands. The HA gain was adjusted for soft and loud input sounds in three frequency bands (0 - 548, 548 - 1000 and >1000 Hz) to match loudness with the CI. This procedure was compared to a simple broadband balancing procedure that reflected current clinical practice. In a three-visit cross-over design with four weeks between sessions, we tested speech understanding in quiet and in noise and administered questionnaires to also assess benefit experienced in the real world. After balancing, gains below 548 Hz were the same for both procedures. Only for 40 and 60 dB SPL input, HA gains in the middle and highest frequency band of the three-band fitting were significantly lower compared to broadband balancing. For speech in noise, a marginal bimodal benefit of 0.3 ± 4 dB was found, with large differences between subjects and spatial sound configurations. Speech understanding in quiet and in noise did not differ between the two loudness balancing procedures. On average, frequency-dependent loudness balancing did not lead to improved speech understanding in quiet and in noise as compared to broadband balancing. However, nine out of fifteen subjects showed significantly better speech understanding with either one of the fittings, suggesting the importance of individual HA fine-tuning.

In *Chapter 3* the three-band loudness balancing procedure was further optimized. In addition, we aimed to improve bimodal benefit by matching the number of compression channels and time constants of the automatic gain control (AGC) of the HA to the CI. Equivalent AGC was hypothesized to support a balanced loudness and improve bimodal benefit for dynamically changing signals like speech. In a three-visit cross-over design with four weeks between sessions, performance was measured using a HA with a standard AGC (syllabic multichannel compression with 1 ms attack time and 50 ms release time) or an AGC that was adjusted to match that of the CI processor (dual AGC broadband compression, 3 and 240 ms attack time, 80 and 1500 ms release time). Significant bimodal benefit over the CI alone was only found for the AGC-matched HA for speech understanding with single-talker noise. Compared to the standard HA, matched AGC

characteristics significantly improved speech understanding by 1.9 dB when noise was presented from the (least favorable SNR) HA side, which possibly suggests better binaural processing with matched AGCs. AGC matching increased bimodal benefit insignificantly by 0.6 dB when single-talker noise was presented from the CI implanted side or by 0.8 (single-talker noise) and 1.1 dB (stationary noise) in the more complex configurations with two simultaneous maskers. In questionnaires, subjects rated the AGC matched HA higher than the standard HA, for understanding one person in quiet and in noise, and for the quality of sounds. Listening to a slightly raised voice, subjects indicated increased listening comfort with matched AGCs. At the end of the study, nine out of fifteen subjects preferred to take home the AGC-matched HA, one preferred the standard HA and five subjects had no preference. These findings encourage the use of a CI processor and HA with matched AGC characteristics for bimodal hearing.

It often remains unclear if bimodal benefit arises from the complementary types of information from both devices or from real binaural cues, including interaural time (ITD) and level (ILD) differences. In *Chapter 4* we tested horizontal sound localization in bimodal listeners to determine the availability and contribution of binaural (ILDs, temporal fine structure and envelope ITDs) and monaural (loudness, spectral) cues, by systematically varying the frequency band, level and envelope of the stimuli. Results showed that the sound bandwidth had a strong effect on the localization behavior of bimodal listeners. Responses could be systematically changed by adjusting the frequency range of the stimulus, or by simply switching the HA and CI on and off. Localization responses were largely biased to one side, typically the CI side for broadband and high-pass filtered sounds, and occasionally to the HA side for low-pass filtered sounds. Overall, localization behavior remained poor. HA-aided thresholds better than 45 dB HL in the frequency range of the stimulus were a prerequisite, but not a guarantee, for the ability to indicate sound-source direction. We argued that bimodal sound localization is likely based on ILD cues, even at frequencies below 1500 Hz, for which the natural ILDs are small. These cues are typically perturbed in bimodal listeners, leading to a biased localization percept of sounds. The improved behavior of some listeners could result from a combination of sufficient spectral overlap and loudness balancing in bimodal hearing.

Residual hearing in the non-implanted ear is typically limited to the low frequencies, while natural ILD cues are more pronounced for high frequencies. The goal of *Chapter 5* was to improve the audibility of high-frequency ILDs by applying extreme frequency compression (FC) in the hearing aid (HA) of twelve bimodal listeners. An experimental signal-adaptive frequency-lowering algorithm was tested, compressing frequencies above 160 Hz into the audible range of residual hearing, but only for consonants, thus protecting vowel formants to preserve speech perception. In a cross-over design with at least five weeks of acclimatization between sessions, bimodal performance with and without FC was compared for horizontal sound localization, speech understanding in quiet and in noise, vowel, consonant and voice pitch perception. Overall, adaptive frequency lowering did not significantly affect bimodal listening performance, possibly because the adaptive switching preserved vowel perception but prevented adaptation to consistent ILD cues. Still, two subjects seemed to improve horizontal sound localization. Besides adaptive FC, four subjects also tested experimental frequency lowering applied to all incoming sounds, resulting in better sound localization performance, but worse speech understanding. After the study, four subjects preferred to use adaptive FC, four preferred not to use FC, and four had no preference. Noteworthy, the subjects preferring FC were the best bimodal listeners, performing superior in all psychophysical tasks, both with and without FC.

In current clinical and research practice, hearing performance is typically determined using traditional measures of speech understanding, including percentage words correct or speech reception thresholds. The study in *Chapter 6* investigated an alternative hearing test that could not be influenced by the subject's linguistic and cognitive skills. We used abstract 'ripple' stimuli, which are broadband noises with sinusoidally modulated amplitudes over time and/or frequency, with similar characteristics as in natural speech. Stimuli consisted of a static harmonic complex between 500 and 1500 ms, immediately followed by a dynamic ripple (modulation depth: 50%) with all possible combinations of 8 spectral (0-8 cyc/oct) and 17 temporal (±[0-64] Hz) modulation rates. Six normal-hearing subjects were instructed to push a button as soon as they heard the sound stimuli change from static to dynamic ripple. Stimuli were presented monaurally and binaurally under normal hearing conditions, and while using simulations of a HA and CI. From measured reaction times we constructed the spectro-temporal transfer function of each listener. Binaural and simulated bimodal responses could not be fully predicted from the monaural performance data, suggesting neural integration of binaural information. The joint spectro-temporal transfer function could be separated in a pure spectral and temporal component for the normal hearing conditions, but to a lesser extent for the simulated hearing-impaired responses. We concluded that reaction times provide a solid objective measure of spectral-temporal ripple sensitivity.

In *Chapter 7* the ripple test, developed in Chapter 6, was performed in a group of fourteen bimodal listeners. Ripples were presented in the bimodal and monaural CI- and HA conditions. On average, bimodal responses were comparable to responses in the monaural CI condition, which elicited faster reaction times than the monaural HA condition. Responses to temporal and spectral modulations followed a band-pass filter pattern with fastest responses around 0.25 cyc/oct and 9 Hz. Bimodal responses could not be predicted from the CI/HA-only performance data, suggesting neural integration of binaural information. A correlation analysis showed that ripple sensitivity is a strong performance indicator of speech understanding in quiet and in noise. The joint MTF was partially inseparable for frequency and time modulations, as the MTF could not be fully described by a pure spectral and temporal component. We also found that binaural performance was always slower than predicted by the benchmark race model of statistical facilitation, which assumes independent monaural processing channels.

NEDERLANDSE SAMENVATTING

Door de snelle technologische ontwikkelingen en positieve effecten op de spraak-perceptie, komen steeds meer slechthorenden in aanmerking voor een cochleair implantaat (CI). Een CI omzeilt het niet meer functionerende binnenoor en geeft geluiden in de vorm van elektrische signalen direct door aan de gehoorzenuw.

De groeiende populatie van CI-gebruikers heeft vaak nog redelijk restgehoor, waardoor een conventioneel gehoorapparaat in het niet-geïmplanteerde oor een extra meerwaarde kan geven. Deze combinatie van een CI en een contralateraal hoortoestel heet 'bimodale stimulatie' en kan een verbetering geven op de geluidskwaliteit, geluidslokalisatie en spraak verstaan in rumoerige omgevingen. Er is echter nog veel onduidelijk over het horen met twee geheel verschillende gehoorapparaten, en de resultaten variëren aanzienlijk tussen individuen. In dit proefschrift is onderzoek gedaan naar bimodaal horen, met als doel de meerwaarde van het hoortoestel te optimalizeren door een betere match te creëren met het CI. Hiervoor zijn een groot aantal experimenten gedaan bij een vaste groep van 15-20 bimodale gebruikers die bereid was om gedurende twee jaar lang mee te doen aan dit onderzoek. Alle participanten gebruikten een Advanced Bionics CI processor in het ene oor en een Phonak Naida S IX UP hoortoestel in het andere oor.

