

Coherent Coding of Enhanced Interaural Cues Improves Sound Localization in Noise With Bilateral Cochlear Implants

Trends in Hearing
Volume 22: 1–18
© The Author(s) 2018
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2331216518781746
journals.sagepub.com/home/tia



Ben Williges¹, Tim Jürgens^{1,2}, Hongmei Hu¹, and Mathias Dietz^{1,3}

Abstract

Bilateral cochlear implant (BCI) users only have very limited spatial hearing abilities. Speech coding strategies transmit interaural level differences (ILDs) but in a distorted manner. Interaural time difference (ITD) information transmission is even more limited. With these cues, most BCI users can coarsely localize a single source in quiet, but performance quickly declines in the presence of other sound. This proof-of-concept study presents a novel signal processing algorithm specific for BCIs, with the aim to improve sound localization in noise. The core part of the BCI algorithm duplicates a monophonic electrode pulse pattern and applies quasistationary natural or artificial ITDs or ILDs based on the estimated direction of the dominant source. Three experiments were conducted to evaluate different algorithm variants: Experiment 1 tested if ITD transmission alone enables BCI subjects to lateralize speech. Results showed that six out of nine BCI subjects were able to lateralize intelligible speech in quiet solely based on ITDs. Experiments 2 and 3 assessed azimuthal angle discrimination in noise with natural or modified ILDs and ITDs. Angle discrimination for frontal locations was possible with all variants, including the pure ITD case, but for lateral reference angles, it was only possible with a linearized ILD mapping. Speech intelligibility in noise, limitations, and challenges of this interaural cue transmission approach are discussed alongside suggestions for modifying and further improving the BCI algorithm.

Keywords

cochlear implants, localization, lateralization, binaural cues

Date received: 22 September 2017; revised: 3 May 2018; accepted: 14 May 2018

Introduction

Providing a second cochlear implant (CI) to unilaterally implanted CI patients has been shown to improve speech intelligibility, spatial awareness (van Schoonhoven et al., 2013), and sound source localization (Kerber & Seeber, 2012). However, many benefits of the second CI lag behind the benefits that normal-hearing (NH) listeners get from the combined usage of their two ears.

The main cues involved by using two ears (or CIs) are interaural time differences (ITDs) and interaural level differences (ILDs). The brain can exploit these binaural cues to create an auditory scene (Bregman, 1999), and to segregate target speech from spatially separated noise (Roman, Wang, & Brown, 2003), to improve speech intelligibility. Although NH listeners enjoy finely tuned neural circuits that extract binaural cues even at negative signal-to-noise ratios (SNRs) (e.g., Licklider, 1948), most forms of hearing impairment have a detrimental impact

on the processes that extract the binaural cues. Furthermore, many hearing devices are known to corrupt binaural cues, because the algorithms used in these devices aim to optimize speech intelligibility for each ear in isolation. Even stronger than typical hearing aids, CIs corrupt or even discard most binaural information: ILD information is reduced due to the limited dynamic range

¹Medizinische Physik and Cluster of Excellence “Hearing4all,” Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

²Institute of Acoustics, University of Applied Sciences Lübeck, Lübeck, Germany

³National Centre for Audiology, School of Communication Sciences and Disorders, Western University, London, Ontario, Canada

Corresponding author:

Ben Williges, Medizinische Physik and Cluster of Excellence “Hearing4all,” Carl von Ossietzky Universität Oldenburg, Kückersweg 74, 26129 Oldenburg, Germany.
Email: ben.williges@uni-oldenburg.de



(DR) available in electric hearing (Kerber & Seeber, 2012), and ITD information is largely disregarded due to the usage of high pulse rate coding strategies and missing time-synchronization of the two independent sound processors. In addition, even if binaural cues are provided with precisely synchronized timing in controlled laboratory settings, the CI user's sensitivity to binaural cues is reduced due to the electric stimulation itself (e.g., Smith & Delgutte, 2007) and other subject-related factors (see, e.g., Dietz, 2016, for an overview).

Bilateral CI (BCI) users have a reduced ITD sensitivity compared to NH listeners and require low pulse rates to detect ITDs provided in the pulse timing (e.g., van Hoesel, 2007). More recently, Laback, Egger, and Majdak (2015) showed that for a given constant ITD, that was provided using low-rate pulse trains at a single electrode on each side, BCI users can lateralize fairly well. However, BCI users show a smaller extent of lateralization than NH listeners (Baumgärtel, Hu, Kollmeier, & Dietz, 2017). Churchill, Kan, Goupell, and Litovsky (2014) showed that lateralization based solely on ITDs was also possible when using multiple (five) electrodes with constant low-rate pulse trains. They reported that pulse rate is a strong trade-off between intelligibility and lateralization in continuous interleaved sampling (CIS)-based coding strategies and suggest mixed-rates as a possible trade-off. As a follow-up to Churchill et al. (2014), we want to investigate if pure ITD-based lateralization is also possible with varying rates on each electrodes, as is the case for non CIS-based coding strategies (e.g., Smith, 2014; van Hoesel, 2007). Our first hypothesis is that BCI listeners are able to lateralize such stimuli when there is no interaural electrode mismatch (Hu & Dietz, 2015).