Omdat er nog geen standaardprocedure is voor het aanpassen van een hoortoestel bij CI gebruikers, was dit het uitgangspunt van de eerste studie, beschreven in ***Hoofdstuk 2***. Het doel was om een balans in luidheid te creëren tussen het CI en hoortoestel. Hiervoor vergeleken we twee verschillende procedures. Bij de simpele 'breedband' methode, die de klinische praktijk weerspiegelde, werd de luidheid gebalanceerd aan de hand van een spraak signaal op 65 dB SPL over het gehele frequentiespectrum. Dit vergeleken we met een zogenaamde 'drie-band' procedure waarin de luidheid tussen CI en hoortoestel werd afgeregeld voor zachte en harde geluiden in drie verschillende frequentiebanden (0 - 548, 548 – 1000, and >1000 Hz). Na een gewenningsperiode van vier weken testten we het spraak verstaan en de subjectieve beleving van bimodaal horen na breedband en drie-band balanceren, volgens een cross-over design. Beide procedures resulteerden in vergelijkbare versterkingen in het hoortoestel voor frequenties < 548 Hz. Voor de hogere frequenties was de versterking voor 40 en 60 dB SPL input lager bij de drie-band methode. Voor het spraak verstaan in stilte en in ruis werden geen verschillen gevonden tussen de twee manieren van luidheidsafregeling. Echter, op individueel niveau, waren vijf proefpersonen beter in het spraak verstaan in ruis met de breedband methode en vier met de drie-band methode. Dit geeft aan dat het individueel afregelen van de bimodale aanpassing mogelijk meer bimodaal voordeel kan geven.

In ***Hoofdstuk 3*** werd de drie-band procedure verder geoptimaliseerd. Daarnaast was het doel in deze studie om de luidheidsbalans te behouden voor fluctuerende geluiden zoals spraak. De compressie, oftewel *automatic gain control (AGC),* van een hoortoestel speelt hierbij een grote rol. Voor dit onderzoek werd het standaard Phonak hoortoestel (syllabische compressie, 1 ms inregeltijd, 50 ms uitregeltijd) voorgeprogrammeerd met dezelfde compressie-eigenschappen als de CI (dubbele AGC loop, 3 en 240 ms inregeltijd, 80 en 1500 ms uitregeltijd). In een cross-over design testten alle proefpersonen zowel het standaard hoortoestel als het aangepaste hoortoestel. Na een gewenningsperiode van vier weken werd het spraak verstaan in stilte en in ruis gemeten. Hierbij vonden we een verbetering van 1.9 dB in bimodale meerwaarde bij gebruik van het aangepaste hoortoestel voor het spraak verstaan met ruis van de hoortoestel zijde. Ook in vragenlijsten kwam het aangepaste hoortoestel als beste naar boven bij het spraak verstaan van één persoon in stilte en in achtergrondruis, alsmede de geluidskwaliteit. Na deze studie wilden negen van de vijftien proefpersonen doorgaan met het gebruik van het aangepaste hoortoestel, één proefpersoon prefereerde het standaard hoortoestel, de overige vier hadden geen voorkeur.

Het is vaak onduidelijk of een hoortoestel in combinatie met een CI meerwaarde geeft door de toevoeging van complementaire informatie, of dat het brein de input aan beide oren daadwerkelijk kan integreren voor bijvoorbeeld geluidslokalisatie. De richting van een geluid in het horizontale vlak wordt bepaald door kleine verschillen in tijd (voor frequenties < 1500 Hz) en luidheid (voor frequenties > 1500 Hz) tussen de twee oren, de zogeheten interaurale verschillen. De vraag in ***Hoofdstuk 4*** was of bimodale gebruikers deze verschillen kunnen waarnemen. Dit hebben we getest door proefpersonen in een donkere ruimte de richting van korte (150 ms) geluiden te laten aangeven. Uit de resultaten bleek dat de bandbreedte van de stimulus grote invloed had op het localisatie-gedrag van de bimodale proefpersonen. De meeste proefpersonen namen geluiden met lage frequenties waar aan de kant van het hoortoestel en hoge frequenties of breedband geluiden aan de kant van het CI. Slechts enkele proefpersonen met (geholpen) gehoordrempels < 45 dB HL in het oor met het hoortoestel waren in staat de juiste richting aan te geven. We concludeerden dat het richting horen bij bimodale stimulatie gebaseerd moet zijn op interaurale luidheidsverschillen, zelfs voor frequenties < 1500 Hz waar luidheidsverschillen tussen beide oren in de normale geluidswereld klein zijn.

Het restgehoor bij slechthorenden is vaak beperkt tot de lage frequenties, terwijl interaurale verschillen in luidheid met name ontstaan in de hoge frequenties. In ***Hoofdstuk 5*** werd gepoogd de hoge frequenties beter hoorbaar te maken in het hoortoestel, met als doel de waarneming van interaurale luidheidsverschillen, en dus het richting horen, te verbeteren. Hiervoor werd een nieuw algoritme in het hoortoestel getest dat frequenties boven 160 Hz comprimeerde tot het frequentiegebied van het restgehoor. Deze compressie werd alleen toegepast bij medeklinkers en niet bij klinkers, om het spraak verstaan te behouden. Na vijf weken gewenning testten we het richting horen en spraak verstaan met en zonder adaptieve frequentie compressie volgens een cross-over design. Slechts bij twee proefpersonen met relatief goed restgehoor en dus zwakke compressie werd een verbetering in het richting horen gevonden. Gemiddeld over proefpersonen werd er geen verbetering of verslechtering gevonden in het richting horen of spraak verstaan. Mogelijk kwam dit doordat het switchen tussen klinkers en medeklinkers geen consistente interaurale informatie gaf. Bij vier proefpersonen testten we ook frequentie compressie op alle inkomende geluiden (zowel klinkers als medeklinkers). Dit leek een verbetering te geven op het richting horen, maar een verslechtering op het spraak verstaan. Na de studie wilden vier proefpersonen adaptieve frequentie compressie blijven gebruiken, vier wilden dit niet, en de overige vier hadden geen voorkeur.

Bij hoortoestel aanpassingen en onderzoek is de mate van spraak verstaan vaak het belangrijkste uitgangspunt. Uitkomstmaten zijn meestal uitgedrukt in een percentage correct herhaalde woorden of een signaal-ruis-verhouding waarbij 50% wordt verstaan. Echter, naast het gehoor zelf kunnnen deze scores ook beïnvloed worden door linguïstische kennis, cognitief vermogen, en de beperkte woordenset van de test. In ***Hoofdstuk 6*** is onderzoek gedaan naar een alternatieve manier om het spraak verstaan te testen. Hierbij is gebruik gemaakt van abstracte *ripple stimuli*, non-linguïstische geluiden met modulaties over tijd en over frequentie, die vergelijkbaar zijn met gewone spraak. Stimuli startten met een statische ruis die na 500 – 1500 ms overging in een gemoduleerde ripple, met een combinatie van 8 verschillende spectrale (0-8 cyc.oct) en 17 temporele (±[0-64] Hz) modulatiesnelheden. Zes normaal-horende proefpersonen moesten op een knop drukken zodra ze de overgang van stationaire ruis naar gemoduleerde ripple hoorden. Stimuli werden binauraal en monauraal gepresenteerd, onder normale omstandigheden alsook via simulaties van een CI en hoortoestel. Uit de resultaten konden we concluderen dat de reactietijd voor ripple-detectie een betrouwbare en gemakkelijk te meten uitkomstmaat biedt voor ripple gevoeligheid, en daardoor voor spectrale en temporele modulatiegevoeligheid. De reactietijden in de binaurale conditie en bimodale simulatie konden niet voorspeld worden aan de hand van de monauraal verkregen data. De spectro-temporele modulatie overdrachtsfunctie kon opgesplitst worden in een puur spectrale en pure temporele component voor de normaal-horende condities, maar in mindere mate voor de simulaties van CI, HA en bimodale stimulatie.

In ***Hoofdstuk 7*** hebben we dezelfde ripple tests toegepast bij een groep bimodale proefpersonen. Ripple detectie werd gemeten in de bimodale conditie, alsook in de monaurale CI en hoortoestel conditie. We testten ripples met een combinatie van zeven verschillende spectrale modulatiesnelheden en veertien temporele modulatiesnelheden. Reactietijden voor deze spectro-temporeel gemoduleerde ripples bleken niet simpelweg separeerbaar in een spectrale en temporele component. Bovendien konden de bimodale reactietijden niet simpelweg afgeleid worden uit de monaurale reactietijden, wat duidt op binaurale integratie van de input aan beide oren. De spectro-temporele modulatie overdrachtsfunctie kon niet geheel voorspeld worden uit een puur spectrale en puur temoporele component. Daarnaast vonden we sterke correlaties tussen de reactietijd van ripple detectie en de resultaten van spraak verstaan uit de vorige hoofdstukken. Snellere reactietijden correleerden met hogere percentages spraak verstaan, en met lagere (betere) signaal-ruis-verhoudingen voor spraak verstaan in achtergrondruis. Dit betekent dat ripples een representatieve maat zijn voor de mate van spraak-perceptie, zonder de nadelen van klassieke tests. Dit kan mogelijk interessant zijn voor de ontwikkeling van een snelle klinische performance-test.

LIST OF ABBREVATIONS

AGC	automatic gain control
AM	amplitude modulation
BB	broadband
BM	basilar membrane
CDF	cumulative distribution function
CI	cochlear implant
CR	compression ratio
DAI	direct audio input
EAS	electroacoustic stimulation
F0	fundamental frequency
FC	frequency compression
HA	hearing aid
HL	hearing loss
HSE	head shadow effect
HP	high-pass filtered noise
ICRA	International Collegium of Rehabilitative Audiology
IFFM	international female fluctuating masker
IHC	inner hair cells
ILD	interaural level difference
ISTS	international speech test signal
ITD	interaural time difference
IPD	interaural phase difference
LP	low-pass filtered noise
LIST	Leuven intelligibility sentence test
LMM	linear mixed model
NH	normal hearing
MTF	modulation transfer function
MAE	mean absolute error
NVA	Nederlanse Vereniging voor Audiologie

OHC	outer hair cells
RAU	rationalized arcsine units
RMS	root mean square
RT	reaction time
S0NCI	speech from front and noise from CI side
S0NHA	speech from front and noise from HA side
S0N0	speech and noise from front
S0N±90	speech from front and noise from both sides
SNR	signal to noise ratio
SNR50	signal to noise ratio at 50% speech reception threshold
SRM	spatial release from masking
SRT	speech reception threshold
SSQ	Speech, Spatial and Qualities of hearing scale
ST	spectro-temporal
STMTF	spectro-temporal modulation transfer function
SVD	singular value decomposition

DANKWOORD

Een grote groep mensen heeft mij geholpen bij het schrijven van dit proefschrift. Dank aan iedereen en in het bijzonder aan onderstaande collega's en vrienden.

Allereerst dank aan mijn promotoren John en Ad. Bedankt voor jullie hulp en kritische blik bij het opzetten, uitvoeren en vastleggen van dit onderzoek. Ondanks een drukke agenda waren jullie altijd beschikbaar voor overleg en adviezen zodra dat nodig was. Het kostte soms wat moeite om de tussenweg tussen fundamenteel onderzoek en klinische relevantie te vinden, maar dat is zeker gelukt. Ad, ik vond het leuk dat je wekelijks een dag gewoon tussen de PhD-studenten op onze kamer zat en zo makkelijk bereikbaar was voor vragen. John, speciale dank voor de eindsprint afgelopen jaar bij het afronden van mijn proefschrift.

Lucas, vaak stond ik versteld van de tijd die je kon vrijmaken voor mijn onderzoek. Bedankt voor het kritisch doorlezen van alle papers en meet-protocollen, je klinische en wetenschappelijke kennis, begeleiding bij het instellen en doormeten van hoortoestellen, uren hulp in de 'dode kamer' met meetopstellingen en proefpersonen en de 'roadtrip' tussen San Francisco en Lake Tahoe.

Marc, bedankt voor je extreem kritische blik bij het schrijven van de verschillende hoofdstukken, het opzetten van metingen en nieuwe ideeën wat betreft data-analyse en figuren. Fijn dat je maar op twee meter afstand zat voor al mijn vragen, ik heb veel van je geleerd!

Thanks to Sonova (Advanced Bionics & Phonak hearing systems) for the financial support of my research. Sepp, I am very thankful for your advices, help and always-quick responses. You have contributed a lot to this research by providing us with interesting and innovative ideas, special fitting software and the possibility to perform applied research by equipping hearing aids with experimental firmware. Many thanks also to Paddy; I will never forget that day you helped me calibrating the Harmony processor.

Martijn, jij was vaak de schakel tussen het fundamentele onderzoek en de klinische praktijk, bedankt voor al je hulp en adviezen, met name bij het richtinghoren.

Günter, Ruurd en Chris-Jan, bedankt voor jullie technische kennis en ondersteuning, de vele (spoed-)reparaties aan de lokalisatie-boog, jullie gezelschap tijdens de lunch, de juiste kabeltjes en adapters, audio-aansluitingen en het solderen van een splitter die cruciaal was voor hoofdstuk 2.

Dank aan alle proefpersonen die tientallen keren op en neer zijn gekomen voor alle onderzoeken en bereid waren om fijne en minder fijne hoortoestel-instellingen uit te proberen. Ondanks opsluitingen in de kelder, klapbanden op de snelweg, uren in de 'dode kamer' en saaie tests om het spraakverstaan in ruis (a.k.a. Zweedse borrelpraat) te meten, kwamen jullie toch altijd weer opdagen bij de volgende meting. Ik ben blij dat we op die manier een hoop waardevolle data hebben kunnen verzamelen! Dankzij jullie zijn er vijf mooie onderzoeken gepubliceerd en heeft Phonak een nieuwe bimodale fitting

kunnen doorvoeren.

Leden van de manuscriptcommissie, dank voor jullie vlotte beoordeling van dit proefschrift.

Zonder hulp van twee goede stagiaires had ik een aantal hoofdstukken van dit proefschrift nooit kunnen afronden. Daisy, bedankt voor alle uren die je hebt doorgebracht zonder daglicht, in het lab onder de grond, tijdens het meten van de ripples voor Hoofdstuk 6. Maartje, bedankt voor een heel jaar lang gezelschap, de vele richting-hoor-metingen en uitgebreide analyses. Leuk dat je in de audiologische hoek bent gebleven!

Thanks to my room mates and all the other people in Wing 8 of the Huygens building. Yoollah, you were there since day one of my PhD trajectory, for good company and stories, thanks for that. Too bad you moved all the way to the UK! Rachel, besides hearing problems, I learned a lot too during the past four years about the meat industry ;). I'm happy we were neighbors for almost the entire four years. Thanks also for organizing a lot of enjoyable game nights! Snandan and Seba, thank you for the bimodal and clinical collaborations, I'm happy you will continue the CI-related research! To the rest, Laurens, Baha, Bahram, Ahmed, Elahe, Katharina, Jose, Marije, Luuk and Guus, thanks for your company during lunch-time and good luck finishing your own thesis!

Dank aan Judith, Gea, en Marie-Louise voor alle organisatorische en administratieve hulp, waaronder het onderscheppen en versturen van vele pakketjes met hoortoestellen.

Dank aan de mensen van de afdeling Klinische Neurofysiologie, Britt-Sofie, Bart van de Warrenburg en Dick Stegeman, die mij enthousiast gemaakt hebben voor onderzoek.

Aan alle vrienden, oud-studiegenoten en roeiers uit Eindhoven dank voor alle afleiding de afgelopen vier jaar. Ik ben blij dat we af en toe de tijd konden vinden om elkaar op te zoeken, ondanks dat ik in het oosten in plaats van het zuiden van het land vertoefde. De bezoekjes aan Aruba, Boekel, de Effenaar en de Efteling hebben me er zeker doorheen gesleept ;). Ook van-A-tot-Zink bedankt voor alle uitjes, van ploeghappen tot carnaval. Speciale dank aan mijn paranimfen, Nadine en Carin, ik ben blij dat jullie naast mij staan tijdens de verdediging!

Tenslotte, dank aan mijn familie, ouders en zusjes, voor jullie plotselinge interesse in gehoorapparaten en het regelmatig informeren naar de voortgang van mijn onderzoek. Mooi dat we the International Hearing Aid Conference konden combineren met een ouderwets geslaagde familie-vakantie in Mexico!