Most BCI users can localize a single sound source with fair accuracy (e.g., van Hoesel & Tyler, 2003). However, while NH listeners can even localize in background noise at negative SNRs, BCI users show a strong decrease in localization performance already at positive SNRs (Kerber & Seeber, 2012). Furthermore, BCI users cannot differentiate early and late room reflections to improve localization and speech intelligibility (Kokkinakis, Hazrati, & Loizou, 2011). Signal processing algorithms have been developed in the past decades primarily to improve speech intelligibility for hearing aid users and have recently also been applied for BCI users (for a review, see Wouters, Doclo, Koning, & Francart, 2013). A very common approach is spatial filters, such as multichannel Wiener filters or beamformers (Luts et al., 2010), which provide a substantial SNR improvement. The beam of such a spatial filter is typically steered toward the front, preserving sound from a frontally located talker and decreasing interferer sound coming from other directions, but it can also be steered to the dominant sound source (e.g., Adiloglu et al., 2015). It is

the nature of these binaural spatial filters that their binaural output consists of high interaural coherence and both target and the residual (attenuated) interferer have the same binaural cues. Therefore, these algorithms have two disadvantages for a typical bilateral hearing aid user: (a) The reduction of spatial perception to only a small section of the full 360° auditory space (toward the beam direction) and (b) the inability to use their own binaural system for efficient spatial filtering, which in turn destroys binaural squelch and reduces spatial release from masking. Modified beamformers have been proposed that better preserve binaural cues of the interferers or provide deliberately reduced coherence (e.g., Marquardt, Hohmann, & Doclo, 2015). In BCI users, the starting point is somewhat different: As argued earlier, BCI users' average performance in localization and speech intelligibility is very susceptible to noise, reverberation or a second source. In addition, little benefit is obtained from spatially separating multiple competing talkers (Hu, Dietz, Williges, & Ewert, 2018) in contrast to hearing aid users (Best, Roverud, Streeter, Mason, & Kidd, 2017). Consequently, for BCI users there may even be an advantage in reducing the perceived auditory space to the most prominent sound source, that is, in the beam direction. Our second hypothesis is therefore that highly coherent and (quasi) stationary binaural cues as provided by a steering binaural beamformer may strongly improve localization and sound direction discrimination in noise.

Preserving the natural binaural cues of the target is a frequently expressed and challenging goal (e.g., Wouters et al., 2013). Despite some clear merit of natural cues, several studies hint that they are not necessarily ideal for sound localization with CIs. A study from Francart, Lenssen, and Wouters (2011) has shown that application of artificial interaural cues can improve localization in bimodal CI users and similar effects can be expected for BCI users (Kelvasa & Dietz, 2015). Furthermore, if ITDs would be provided at low enough pulse rates, BCI users would likely benefit from enlarged ITDs rather than natural ITDs (Baumgärtel et al., 2017). Given this, the third and final hypothesis of this study is that enhanced ILDs and, if available, enlarged ITDs, improve BCI listeners sound direction discrimination compared to natural cues.

Here we present a BCI algorithm whose development was inspired by the three hypotheses mentioned earlier. While describing and discussing the algorithm are the focus of this article, we also present data from three experiments, primarily to identify the practical challenges but also to provide a proof-of-concept. Experiment 1 does not employ the full BCI algorithm but focusses only on the first hypothesis: ITDs are applied to the output of a low-rate speech coding strategy and we test ITD-based lateralization. In Experiment 2, we ignore ITDs using a conventional continuous interleaved sampling (CIS) strategy together with the new BCI algorithm to test

Hypothesis 2 (stationary ILDs from beamformer) and Hypothesis 3 (ILD modifications). The experiment is sound direction discrimination in noise. Finally, Experiment 3 is the same task as Experiment 2 but with the low-rate speech coding strategy from Experiment 1 and with manipulating ITD rather than ILD. As such, the speech coding of Experiment 3 includes three deviations from classical CIS: low-rate speech coding, beamforming with interaurally coherent output, and ITD enhancement.

Methods

Subjects

Nine bilateral MED-EL CI users (age: 18–74 years, median 58 years) participated in this study. BCI1, BCI3, and BCI4 also participated in previous studies on single-electrode extent of lateralization (Baumgärtel et al., 2017) and in an interaural electrode matching study (Hu & Dietz, 2015). Demographic details of the subjects can be found in Table 1. The experiments were conducted within two sessions, each of approximately 2 hours on different days. All subjects provided voluntary written consent and were paid a stipend for participation. The study was conducted with approval of the Ethics committee of the University of Oldenburg.

Apparatus

This study used direct electric stimulation, bypassing the subjects' own speech processors. Instead, electrical stimulation patterns for both implants were generated using customized scripts in Matlab (The MathWorks) running on a standard PC. These patterns were transferred to a Research Interface Box II (RIB-II, University of Innsbruck, Austria) that was connected to this PC via

a National Instruments I/O-card. The RIB-II transmitted the stimulation patterns to the internal parts of the user's CIs via an optical isolation interface and telemetry coils, that is, the subjects did not wear their own speech processors. Stimulus verification was carried out using two implants-in-a-box (RIB i100 detector box, University of Innsbruck, Austria) connected to an oscilloscope. In addition, the detector box output was recorded with a sound-card, to verify correct across-channel sequential stimulation timing. All the electric stimuli used in this study had a pulse phase duration of 32 μ s and stimulation was performed with anodic-first biphasic pulse trains with an interphase gap of 2.1 μ s. All multielectrode stimulation was done sequentially.

For comparison, all BCI users were also tested with over-ear headphones covering their own speech processor microphones, that is, providing the acoustic signal to the speech processors that were used every day. These acoustic stimuli were presented via Sennheiser HD 600 Headphones and a RME Fireface UC/X sound-card at 65 dB SPL.

Pretests

Individual maximum-comfort (MC) and hearing threshold (THR) levels for each electrode were determined using the following procedure: The impedance for each of the 12 electrodes was measured and for each side (left or right) the electrode with the highest impedance served as starting electrode. MC and THR levels were measured using a one-up-one-down procedure with a pulse train of 1 sec duration at a constant stimulation rate depending on coding strategy employed (see Table 2). The subjects were asked to change the current level of the pulse train in linear factors of 0.95 (decrease) or 1.05 (increase) of the preceding pulse train, until the perceived sound was just audible (THR level) or loud enough to be acceptable

Table 1. Demographic Details of the BCI Subjects Included in the Study.

ID	Sex	Age (years)	Etiology	Years exp HA (left/right)	Years exp CIs (left/right)	OISa SRT (S0°N0° [dB])	Implant type (left/right)
BCI1	F	57	Measles	16/12	11/15	0.4	Pulsar/Pulsar
BCI2	M	74	Unknown	16/10	2/7	2.3	Concerto/Sonata
BCI3	F	61	Sudden hearing loss	0/2	7/6	3.1	Sonata/Sonata
BCI4	M	58	Noise	15/15	13/9	0.9	Pulsar/Sonata
BCI5	M	56	Progressive	47/47	1.8/2.7	6.2	Synchrony/Concerto
BCI6	M	23	Genetic	0/0.5	21/19	−1.8	Pulsar/Concerto
BCI7	M	71	Noise	40/40	3.5/1	2.8	Synchrony/Synchrony
BCI8	F	18	Meningitis	0/0	0.8/0.8	−1.3	Synchrony/Synchrony
BCI9	M	60	Noise	15/15	2.3/1.4	3.8	Synchrony/Synchrony

Note. HA = hearing aid; CI = cochlear implant; BCI = bilateral cochlear implant.

Table 2. Frequency-to-Electrode Allocations and the Pulses Rates Used to Determine THR and MC Levels of Each Electrodes for Both Coding Strategies.

Electrode no.	1	2	3	4	5	6	7	8	9	10	11	12
Center frequency (Hz)	120	235	384	579	836	1,175	1,624	2,222	3,019	4,084	5,507	7,410
Maximum pulse rate of PP (pps)	173	279	396	190	231	293	333	273	294	334	384	330
Pulse rate of CIS (pps)	800	800	800	800	800	800	800	800	800	800	800	800

Note. For the PP strategy, the pulse rates were selected to be the maximal occurring pulse rates. PP = peak picking; CIS = continuous interleaved sampling.

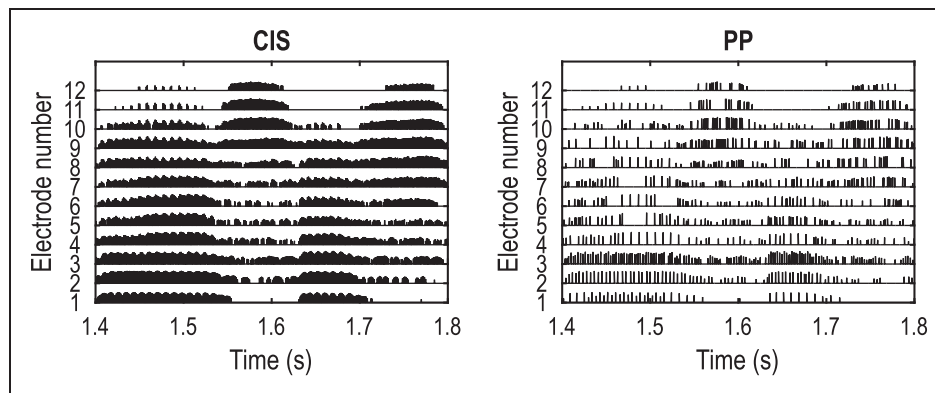


Figure 1. Electrodegrams of the two different CI sound coding strategies processing the German word “nasse” taken from a sentence of the Oldenburg sentence test (OISa “Doris bekommt acht nasse Steine,” English: “Doris gets eight wet stones”). For the PP strategy, note the increasing rate from Electrodes 1 to 3 (“fine-structure” electrodes) and the lower rates for Electrode 4 and higher, where the pulses correspond to temporal envelope maxima. CI = cochlear implant.

for a few minutes (MC level). The 0.8 (MC) or 1 (THR) level of the neighboring electrode served as starting value for the next (adjacent) electrode.

After collecting MC and THR levels for all electrodes, a subject-specific volume setting V to achieve comfortable binaural loudness was determined. Starting at $V=0.5$ (min=0, max=1, see Equation 2), the test stimulus (/ana/ with zero ITD for Experiment 1, /ana/ from 0° with isotropic noise at +5dB SNR for Experiments 2 and 3, and speech-shaped noise for the speech in noise experiment) was presented binaurally. The subject was asked to adjust the volume in step of 0.1 on a linear scale until the loudness was at medium comfortable loudness. Loudness across ears was then fine-tuned to achieve a centralized percept. The subject was asked to indicate, whether the perceived sound image was located centrally or shifted to the right or the left. When the subject responded left, the volume in the right ear was increased by 0.05 and the volume in the left ear was reduced by 0.05. Thus the subjects were able to shift the perceived image to the right or the left, until they perceived it as lateralized centrally. For measuring via headphones, the subjects own remote control was used to vary the loudness across ears, until a comfortable loudness in both ears was achieved.

CI Speech Coding Strategies

This subsection describes the conversion of a mono acoustic signal to an electric pulse pattern. In case of a two-channel input, the respective strategy is applied independently to each channel.