Roel(ie), bedankt voor je dagelijkse belangstelling en voor de splitters uit je oude DJ set die eindelijk een keer van pas kwamen. Ik ben heel blij dat we gaan trouwen terwijl dit boekje gedrukt wordt! Wel jammer dat daarmee de echte auteur van alle papers mevrouw Castermans wordt terwijl ene Veugen vermeld staat ;)

CURRICULUM VITAE

Lidwien Veugen was born on December 13th 1987 in Waalre, the Netherlands. She completed her secondary education (Gymnasium-Beta) in 2006 at the *Van Maerlant Lyceum* in Eindhoven. Subsequently, Lidwien decided to study *Biomedical Engineering* at the *University of Technology* in Eindhoven. During her master *Medical Engineering* she became interested in research with a clinical application. She did a three-month internship on medical image registration at the *Brigham & Women's Hospital,* affiliated with *Harvard Medical School* in Boston, USA. For her Master's internship she performed research on transcranial magnetic stimulation in patients with writer's cramp at the *Neurology department* of the *Radboud University Nijmegen Medical Centre*. After her graduation, she started as a research assistant at the same department in the end of 2011. From May 2012 to 2016 Lidwien did her PhD research about 'bimodal cochlear implant fitting' at the *Biophysics department* of the *Radboud University Nijmegen*, in collaboration with the *department of Otorhinolaryngology* at the *Radboudumc* and *Advanced Bionics/Phonak Hearing Systems*. In June 2016 she started working on the edge of research and innovative software applications at *Eaglescience Software BV* in Amsterdam.

LIST OF PUBLICATIONS

Veugen, L.C., Chalupper, J., Mens, L.H., Snik, A.F. and van Opstal, A.J., 2017. Effect of extreme adaptive frequency compression in bimodal listeners on sound localization and speech perception. Cochlear Implants International, pp.1-12.

Veugen, L.C., Hendrikse, M.M., van Wanrooij, M.M., Agterberg, M.J., Chalupper, J., Mens, L.H., Snik, A.F. and van Opstal, A.J., 2016. Horizontal sound localization in cochlear implant users with a contralateral hearing aid. Hearing research, 336, pp.72-82.

Veugen, L.C., Chalupper, J., Snik, A.F., van Opstal, A.J. and Mens, L.H., 2016. Matching automatic gain control across devices in bimodal cochlear implant users. Ear and hearing, 37(3), pp.260-270.

Veugen, L.C., Chalupper, J., Snik, A.F., van Opstal, A.J. and Mens, L.H., 2016. Frequency-dependent loudness balancing in bimodal cochlear implant users. Acta oto-laryngologica, 136(8), pp.775-781.

Hoffland, B.S., Veugen, L.C., Janssen, M.M.H.P., Pasman, J.W., Weerdesteyn, V. and van de Warrenburg, B.P., 2014. A gait paradigm reveals different patterns of abnormal cerebellar motor learning in primary focal dystonias. The Cerebellum, 13(6), pp.760-766.

Delnooz, C.C.S., Veugen, L.C., Pasman, J.W., Lapatki, B.G., Dijk, J.P. and Warrenburg, B.P.C., 2014. The clinical utility of botulinum toxin injections targeted at the motor endplate zone in cervical dystonia. European journal of neurology, 21(12), p.1486.

Janssen, S., Veugen, L.C., Hoffland, B.S., Kassavetis, P., van Rooijen, D.E., Stegeman, D.F., Edwards, M.J., van Hilten, J.J. and van de Warrenburg, B.P., 2014. Normal eyeblink classical conditioning in patients with fixed dystonia. Experimental brain research, 232(6), pp.1805-1809.

Veugen, L.C., Hoffland, B.S., Stegeman, D.F. and van de Warrenburg, B.P., 2013. Inhibition of the dorsal premotor cortex does not repair surround inhibition in writer's cramp patients. Experimental brain research, 225(1), pp.85-92.

8

REFERENCES

Agterberg, M.J., Snik, A.F., Hol, M.K., Van Wanrooij, M.M., Van Opstal, A.J. 2012. Contribution of monaural and binaural cues to sound localization in listeners with acquired unilateral conductive hearing loss: improved directional hearing with a bone-conduction device. Hearing research 286, 9-18.

Agterberg, M.J., Snik, A.F., Hol, M.K., van Esch, T.E., Cremers, C.W., Van Wanrooij, M.M., Van Opstal, A.J. 2011. Improved horizontal directional hearing in bone conduction device users with acquired unilateral conductive hearing loss. Journal of the Association for Research in Otolaryngology : JARO 12, 1-11.

Alexander, J.M. 2013. Individual variability in recognition of frequency-lowered speech. Seminars in Hearing 34, 86-109.

Anderson, E.S., Oxenham, A.J., Nelson, P.B., Nelson, D.A. 2012. Assessing the role of spectral and intensity cues in spectral ripple detection and discrimination in cochlear-implant users. The Journal of the Acoustical Society of America 132, 3925-34.

Armstrong, M., Pegg, P., James, C., Blamey, P. 1997. Speech perception in noise with implant and hearing aid. The American journal of otology 18, S140-1.

Bacon, S.P., Viemeister, N.F. 1985. Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners. Audiology : official organ of the International Society of Audiology 24, 117-34.

Bacon, S.P., Gleitman, R.M. 1992. Modulation detection in subjects with relatively flat hearing losses. J Speech Hear Res 35, 642-53.

Baer, T., Moore, B.C. 1994. Effects of spectral smearing on the intelligibility of sentences in the presence of interfering speech. The Journal of the Acoustical Society of America 95, 2277-80.

Beijen, J., Snik, A.F., Straatman, L.V., Mylanus, E.A., Mens, L.H. 2010. Sound localization and binaural hearing in children with a hearing aid and a cochlear implant. Audiol Neurootol 15, 36-43.

Berenstein, C.K., Mens, L.H., Mulder, J.J., Vanpoucke, F.J. 2008. Current steering and current focusing in cochlear implants: comparison of monopolar, tripolar, and virtual channel electrode configurations. Ear and hearing 29, 250-60.

Bernstein, J.G.W., Mehraei, G., Shamma, S., Gallun, F.J., Theodoroff, S.M., Leek, M.R. 2013. Spectrotemporal Modulation Sensitivity as a Predictor of Speech Intelligibility for Hearing-Impaired Listeners. Journal of the American Academy of Audiology 24, 293-306.

Bernstein, L.R., Green, D.M. 1987. Detection of simple and complex changes of spectral shape. The Journal of the Acoustical Society of America 82, 1587-92.

Blamey, P.J., Dooley, G.J., James, C.J., Parisi, E.S. 2000. Monaural and binaural loudness measures in cochlear implant users with contralateral residual hearing. Ear and hearing 21, 6-17.

Blauert, J. 1997. Spatial Hearing–The Psychophysics of Human Sound Localization. MIT Press, Cambridge, MA.

Boersma, P., Weenink, D. 2001. Praat, a system for doing phonetics by computer. Glot International 5(9/10), 341-345.

Bohnert, A., Nyffeler, M., Keilmann, A. 2010. Advantages of a non-linear frequency compression algorithm in noise. European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery 267, 1045-53.

Bosman, A.J., Smoorenburg, G.F. 1992. Woordenlijst voor spraakaudiometrie [compact disc]. Nederlandse Vereniging voor Audiologie. Doesburg NL: Electro Medical Instruments b.v. & Gouda NL: Veenhuis Medical Audio b.v.

Boyle, P.J., Buchner, A., Stone, M.A., Lenarz, T., Moore, B.C. 2009. Comparison of dual-time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants. International journal of audiology 48, 211-21.

Brand, T., Kollmeier, B. 2002. Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. The Journal of the Acoustical Society of America 111, 2801-10.

Bremen, P., van Wanrooij, M.M., van Opstal, A.J. 2010. Pinna cues determine orienting response modes to synchronous sounds in elevation. The Journal of neuroscience : the official journal of the Society for Neuroscience 30, 194-204.

Brokx, J.P.L., Nooteboom, S.G. 1982. Intonation and the Perceptual Separation of Simultaneous Voices. J Phonetics 10, 23-36.

Brown, C.A. 2014. Binaural enhancement for bilateral cochlear implant users. Ear and hearing 35, 580-4.

Brown, K.D., Balkany, T.J. 2007. Benefits of bilateral cochlear implantation: a review. Current opinion in otolaryngology & head and neck surgery 15, 315-8.

Carpenter, R.H., Reddi, B.A., Anderson, A.J. 2009. A simple two-stage model predicts response time distributions. The Journal of physiology 587, 4051-62.

Chalupper, J., Geissler, G., Büchner, A., Battmer, R. 2014. Tolerable processing delay in electro-acoustic stimulation. Poster presented at the 13th Conference on Cochlear Implants and Other Implantable Auditory Technologies, Munich, Germany.

Chalupper, J., Agrawal, S., Fredelake, S., Spahr, T. 2015. Rationale and Implementation of a Bimodal Fitting Formula. Figshare: http://dx.doi.org/10.6084/m9.figshare.1340020.

Chi, T., Gao, Y., Guyton, M.C., Ru, P., Shamma, S. 1999. Spectro-temporal modulation transfer functions and speech intelligibility. The Journal of the Acoustical Society of America 106, 2719-32.

Ching, T.Y., Incerti, P., Hill, M. 2004. Binaural benefits for adults who use hearing aids and cochlear implants in opposite ears. Ear and hearing 25, 9-21.

Ching, T.Y., van Wanrooy, E., Dillon, H. 2007. Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: a review. Trends in amplification 11, 161-92.

Ching, T.Y., Dillon, H., Katsch, R., Byrne, D. 2001a. Maximizing effective audibility in hearing aid fitting. Ear and hearing 22, 212-24.

Ching, T.Y., Psarros, C., Hill, M., Dillon, H., Incerti, P. 2001b. Should children who use cochlear implants wear hearing aids in the opposite ear? Ear and hearing 22, 365-380.

Ching, T.Y., Hill, M., Brew, J., Incerti, P., Priolo, S., Rushbrook, E., Forsythe, L. 2005. The effect of auditory experience on speech perception, localization, and functional performance of children who use a cochlear implant and a hearing aid in opposite ears. International journal of audiology 44, 677-90.

Corneil, B.D., Van Wanrooij, M., Munoz, D.P., Van Opstal, A.J. 2002. Auditory-visual interactions subserving goal-directed saccades in a complex scene. J Neurophysiol 88, 438-54.

Cullington, H.E., Zeng, F.G. 2010. Bimodal hearing benefit for speech recognition with competing voice in cochlear implant subject with normal hearing in contralateral ear. Ear and hearing 31, 70-3.

Cullington, H.E., Zeng, F.G. 2011. Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification. Ear and hearing 32, 16-30.

Davidson, L.S., Firszt, J.B., Brenner, C., Cadieux, J.H. 2015. Evaluation of hearing aid frequency response fittings in pediatric and young adult bimodal recipients. Journal of the American Academy of Audiology 26, 393-407.

Davies-Venn, E., Nelson, P., Souza, P. 2015. Comparing auditory filter bandwidths, spectral ripple modulation detection, spectral ripple discrimination, and speech recognition: Normal and impaired hearing. The Journal of the Acoustical Society of America 138, 492-503.

Depireux, D.A., Simon, J.Z., Klein, D.J., Shamma, S.A. 2001. Spectro-temporal response field characterization with dynamic ripples in ferret primary auditory cortex. Journal of neurophysiology 85, 1220-34.

Dillon, H. 2001. Hearing Aids. Sydney Boomerang Press:159–186

Dorman, M., Spahr, A., Gifford, R.H., Cook, S., Zhang, T., Loiselle, L., Whittingham, J., Schramm, D. 2012.

Bilateral and bimodal benefit as a function of age for adults. Journal of hearing science 2, 37-39.

Dorman, M.F., Gifford, R.H., Spahr, A.J., McKarns, S.A. 2008. The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. Audiology & neuro-otology 13, 105-12.

Dorman, M.F., Loizou, P., Wang, S., Zhang, T., Spahr, A., Loiselle, L., Cook, S. 2014a. Bimodal cochlear implants: the role of acoustic signal level in determining speech perception benefit. Audiology & neuro-otology 19, 234-8.

Dorman, M.F., Loiselle, L., Stohl, J., Yost, W.A., Spahr, A., Brown, C., Cook, S. 2014b. Interaural level differences and sound source localization for bilateral cochlear implant patients. Ear and hearing 35, 633-40.

Dorman, M.F., Zeitler, D., Cook, S.J., Loiselle, L., Yost, W.A., Wanna, G.B., Gifford, R.H. 2015. Interaural level difference cues determine sound source localization by single-sided deaf patients fit with a cochlear implant. Audiol Neurootol 20, 183-8.

Drennan, W.R., Anderson, E.S., Won, J.H., Rubinstein, J.T. 2014. Validation of a clinical assessment of spectral-ripple resolution for cochlear implant users. Ear and hearing 35, e92-8.

Drennan, W.R., Won, J.H., Timme, A.O., Rubinstein, J.T. 2015. Nonlinguistic Outcome Measures in Adult Cochlear Implant Users Over the First Year of Implantation. Ear and hearing.

Dreschler, W.A., Verschuure, H., Ludvigsen, C., Westermann, S. 2001. ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology. Audiology : official organ of the International Society of Audiology 40, 148-57.

Drullman, R., Festen, J.M., Plomp, R. 1994. Effect of reducing slow temporal modulations on speech reception. The Journal of the Acoustical Society of America 95, 2670-80.

Dunn, C.C., Tyler, R.S., Witt, S.A. 2005. Benefit of wearing a hearing aid on the unimplanted ear in adult users of a cochlear implant. Journal of speech, language, and hearing research : JSLHR 48, 668-80.

Dunn, C.C., Perreau, A., Gantz, B., Tyler, R.S. 2010. Benefits of localization and speech perception with multiple noise sources in listeners with a short-electrode cochlear implant. Journal of the American Academy of Audiology 21, 44-51.

Eddins, D.A., Bero, E.M. 2007. Spectral modulation detection as a function of modulation frequency, carrier bandwidth, and carrier frequency region. Journal of the Acoustical Society of America 121, 363-372.

EHIMA. Retrieved 17 April, 2013. European Hearing Instrument Manufacturers Association. www.ehima.com/ehima2.

Elliott, T.M., Theunissen, F.E. 2009. The Modulation Transfer Function for Speech Intelligibility. PLoS Comput Biol 5.

Firszt, J.B., Reeder, R.M., Skinner, M.W. 2008. Restoring hearing symmetry with two cochlear implants or one cochlear implant and a contralateral hearing aid. Journal of rehabilitation research and development 45, 749-67.

Fitzpatrick, E.M., Leblanc, S. 2010. Exploring the factors influencing discontinued hearing aid use in patients with unilateral cochlear implants. Trends in amplification 14, 199-210.

Francart, T., Wouters, J. 2007. Perception of across-frequency interaural level differences. J Acoust Soc Am 122, 2826-31.

Francart, T., McDermott, H. 2012a. Speech perception and localisation with SCORE bimodal: a loudness normalisation strategy for combined cochlear implant and hearing aid stimulation. PloS one 7, e45385.

Francart, T., McDermott, H.J. 2012b. Development of a loudness normalisation strategy for combined cochlear implant and acoustic stimulation. Hearing research 294, 114-24.

Francart, T., McDermott, H.J. 2013. Psychophysics, fitting, and signal processing for combined hearing aid and cochlear implant stimulation. Ear and hearing 34, 685-700.

Francart, T., Brokx, J., Wouters, J. 2008. Sensitivity to interaural level difference and loudness growth with bilateral bimodal stimulation. Audiology & neuro-otology 13, 309-19.

Francart, T., Brokx, J., Wouters, J. 2009. Sensitivity to interaural time differences with combined cochlear implant and acoustic stimulation. Journal of the Association for Research in Otolaryngology : JARO 10, 131-41.

Francart, T., Lenssen, A., Wouters, J. 2011. Sensitivity of Bimodal Listeners to Interaural Time Differences with Modulated Single- and Multiple-Channel Stimuli. Audiol Neuro-Otol 16, 82-92.

Fu, Q.J. 2002. Temporal processing and speech recognition in cochlear implant users. Neuroreport 13, 1635-1639.

Gatehouse, S., Noble, W. 2004. The Speech, Spatial and Qualities of Hearing Scale (SSQ). International journal of audiology 43, 85-99.

Gatehouse, S., Akeroyd, M. 2006. Two-eared listening in dynamic situations. International journal of audiology 45 Suppl 1, S120-4.

Gielen, S.C., Schmidt, R.A., Van den Heuvel, P.J. 1983a. On the nature of intersensory facilitation of reaction time. Perception & Psychophysics 34, 161-8.

Gielen, S.C., Schmidt, R.A., Van den Heuvel, P.J. 1983b. On the nature of intersensory facilitation of reaction time. Perception & psychophysics 34, 161-8.