Two different CI sound coding strategies are used in this study: the continuous interleaved sampling (CIS, Wilson et al., 1991) coding strategy and a novel strategy, called peak picking (PP) strategy. The PP strategy combines elements of the fundamental asynchronous stimulus timing (FAST) strategy (Smith, 2014), Peak-derived timing strategy (PDT, van Hoesel, 2003), and the fine-structure processing (FSP) coding strategy (Hochmair et al., 2006), which all are in principle able to encode ITD information in the form of time differences between single stimulation pulses across left and right CI. Example electrodegrams of the CIS and PP coding strategy are displayed in Figure 1. The constant relatively high pulse rate in CIS makes it hard to identify single pulses in the left panel of Figure 1, whereas single pulses are clearly visible in the electrodegram of PP (right panel) because of its lower average pulse rate.

The first steps of the signal processing in both coding strategies (CIS and PP) are identical: The signal

is pre-emphasized with a first-order infinite impulse response (IIR)-Butterworth high-pass filter with a cutoff frequency of 1,200 Hz. Afterwards, a gammatone filterbank (Hohmann, 2002) with third order, three ERB wide filters is used to split the signal into 12 channels. The center frequencies of the channels (see Table 2) are in agreement with the center frequencies used in MED-EL devices (Nobbe, Schleich, Zierhofer, & Nopp, 2007). The use of the complex-valued gammatone filterbank allows the extraction of both the Hilbert envelope and the real-valued temporal bandpass-signal in each channel. The temporal envelopes are frequency weighted, by adapting the 22-channel weights from Nogueira, Büchner, Lenarz, and Edler (2005) to 12 channels for better speech perception (Hu et al., 2018). Each temporal envelope is low-pass filtered using a first-order IIR-Butterworth filter with a cutoff frequency of 200 Hz.

For the CIS-coding strategy, the low-pass filtered envelopes in the 12 channels are sampled at a fixed rate of 800 pps per channel. The stimulation is done sequentially with this rate from basal to apical electrodes.

The peak-picking coding strategy extracts pulse timing and amplitude of each channel’s local temporal maxima. For the nine most basal channels, the maxima of the low-pass filtered temporal envelope are extracted, similar to the FAST coding strategy. For the three most apical channels (i.e., up to a center frequency of 384 Hz), instead of the envelope, the real part of the complex-valued gammatone-filter output is used as the signal, which contains the temporal fine structure of the signal. Pulse timing and amplitude of this signal’s local maxima are used for the stimulation. The PP in each channel is independent of the PP in all the other channels; therefore, in principle, a simultaneous stimulation in two channels is possible. To prevent simultaneous stimulation in such a situation, the most apical pulse is given priority and any potential simultaneous pulse in more basal electrodes is omitted.

Loudness-Growth Function. For both the CIS and the PP coding strategies, the same loudness-growth-function (Equation 1, see Nogueira et al., 2005) was applied to each stimulation amplitude $s(k)$. This loudness-growth function set the output $DR(k)$ to 0 for input $s(k)$ below $b=0.0156$, a compressive logarithmic function with $\alpha=340.83$ between b and m , and a hard clipping (setting to m) for $s(k)$ bigger than $m=1.5859$ (Harczos, 2015, p. 18). To ensure an optimal fit of the stimulation amplitudes into the DR between m and b levels, the electrodiagram was scaled, such that the global (across all electrodes) root mean square was $(m-b)/20 + b$, corresponding to -20 dB full scale (FS). The output of this stage is subject-independent and corresponds to a percent DR

amplitude (b to 0%; m to 100%).

$$DR(k) = \begin{cases} 0, & s(k) < b, \\ \frac{\ln(1 + \alpha * (s(k) - b)/(m - b))}{\ln(1 + \alpha)} & b \leq s(k) \leq m, \\ 1, & s(k) > m \end{cases} \quad (1)$$

Individual Loudness Mapping. As a last step, the two % DR electrodiagrams were transformed to current units (CU) using Equation 2 with individually determined MC and THR levels for each electrode k and an electrode-independent volume setting V , ranging from 0 to 1 (see “Pretests” section).

$$CU(k) = THR(k) + (MC(k) - THR(k)) * DR(k) * V \quad (2)$$

Note that the current implementation was offline, that is, the whole stimulus was digitally generated, converted into an electrodiagram, then ITDs or ILDs were applied and both electrodiagrams were then transferred to the BCI subject’s implants.

Binaural Beamformer

A four-channel binaural beamformer was used in this study for the angle discrimination experiments with a binaural minimum variance distortionless response-design (Marquardt et al., 2015). Such an algorithm can be used to process the signals from the two microphones of each CI to one or two audio output signals. Upon specifying a target azimuthal beam direction, the beamformer provides effectively a spatial filtering, transmitting the target as undistorted as possible, and adaptively filtering out the other spatial directions, thus enhancing the SNR of sound originating from the target direction. The binaural cues for the target direction are preserved, by filtering all input signals with the same filter. The resulting output channels have a highly improved SNR (see, e.g., Baumgärtel et al., 2015). The output SNR (not the SNR improvement) is very similar in all output channels. The spatial filtering destroys the binaural cues of sounds from other directions than the target direction such that (independent of the input) the output is always perceived as a unidirectional sound field. In this study, the beam direction was set to the direction of the dominant sound, which was always the target sound. One SNR-improved output channel was then fed into the CI speech coding stage (see “CI speech coding strategies” section).

To specify the target azimuthal direction, a signal-driven broadband direction of arrival (DOA) estimator was used. The DOA cross-correlates the input signals

from the microphones. The cross-correlations are then classified using support-vector machines and mapped to a source presence probability using linear generalized models (Kayser & Anemüller, 2014). In a postprocessing step, the target direction is detected by averaging over 300 ms timeframes and settings constraints like a maximum speed of target movement. The DOA estimation time window allows for following a moving talker or alternating talkers, as well as following slow head movements. It operates very robustly even at -10 dB SNR in a diffuse noise field (Adiloglu et al., 2015).