Gifford, R.H. 2011. Who is a cochlear implant candidate? Hearing Journal 64, 16-22.

Gifford, R.H., Dorman, M.F., Sheffield, S.W., Teece, K., Olund, A.P. 2014. Availability of binaural cues for bilateral implant recipients and bimodal listeners with and without preserved hearing in the implanted ear. Audiology & neuro-otology 19, 57-71.

Glasberg, B.R., Moore, B.C. 1986. Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. The Journal of the Acoustical Society of America 79, 1020-33.

Glista, D., Scollie, S., Bagatto, M., Seewald, R., Parsa, V., Johnson, A. 2009. Evaluation of nonlinear frequency compression: Clinical outcomes. International journal of audiology 48, 632-644.

Golub, J.S., Won, J.H., Drennan, W.R., Worman, T.D., Rubinstein, J.T. 2012. Spectral and temporal measures in hybrid cochlear implant users: on the mechanism of electroacoustic hearing benefits. Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology 33, 147-53.

Goupell, M.J., Kan, A., Litovsky, R.Y. 2013a. Mapping procedures can produce non-centered auditory images in bilateral cochlear implantees. Journal of the Acoustical Society of America 133, El101-El107.

Goupell, M.J., Stoelb, C., Kan, A., Litovsky, R.Y. 2013b. Effect of mismatched place-of-stimulation on the salience of binaural cues in conditions that simulate bilateral cochlear-implant listening. The Journal of the Acoustical Society of America 133, 2272-87.

Grantham, D.W., Ashmead, D.H., Ricketts, T.A., Labadie, R.F., Haynes, D.S. 2007. Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. Ear and hearing 28, 524-41.

Henning, G.B. 1974. Detectability of interaural delay in high-frequency complex waveforms. The Journal of the Acoustical Society of America 55, 84-90.

Henry, B.A., Turner, C.W. 2003. The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. The Journal of the Acoustical Society of America 113, 2861-73.

Henry, B.A., Turner, C.W., Behrens, A. 2005. Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. The Journal of the Acoustical Society of America 118, 1111-21.

Hershenson, M. 1962. Reaction-Time as a Measure of Intersensory Facilitation. J Exp Psychol 63, 289-&.

Hofman, P.M., Van Riswick, J.G.A., Van Opstal, A.J. 1998. Relearning sound localization with new ears. Nature Neuroscience 1, 417-421.

Holube, I., Fredelake, S., Vlaming, M., Kollmeier, B. 2010. Development and analysis of an International Speech Test Signal (ISTS). International journal of audiology 49, 891-903.

Hoth, S. 2007. Indication for the need of flexible and frequency specific mapping functions in cochlear implant speech processors. European archives of oto-rhino-laryngology : official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS) : affiliated with the German Society for

Oto-Rhino-Laryngology - Head and Neck Surgery 264, 129-38.

Houben, R., Koopman, J., Luts, H., Wagener, K.C., van Wieringen, A., Verschuure, H., Dreschler, W.A. 2014. Development of a Dutch matrix sentence test to assess speech intelligibility in noise. International journal of audiology 53, 760-3.

House, W.F., K.I., B. 1991. Cochlear implants: From idea to clinical practice. Cochlear Implants: A practical guide, ed. H. Cooper. London: Whurr Publishers Litd.

Hua, H., Johansson, B., Jonsson, R., Magnusson, L. 2012. Cochlear implant combined with a linear frequency transposing hearing aid. Journal of the American Academy of Audiology 23, 722-32.

Ibrahim I, P.V., Macpherson E, Cheesman M. 2012. Evaluation of speech intelligibility and sound localization abilities with hearing aids using binaural wireless technology. Audiology Research 3.

Kan, A., Litovsky, R.Y., Goupell, M.J. 2015. Effects of interaural pitch matching and auditory image centering on binaural sensitivity in cochlear implant users. Ear and hearing 36, e62-8.

Kewley-Port, D., Burkle, T.Z., Lee, J.H. 2007. Contribution of consonant versus vowel information to sentence intelligibility for young normal-hearing and elderly hearing-impaired listeners. The Journal of the Acoustical Society of America 122, 2365-75.

Kong, Y.Y., Stickney, G.S., Zeng, F.G. 2005. Speech and melody recognition in binaurally combined acoustic and electric hearing. Journal of the Acoustical Society of America 117, 1351-1361.

Kreisman, B.M., Mazevski, A.G., Schum, D.J., Sockalingam, R. 2010. Improvements in speech understanding with wireless binaural broadband digital hearing instruments in adults with sensorineural hearing loss. Trends in amplification 14, 3-11.

Krueger, B., Joseph, G., Rost, U., Strauss-Schier, A., Lenarz, T., Buechner, A. 2008. Performance groups in adult cochlear implant users: speech perception results from 1984 until today. Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology 29, 509-12.

Kumpik, D.P., Oliver Kacelnik, and Andrew J. King. 2010. Adaptive reweighting of auditory localization cues in response to chronic unilateral earplugging in humans. The Journal of Neuroscience 30, 4883-4894.

Laszig, R., Aschendorff, A., Stecker, M., Muller-Deile, J., Maune, S., Dillier, N., Weber, B., Hey, M., Begall, K., Lenarz, T., Battmer, R.D., Bohm, M., Steffens, T., Strutz, J., Linder, T., Probst, R., Allum, J., Westhofen, M., Doering, W. 2004. Benefits of bilateral electrical stimulation with the nucleus cochlear implant in adults: 6-month postoperative results. Otology & neurotology : official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology 25, 958-68.

LB, H. 2008. Epoq studymeasures user benefit. Hear Jour 61, 47-49.

Li, N., Loizou, P.C. 2008. A glimpsing account for the benefit of simulated combined acoustic and electric hearing. The Journal of the Acoustical Society of America 123, 2287-94.

Liberman, A.M. 1997. Speech: a special code Perceptual and Motor Skills. MIT Press, Cambridge, MA 85, 383-383.

Litovsky, R.Y., Johnstone, P.M., Godar, S.P. 2006. Benefits of bilateral cochlear implants and/or hearing aids in children. International journal of audiology 45 Suppl 1, S78-91.

Litovsky, R.Y., Parkinson, A., Arcaroli, J. 2009. Spatial hearing and speech intelligibility in bilateral cochlear implant users. Ear and hearing 30, 419-31.

Litvak, L.M., Spahr, A.J., Saoji, A.A., Fridman, G.Y. 2007. Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. Journal of the Acoustical Society of America 122, 982-991.

Loizou, P.C., Dorman, M., Tu, Z. 1999. On the number of channels needed to understand speech. The Journal of the Acoustical Society of America 106, 2097-103.

Luntz, M., Shpak, T., Weiss, H. 2005. Binaural-bimodal hearing: concomitant use of a unilateral cochlear implant and a contralateral hearing aid. Acta oto-laryngologica 125, 863-9.

Macpherson, E.A., Middlebrooks, J.C. 2002. Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. The Journal of the Acoustical Society of America 111, 2219-36.

Massoudi, R., Van Wanrooij, M.M., Van Wetter, S.M., Versnel, H., Van Opstal, A.J. 2013. Stable bottom-up processing during dynamic top-down modulations in monkey auditory cortex. The European journal of neuroscience 37, 1830-42.

Massoudi, R., Van Wanrooij, M.M., Van Wetter, S.M., Versnel, H., Van Opstal, A.J. 2014. Task-related preparatory modulations multiply with acoustic processing in monkey auditory cortex. Eur J Neurosci 39, 1538-50.

McCreery, R.W., Brennan, M.A., Hoover, B., Kopun, J., Stelmachowicz, P.G. 2013. Maximizing Audibility and Speech Recognition With Nonlinear Frequency Compression by Estimating Audible Bandwidth. Ear and hearing 34, E24-E27.

McDermott, H., Henshall, K. 2010. The use of frequency compression by cochlear implant recipients with postoperative acoustic hearing. Journal of the American Academy of Audiology 21, 380-9.

Mehraei, G., Gallun, F.J., Leek, M.R., Bernstein, J.G.W. 2014. Spectrotemporal modulation sensitivity for hearing-impaired listeners: Dependence on carrier center frequency and the relationship to speech intelligibility. Journal of the Acoustical Society of America 136, 301-316.

Miller, J. 1982. Divided attention: evidence for coactivation with redundant signals. Cognitive psychology 14, 247-79.