Combination of Beamformer and CI Coding Strategy to the Full BCI Algorithm

This study used a novel BCI algorithm (Dietz & Backus, 2015; see Figure 2) specific for a binaural CI system that receives input signals from both left and right microphones, for example, through a wireless link. The algorithm essentially consists of two parts: a binaural beamformer with DOA estimator, and a novel conversion

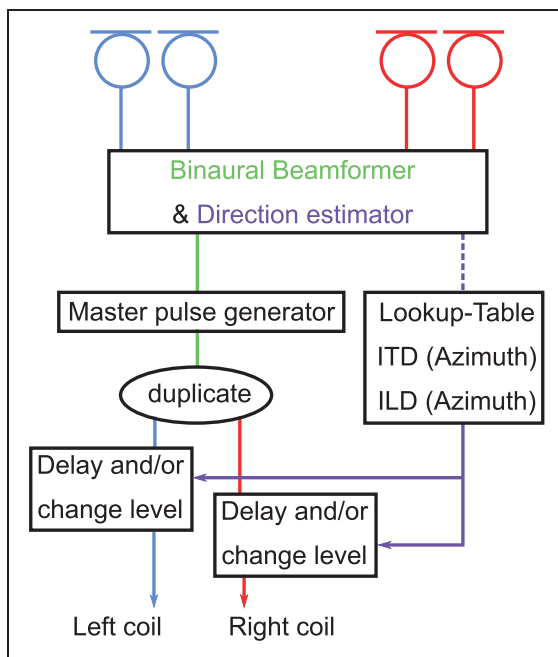


Figure 2. Signal flow diagram of the full BCI algorithm. The acoustic signals from the four microphones (two left, two right) are processed using a DOA estimator and a binaural beamformer. Afterwards, either the right or the left channel (corresponding to the DOA) is selected and transformed to a multichannel electrogram (“master pulse generator;” this can be done using any speech coding strategy). The resulting pulse pattern is duplicated and ITD or ILD information is applied on a pulse-by-pulse basis before being sent to the left and right research processor coils. BCI = bilateral cochlear implant; ILD = interaural level difference; ITD = interaural time difference; DOA = direction of arrival.

to binaural electrode pulse patterns which allows for mapping of natural or artificial (manipulated) ITDs and ILDs.

First, the signal is passed through the DOA and the steered beamformer described in the previous subsection. The required output for the novel BCI algorithm is the DOA and a single SNR improved channel. As stated earlier, the beamformer output channels have an almost identical SNR, so the selection which of the channels to use is somewhat arbitrary. In this study, the channel from the microphone closest to the dominant sound source was chosen based on the location estimation of the DOA.

The novel core part of the algorithm is that only one acoustic channel is converted to an electrogram using the desired speech coding strategy, such as CIS, PP or any other. The resulting electrogram is then duplicated and each pulse is shifted in time or altered in % DR value according to ITD or ILD azimuth lookup tables. In contrast to previous attempts to modify the highly dynamic interaural differences between the left and right acoustic signal (Brown, 2018; Francart et al., 2011; Kollmeier & Peissig, 1990), quasistationary interaural cues are applied in this study. This procedure avoids artifacts due to altering of the acoustic waveforms and results in an interaurally fully coherent output. The applied binaural cues are generally not expected to be the same as the original ITDs and ILDs at the level of the sound receivers. Instead, optimized ITDs or ILDs or combinations of the two can be applied that offer the most accurate localization for an individual BCI user, as suggested by Brown (2016, 2018).

ILD and ITD Mapping. To determine the frequency-dependent ITDs and ILDs, as they would occur naturally in CI users, speech-shaped noise was convolved with anechoic binaural head-related impulse responses (HRIRs; distance: 80 cm, elevation: 0° , azimuthal resolution: 5° ; Kayser et al., 2009). To determine the naturally occurring ILDs in BCIs, the binaural signal was processed using the CIS strategy (with independent processors) at 65 dB SPL. The left and right electrogram (% of DR values) were then subtracted for each channel to get the ILD in percentage of the DR (see dashed lines in Figure 3(a)). Note that the ILD’s change with azimuth is a nonmonotonic function, that is, there are different angles resulting in the same ILD (e.g., for 20° and 60° sound incident angle in the 836 and 1,624 Hz frequency channels). In this study, a linear and frequency-independent mapping function was employed. The linear mapping maximum was set to 60% of the total DR at 90° , that is, 1° azimuth change corresponded to 0.67% ILD (DR) change. While naturally occurring ILDs are highly frequency-dependent, the motivation for choosing a frequency-independent map is twofold: First, the human ILD discrimination ability is mostly frequency-independent (e.g., Rowland & Tobias,