Mok, M., Grayden, D., Dowell, R.C., Lawrence, D. 2006. Speech perception for adults who use hearing aids in conjunction with cochlear implants in opposite ears. J Speech Lang Hear Res 49, 338-51.

Monaghan, J.J., Krumbholz, K., Seeber, B.U. 2013. Factors affecting the use of envelope interaural time differences in reverberation. The Journal of the Acoustical Society of America 133, 2288-300.

Moore, B.C. 2004. Dead regions in the cochlea: conceptual foundations, diagnosis, and clinical applications. Ear and hearing 25, 98-116.

Moore, B.C. 2007. Binaural sharing of audio signals: Prospective benefits and limitations. The Hearing Journal 60, 46–48.

Moore, B.C. 2008. The choice of compression speed in hearing AIDS: theoretical and practical considerations and the role of individual differences. Trends in amplification 12, 103-12.

Moore, B.C., Glasberg, B.R. 2001. Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners. The Journal of the Acoustical Society of America 110, 1067-73.

Morera, C., Manrique, M., Ramos, A., Garcia-Ibanez, L., Cavalle, L., Huarte, A., Castillo, C., Estrada, E. 2005. Advantages of binaural hearing provided through bimodal stimulation via a cochlear implant and a conventional hearing aid: A 6-month comparative study. Acta oto-laryngologica 125, 596-606.

Noble, W., Tyler, R., Dunn, C., Bhullar, N. 2008. Unilateral and bilateral cochlear implants and the implant-plus-hearing-aid profile: comparing self-assessed and measured abilities. International journal of audiology 47, 505-14.

O'Brien, A., Yeend, I., Hartley, L., Keidser, G., Nyffeler, M. 2010. Evaluation of frequency compression and high-frequency directionality. The Hearing Journal 63, 32,34-37.

OPCI. 2016. www.opciweb.nl/ci-centra/aantal-implantaties-in-nederland.

Otte, R.J., Agterberg, M.J., Van Wanrooij, M.M., Snik, A.F., Van Opstal, A.J. 2013. Age-related hearing loss and ear morphology affect vertical but not horizontal sound-localization performance. Journal of the Association for Research in Otolaryngology : JARO 14, 261-73.

Owren, M.J., Cardillo, G.C. 2006. The relative roles of vowels and consonants in discriminating talker identity versus word meaning. The Journal of the Acoustical Society of America 119, 1727-39.

Oxenham, A.J., Kreft, H.A. 2014. Speech perception in tones and noise via cochlear implants reveals influence of spectral resolution on temporal processing. Trends in hearing 18.

Park, L.R., Teagle, H.F.B., Buss, E., Roush, P.A., Buchman, C.A. 2012. Effects of Frequency Compression Hearing Aids for Unilaterally Implanted Children With Acoustically Amplified Residual Hearing in the Nonimplanted Ear. Ear and hearing 33, E1-E12.

Perreau, A.E., Bentler, R.A., Tyler, R.S. 2013. The contribution of a frequency-compression hearing aid to contralateral cochlear implant performance. Journal of the American Academy of Audiology 24, 105-20.

Phonak. 2013. Adaptive Phonak Digital (APD). https://www.phonakpro.com/content/dam/phonak/gc_hq/b2b/en/evidence/compendium/Compendium_No4_Adaptive_Phonak_Digital.pdf.

Potts, L.G., Skinner, M.W., Litovsky, R.A., Strube, M.J., Kuk, F. 2009. Recognition and localization of speech by adult cochlear implant recipients wearing a digital hearing aid in the nonimplanted ear (bimodal hearing). Journal of the American Academy of Audiology 20, 353-73.

Qin, M.K., Oxenham, A.J. 2006. Effects of introducing unprocessed low-frequency information on the reception of envelope-vocoder processed speech. Journal of the Acoustical Society of America 119, 2417-2426.

Raab, D.H. 1962. Statistical facilitation of simple reaction times. Transactions of the New York Academy of Sciences 24, 574-90.

Ratcliff, R., Smith, P.L. 2004. A comparison of sequential sampling models for two-choice reaction time. Psychological review 111, 333-67.

Rayleigh, L. 1907. On our perception of sound direction. Philosophical Magazine 13.

Reddi, B.A., Asrress, K.N., Carpenter, R.H. 2003. Accuracy, information, and response time in a saccadic decision task. Journal of neurophysiology 90, 3538-46.

Reiss, L.A., Turner, C.W., Erenberg, S.R., Gantz, B.J. 2007. Changes in pitch with a cochlear implant over time. Journal of the Association for Research in Otolaryngology : JARO 8, 241-57.

Reiss, L.A., Turner, C.W., Karsten, S.A., Gantz, B.J. 2014a. Plasticity in human pitch perception induced by tonotopically mismatched electro-acoustic stimulation. Neuroscience 256, 43-52.

Reiss, L.A., Ito, R.A., Eggleston, J.L., Wozny, D.R. 2014b. Abnormal binaural spectral integration in cochlear implant users. Journal of the Association for Research in Otolaryngology : JARO 15, 235-48.

Reiss, L.A., Ito, R.A., Eggleston, J.L., Liao, S., Becker, J.J., Lakin, C.E., Warren, F.M., McMenomey, S.O. 2015. Pitch adaptation patterns in bimodal cochlear implant users: over time and after experience. Ear and hearing 36, e23-34.

Robinson, D.A. 1963. A method of measuring eye movements using a sclera search coil in a magnetic field. IEEE transactions on bio-medical engineering 10, 137-145.

Rubinstein, J.T. 2004. How cochlear implants encode speech. Curr Opin Otolaryngol Head Neck Surg. 12, 444-8.

Schroter, H., Ulich, R., Miller, J. 2007. Effects of redundant auditory stimuli on reaction time. Psychonomic bulletin & review 14, 39-44.

Seeber, B.U., Fastl, H. 2008. Localization cues with bilateral cochlear implants. Journal of the Acoustical Society of America 123, 1030-1042.

Seeber, B.U., Baumann, U., Fastl, H. 2004. Localization ability with bimodal hearing aids and bilateral cochlear implants. The Journal of the Acoustical Society of America 116, 1698-709.

Shaw, E.A. 1974. Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. The Journal of the Acoustical Society of America 56, 1848-61.

Shim, H.J., Won, J.H., Moon, I.J., Anderson, E.S., Drennan, W.R., McIntosh, N.E., Weaver, E.M., Rubinstein, J.T. 2014. Can Unaided Non-Linguistic Measures Predict Cochlear Implant Candidacy? Otology & Neurotology 35, 1345-1353.

Shpak, T., Most, T., Luntz, M. 2014. Fundamental frequency information for speech recognition via bimodal stimulation: cochlear implant in one ear and hearing aid in the other. Ear and hearing 35, 97-109.

Simpson, A., Hersbach, A.A., McDermott, H.J. 2005. Improvements in speech perception with an experimental nonlinear frequency compression hearing device. International journal of audiology 44, 281-292.

Simpson, A., Hersbach, A.A., McDermott, H.J. 2006. Frequency-compression outcomes in listeners with steeply sloping audiograms. International journal of audiology 45, 619-629.

Smith P, D.A., Day J, Unwin S, Day G, Chalupper, J. 2008. Real-World Preferences for Linked Bilateral Processing. The Hearing Journal 61, 33-38.

Sockalingam R, H.M., Eneroth K, Schulte M. 2009. Binaural hearing aid communication shown to improve sound quality and localization. Hearing Journal 62, 46-47.

Spriet, A., Van Deun, L., Eftaxiadis, K., Laneau, J., Moonen, M., van Dijk, B., van Wieringen, A., Wouters, J. 2007. Speech understanding in background noise with the two-microphone adaptive beamformer BEAM in the Nucleus Freedom Cochlear Implant System. Ear and hearing 28, 62-72.

Stone, M.A., Moore, B.C. 2007. Quantifying the effects of fast-acting compression on the envelope of speech. The Journal of the Acoustical Society of America 121, 1654-64.

Straatman, L.V., Rietveld, A.C., Beijen, J., Mylanus, E.A., Mens, L.H. 2010. Advantage of bimodal fitting in prosody perception for children using a cochlear implant and a hearing aid. The Journal of the Acoustical Society of America 128, 1884-95.

Studebaker, G.A. 1985. A "rationalized" arcsine transform. J Speech Hear Res 28, 455-62.

Sucher, C.M., McDermott, H.J. 2009. Bimodal stimulation: benefits for music perception and sound quality. Cochlear implants international 10 Suppl 1, 96-9.