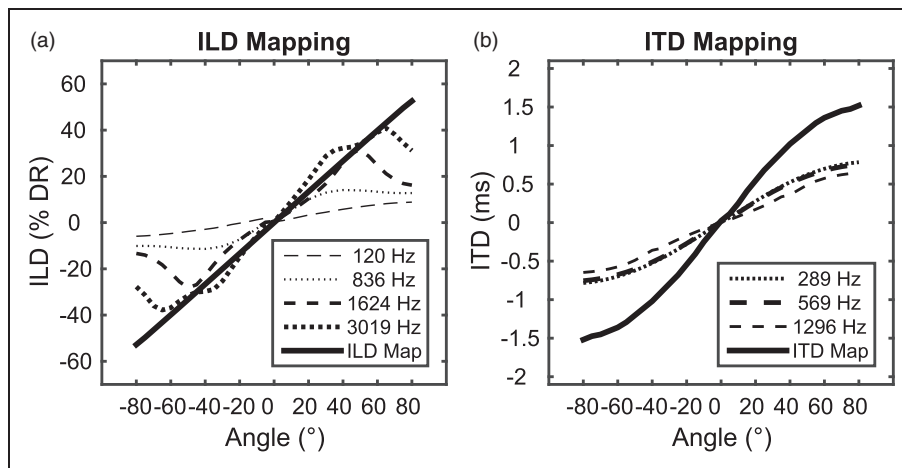


Figure 3. KEMAR-recorded ILDs (a) and ITDs (b) in different frequency bands (dashed lines) together with the employed lookup tables for ITDs and ILD maps (solid lines). Note that the ILDs are not presented in dB, because the ILD mapping is applied in percent dynamic range. Here, the input stimulus was speech shaped noise at 65 dB SPL. ILD = interaural level difference; ITD = interaural time difference.

1967). Secondly, the loudness growth and average DR of CI users is on average independent of the tonotopic place of stimulation (Busby, Whitford, Blamey, Richardson, & Clark, 1994). When a given % DR pulse was going to be reduced to below 0 % DR within the ILD mapping procedure, no stimulation was done for this pulse at the respective side. Similarly, stimulation was capped at 100% DR at the ipsilateral side.

The natural ITD was determined using speech-shaped noise as an input to the binaural model of Dietz, Ewert, and Hohmann (2011). The model output chosen here was the time-varying and frequency-dependent ITD. The ITDs (median over time) for some exemplary frequency channels are depicted in Figure 3(b). Similar to the ILD concept, a frequency-independent ITD was employed for this study, extracted from an auditory filter centered at 569 Hz. This ITD was then multiplied by two and employed for all 12 frequency bands as ITD lookup-table (see Figure 3(b)). The factor two was chosen based on the findings of Baumgärtel et al. (2017) that CI users require on average about twice the NH ITD to perceive a fully lateralized sound image.

Experiment 1: ITD-Based Lateralization in Quiet

To test whether the core part of the BCI algorithm alone (without beamformer) is able to provide ITD-based cues for lateralization in quiet, a visual pointer lateralization procedure was conducted (Baumgärtel et al., 2017; Churchill et al., 2014; Kan, Stoelb, Litovsky, & Goupell, 2013). As stimulus, the speech logatome /ana/ (Wesker et al., 2005) was used and processed with the PP speech coding strategy. The resulting electrodogram was duplicated and one out of 13 different ITDs (± 3 ms, ± 2 ms, ± 1.4 ms, ± 1 ms, ± 0.6 ms, ± 0.2 ms, 0 ms) was applied. The

order of presentation was randomized and each ITD was presented three times. Subjects were asked to indicate where they perceived (lateralized) the sound on a line across the head ranging from beyond the left ear to beyond the right ear. The continuous line was displayed on a touch screen and subjects were asked to press the perceived position on that line. This position was then interpreted in terms of lateralization units, where -10 corresponds to the left ear and $+10$ corresponds to the right ear. Subjects could replay the presented signal to confirm their perceived lateralization. The data for each subject was fitted with a sigmoidal function (Equation 3) in the range of $ITD = \pm 2$ ms (Baumgärtel et al., 2017).

$$P(ITD) = P_l + \frac{P_r - P_l}{1 + 10^{P_{slope}(ITD_c - ITD)}} \quad (3)$$

P_l and P_r denote the individual BCI users' maximum lateralization for right and left side in lateralization units. ITD_c denotes the ITD of centered sound percept (in ms). P_{slope} describes the slope at ITD_c . In addition, the following parameters are then calculated from the fit: LatRatio describes the ratio of the lateralization range within the naturally occurring ITD range of $\pm 600 \mu s$ to the maximum lateralization range. The variability of the lateralization relative to the maximum lateralization range is described by dividing the standard deviation across each ITD condition by the maximum lateralization range (SDev/Range). SDev/Slope is a measure of ITD sensitivity. Similar to Monaghan and Seeber (2016), the mean standard deviation in perceived lateralization for each presented ITD is divided by the slope of the fitted function.

At the beginning of the experiment, a brief familiarization to the stimuli was conducted. The subjects

were presented with seven to eight randomly selected ITDs and asked to confirm whether the stimulus was intelligible and whether any loudness differences were perceivable between the ears. If there were any loudness differences, the binaural loudness balancing and centralization procedure from the pretest section was conducted again.

Experiment 2: Angle Discrimination With Varying Access to ITD Cues

This experiment was designed to investigate the abilities of the BCI users to discriminate sound sources in noise with different access to ITD cues. Angle discrimination performance was measured using a two-interval, two-alternative forced choice procedure (Hartmann & Rakerd, 1989). The subject was presented with a pair of two speech signals (logatome /ana/, 416 msec each) separated by a silent interval of 220 msec. The pair of signals represented two randomized sound source azimuths with a fixed spatial separation of 20° from the set of either -10° versus 10°, -65° versus -45°, or 45° versus 65°. The azimuths were realized by convolving the speech signal with HRIRs from behind-the-ear hearing aid microphones (Kayser et al., 2009). Signals within each pair were presented in pseudo-randomized order, added to isotropic noise at an SNR of +5 dB at the input of the microphones. The isotropic noise was generated by convolving randomly sampled parts of speech-shaped noise with one HRIR for each possible direction (0°–360° horizontal angle, 5° separation, and 0°

elevation) and adding the noises for the different directions together. The task of the subject was to indicate on a touch screen with buttons within a graphical user interface whether the second signal was to the left or to the right of the first signal. Feedback was provided. The result of this experiment was the percentage score of correctly identified relative azimuthal positions judgments using 30 repetitions for each direction set. Each direction set was tested for three algorithms resulting in a total of 270 trials (three algorithms × three directions × 30 repetitions).

Three processing variants were tested, all built on the PP coding strategy. They differed in the amount of transmitted interaural cues and whether or not a beamformer was used. Due to the low pulse rates, the PP coding strategy allowed to transmit perceptually relevant ITDs in the pulse timing. Expected transmission of ILDs, ITDs within the envelope of the signals (ITD_{env}) and within the pulse timing (ITD_{pulse}) are indicated in Table 3 (right part). The processing variants were (a) binaurally independent PP without beamformer (“PP”), providing natural ILD and ITD cues that are, however, very susceptible to noise; (b) PP with beamformer (“PP + beam”), providing very coherent ILD and ITD cues corresponding to the target direction; and (c) PP coding strategy with the BCI algorithm, that is, beamformer and artificial ITD mapping (discarding any ILD cues), as explained in section ILD and ITD Mapping (“PP + ITD”).

As a comparison, and for familiarization with the measurement procedure, the performance with the BCI

Table 3. Overview of the Tested Conditions and Their Expected Amount of ILD or ITD Envelope or ITD Fine Structure Transmission in This Experiment Setup in Quiet Conditions.

Algorithm	Coding strategy	Beamformer	Mapping	ILD	ITD_{env}	ITD_{pulse}
PP	Peak Picking	–	–	Preserved, but susceptible to noise	Preserved, but susceptible to noise	Preserved, but susceptible to noise
PP + Beam	Peak Picking	+	–	Preserved	Preserved	Preserved
PP + ITD	Peak Picking	+	ITD	Not transmitted in current version	Individualized transmission possible	Individualized transmission possible
Own CI devices (via headphone)		–	–	Distorted transmission	Preserved, but susceptible to noise	Not transmitted (except for FS4 channels)
CIS	CIS	–	–	Preserved, but susceptible to noise	Preserved, but susceptible to noise	Not transmitted
CIS + Beam	CIS	+	–	Preserved	Preserved	Not transmitted
CIS + ILD	CIS	+	ILD	Individualized transmission possible	Not transmitted in current version	Not transmitted

Note. “+” denote algorithms steps, that are present in the specific tested conditions, “–” denote steps, which are not present. PP = peak picking; ILD = interaural level difference; CIS = continuous interleaved sampling; ITD = interaural time difference; CI = cochlear implant; ITD_{env} = ITDs between the envelopes of the signals; ITD_{pulse} = ITDs within the pulse timing.

listener's own speech processors via headphones was tested before using the PP coding strategy. For familiarization with the PP coding strategy five to six stimulus pairs coming from different randomly chosen angles were presented before the experiment, and feedback was provided on the responses. This familiarization also served to confirm that the binaural loudness balancing procedure was correct and that the subjects were able to perceive a shift in angle direction when playing -45° and -65° compared to 10° and -10° or compared to 65° and 45° .

Experiment 3: Angle Discrimination With Varying Access to ILD Cues

The same angle discrimination experiment as with ITD cues (Experiment 2) was conducted with varying access to ILD cues using the CIS coding strategy. Three different CIS processing variants of the BCI algorithm were tested, again differing in the amount of transmitted interaural cues and whether or not a beamformer was used. These CIS-based variants do not convey ITD_{pulse} cues (see Table 3). The ITD_{env} information can be transmitted but is perceptually almost irrelevant in case of speech signals (Seeber & Fastl, 2008). The algorithm variants were (a) a binaurally independent CIS coding strategy in isolation, that is, without beamformer ("CIS"), providing natural ILD and ITD_{env} cues that are, however, very susceptible to noise; (b) CIS with beamformer ("CIS + beam"), applying the natural ILD and ITD_{env} cues of the target direction; and (c) CIS with the BCI algorithm, that is, beam former and artificial ILD mapping (disregarding any ITD_{env} cues), as explained in "ILD and ITD Mapping" section ("CIS + ILD").

Also preceding this experiment, five to six stimulus pairs coming from different randomly chosen sound incidents were presented for familiarization with the CIS coding strategy as implemented in this study. Again, feedback was provided.

Results

Experiment 1: ITD-Based Lateralization in Quiet

Figure 4 shows the extent of lateralization (means and standard deviation) for the logatome /ana/ in quiet as a function of the ITD. The extent of lateralization is quantified as the intracranial position of the perceived sound image of the subject, expressed as a numeric value between left ear (-10) and right ear ($+10$). Out of the nine subjects, three (BCI2, BCI5, and BCI9) did not have an ITD-dependent lateralization, that is, they lateralized consistently always to one side and reported that they never perceived a change in sound lateralization, even after multiple attempts of redoing the loudness and centralization task. Thus, the measurements of

Experiment 1 were aborted for these subjects and the (incomplete) data are not reported in Figure 4. The other six subjects showed an ITD-dependent lateralization. Pooling the answers for negative ITD values and comparing them to positive ITD values revealed significant differences between the medians using a Wilcoxon rank test (see Figure 4), indicating a significant right or left discrimination based on ITD. For all BCI subjects, the measured extents of lateralization for the different ITDs did not cover the complete range between the ears (P_l and P_r are not at -10 or $+10$ lateralization units, see Table 4). BCI1 and BCI6 showed high standard deviation in their responses whereas the other subjects could lateralize reliably. Except for BCI1 and BCI3, subjects showed the maximum extent of lateralization beyond the largest naturally occurring ITD (see parameter LatRatio in Table 4).