Summers, V., Leek, M.R. 1994. The internal representation of spectral contrast in hearing-impaired listeners. The Journal of the Acoustical Society of America 95, 3518-28.

Supin, A., Popov, V.V., Milekhina, O.N., Tarakanov, M.B. 1994. Frequency resolving power measured by rippled noise. Hearing research 78, 31-40.

Theelen-van den Hoek, F.L., Houben, R., Dreschler, W.A. 2014. Investigation into the applicability and optimization of the Dutch matrix sentence test for use with cochlear implant users. International journal of audiology 53, 817-28.

Townsend, J.T., Nozawa, G. 1995. Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. J Math Psychol 39, 321-359.

Tyler, R.S., Dunn, C.C., Witt, S.A., Preece, J.P. 2003. Update on bilateral cochlear implantation. Current opinion in otolaryngology & head and neck surgery 11, 388-93.

Tyler, R.S., Parkinson, A.J., Wilson, B.S., Witt, S., Preece, J.P., Noble, W. 2002. Patients utilizing a hearing aid and a cochlear implant: speech perception and localization. Ear and hearing 23, 98-105.

Van Barneveld, D.C., Van Wanrooij, M.M. 2013. The influence of static eye and head position on the ventriloquist effect. The European journal of neuroscience 37, 1501-10.

Van Deun, L., van Wieringen, A., Wouters, J. 2010. Spatial speech perception benefits in young children with normal hearing and cochlear implants. Ear and hearing 31, 702-13.

van Hoesel, R.J. 2012. Contrasting benefits from contralateral implants and hearing aids in cochlear implant users. Hearing research 288, 100-13.

Van Opstal, A.J. 2016. Sound Source Localization in the Human Auditory System. Elsevier Science & Technology Books.

Van Wanrooij, M.M., Van Opstal, A.J. 2004. Contribution of head shadow and pinna cues to chronic monaural sound localization. The Journal of neuroscience : the official journal of the Society for Neuroscience 24, 4163-71.

Van Wanrooij, M.M., Van Opstal, A.J. 2005. Relearning sound localization with a new ear. J Neurosci 25, 5413-24.

Van Wanrooij, M.M., Van Opstal, A.J. 2007. Sound localization under perturbed binaural hearing. Journal of neurophysiology 97, 715-26.

Van Wanrooij, M.M., Bremen, P., John Van Opstal, A. 2010. Acquired prior knowledge modulates audiovisual integration. Eur J Neurosci 31, 1763-71.

Van Wanrooij, M.M., Bell, A.H., Munoz, D.P., Van Opstal, A.J. 2009. The effect of spatial-temporal audiovisual disparities on saccades in a complex scene. Exp Brain Res 198, 425-37.

van Wieringen, A., Wouters, J. 2008. LIST and LINT: sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and the Netherlands. International journal of audiology 47, 348-55.

Versfeld, N.J., Daalder, L., Festen, J.M., Houtgast, T. 2000. Method for the selection of sentence materials for efficient measurement of the speech reception threshold. The Journal of the Acoustical Society of America 107, 1671-84.

Veugen, L.C., Chalupper, J., Snik, A.F., van Opstal, A.J., Mens, L.H. 2016a. Frequency-dependent loudness balancing in bimodal cochlear implant users. Acta oto-laryngologica 136, 775-81.

Veugen, L.C., Chalupper, J., Snik, A.F., Opstal, A.J., Mens, L.H. 2016b. Matching Automatic Gain Control Across Devices in Bimodal Cochlear Implant Users. Ear and hearing 37, 260-70.

Veugen, L.C.E., Hendrikse, M.M.E., Van Wanrooij, M.M., Agterberg, M.J., Chalupper, J., Mens, L.H.M., Snik, A.F.M., van Opstal, A.J. 2016c. Horizontal Sound Localization in Cochlear Implant Users with a Contralateral Hearing Aid. Hearing research Submitted for publication.

Veugen, L.C., Van Wanrooij, M.M., Louvet, D., Van Opstal, A.J. 2016d. Monaural and binaural response reaction times to spectrotemporal modulations: normal-hearing and impaired-hearing simulation. in prep.

Vickers, D., De Raeve, L., Graham, J. 2016. International survey of cochlear implant candidacy. Cochlear implants international 17 Suppl 1, 36-41.

Viemeister, N.F. 1979. Temporal modulation transfer functions based upon modulation thresholds. The Journal of the Acoustical Society of America 66, 1364-80.

Wiggins, I.M., Seeber, B.U. 2013. Linking dynamic-range compression across the ears can improve speech intelligibility in spatially separated noise. The Journal of the Acoustical Society of America 133, 1004-16.

Wilson, B.S., Dorman, M.F. 2008. Cochlear implants: a remarkable past and a brilliant future. Hearing research 242, 3-21.

Wolfe, J., John, A., Schafer, E., Nyffeler, M., Boretzki, M., Caraway, T., Hudson, M. 2011. Long-term effects of non-linear frequency compression for children with moderate hearing loss. International journal of audiology 50, 396-404.

Won, J.H., Drennan, W.R., Rubinstein, J.T. 2007. Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. Jaro-Journal of the Association for Research in Otolaryngology 8, 384-392.

Won, J.H., Drennan, W.R., Nie, K., Jameyson, E.M., Rubinstein, J.T. 2011. Acoustic temporal modulation detection and speech perception in cochlear implant listeners. J Acoust Soc Am 130, 376-88.

Won, J.H., Moon, I.J., Jin, S., Park, H., Woo, J., Cho, Y.S., Chung, W.H., Hong, S.H. 2015. Spectrotemporal Modulation Detection and Speech Perception by Cochlear Implant Users. PloS one 10, e0140920.

www.opciweb.nl. 2012. http://www.opciweb.nl/fileadmin/default/bestanden/standpuntswijziging_cochleaire-implantaten-v2.pdf.

www.zorginstituutnederland.nl. 2014. https://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/rapporten-en-standpunten/2014/1401-tweede-cochleair-implantaat-voor-kinderen-tussen-5-en-18-jaar/1401-tweede-cochleair-implantaat-voor-kinderen-tussen-5-en-18-.

Yoon, Y.S., Shin, Y.R., Gho, J.S., Fu, Q.J. 2015. Bimodal benefit depends on the performance difference between a cochlear implant and a hearing aid. Cochlear implants international 16, 159-67.

Zhang, T., Dorman, M.F., Spahr, A.J. 2010. Information from the voice fundamental frequency (F0) region accounts for the majority of the benefit when acoustic stimulation is added to electric stimulation. Ear and hearing 31, 63-9.

Zhang, T., Spahr, A., Dorman, M. and Saoji, A. 2013. The relationship between auditory function of non-implanted ears and bimodal benefit. Ear and hearing 34, 133-141.

DONDERS GRADUATE SCHOOL FOR COGNITIVE NEUROSCIENCE

For a successful research Institute, it is vital to train the next generation of young scientists. To achieve this goal, the Donders Institute for Brain, Cognition and Behaviour established the Donders Graduate School for Cognitive Neuroscience (DGCN), which was officially recognised as a national graduate school in 2009. The Graduate School covers training at both Master's and PhD level and provides an excellent educational context fully aligned with the research programme of the Donders Institute.

The school successfully attracts highly talented national and international students in biology, physics, psycholinguistics, psychology, behavioral science, medicine and related disciplines. Selective admission and assessment centers guarantee the enrolment of the best and most motivated students.

The DGCN tracks the career of PhD graduates carefully. More than 50% of PhD alumni show a continuation in academia with postdoc positions at top institutes worldwide, e.g. Stanford University, University of Oxford, University of Cambridge, UCL London, MPI Leipzig, Hanyang University in South Korea, NTNU Norway, University of Illinois, North Western University, Northeastern University in Boston, ETH Zürich, University of Vienna etc.

Positions outside academia spread among the following sectors:

- specialists in a medical environment, mainly in genetics, geriatrics, psychiatry and neurology,

- specialists in a psychological environment, e.g. as specialist in neuropsychology, psychological diagnostics or therapy,

- higher education as coordinators or lecturers.

A smaller percentage enters business as research consultants, analysts or head of research and development. Fewer graduates stay in a research environment as lab coordinators, technical support or policy advisors. Upcoming possibilities are positions in the IT sector and management position in pharmaceutical industry. In general, the PhDs graduates almost invariably continue with high-quality positions that play an important role in our knowledge economy.

For more information on the DGCN as well as past and upcoming defenses please visit:

http://www.ru.nl/donders/graduate-school/phd/