The parameters of the individually fitted curves in Figure 4 are shown in Table 3. The extremely steep slope for BCI1 and BCI3 indicated a binary or step-function-type lateralization. Especially BCI3 responded consistently either between -8 and -7 or very close to 0 lateral units, that is, either left or middle. In contrast, the other four subjects had a continuously changing lateralization, indicated by a shallower slope ($P_{slope} = 0.66$ and 1.16 lateralization units). BCI6 showed a considerable bias to the left side ($ITD_c = 1.1$ ms). BCI4 and BCI8 had a centralized ITD percept ($ITD_c < 0.2$ ms).

BCI7 was a special case. Initially, this subject did not show ITD-based lateralization and a strong right bias for all ITDs (gray data points), despite careful loudness balancing. Repeating the lateralization experiment with an electrode shift (lowest frequency band presented to Electrode 1 left and Electrode 2 right; highest band presented to 11 left and 12 right) showed a little bit of bias reduction and a small but systematically ITD-dependent lateralization could be measured. See "Discussion" section for further reasoning.

Experiment 2: Angle Discrimination With Varying Access to ITD Cues

The results of the angle discrimination experiment with varying access to ITD cues are displayed in Figure 5. Nine panels in the first two columns show individual percent correct for each sound incident (-65 vs. -45° , -10° vs. 10° , and 45° vs. 65°) and each processing variant. These variants, employing the PP speech coding strategy are natural cues ("PP"), beamformer with target preserved cues ("PP + Beam") and modified ITD cues ("PP + ITD").

Across all variants, all subjects showed better angle discrimination performance when the sound was presented frontally and chance level performance when the sounds were presented from the sides. To assess the performance differences, a two-factor repeated measures analysis of

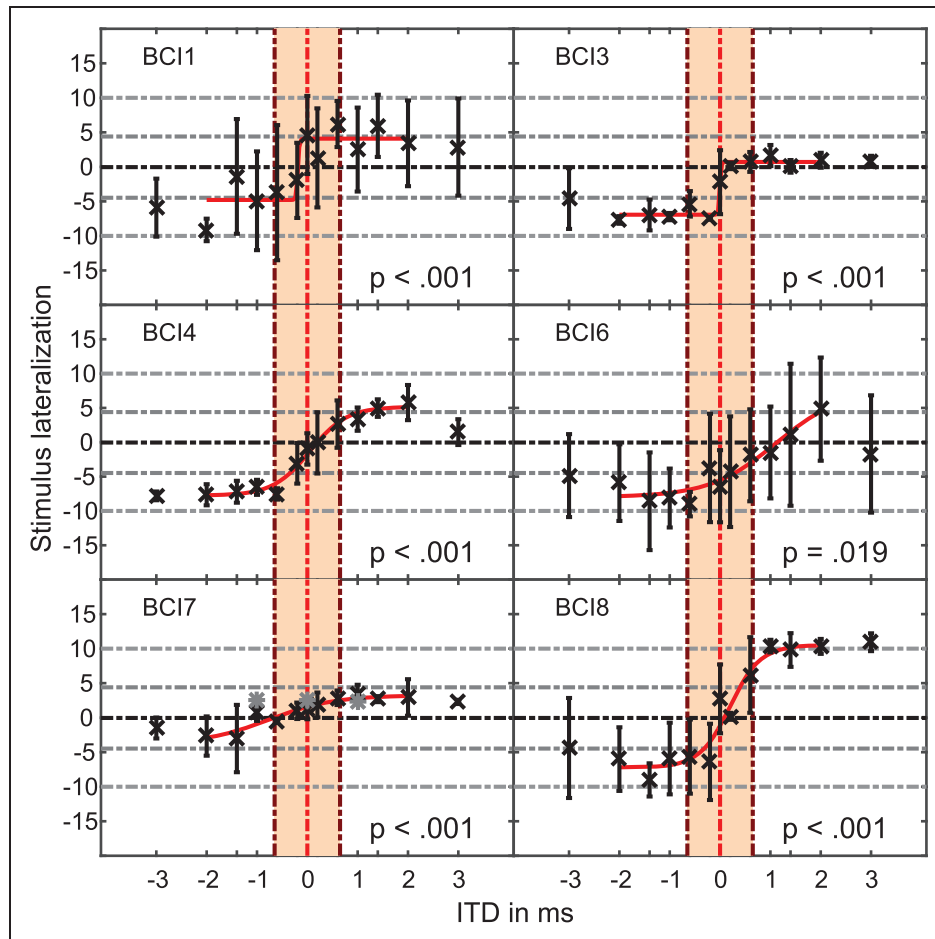


Figure 4. Perceived lateralization (mean + standard deviation) for the logatome /ana/ in quiet using the PP strategy as a function of ITD. Each panel shows data of one BCI subject. Negative ITD values on the abscissa denote a left leading, positive values a right leading ITD. The perceived lateralization indicated on a bar of graphical user interface is mapped to the two anchor points at -10 (left ear) and $+10$ (right ear). Dashed horizontal lines indicate the positions of ears, eyes, and nose (horizontal thick dashed line), and the vertical colored area shows the natural occurring range of ITDs ($\pm 600 \mu\text{s}$). The solid red line is the sigmoidal fit for each subject. p values in the left corner denote a significant difference between the pooled answers of the subjects for negative ITD values versus positive ITD values. BCI = bilateral cochlear implant; ITD = interaural time difference.

Table 4. Fit Parameter Results for the BCI Lateralization Task for Each Subject.

ID	P_l (Lat Units)	P_r (Lat Units)	ITD_c (msec)	P_{slope} (Lat Units/msec)	LatRatio (%)	SDev/Range (%)	SDev/Slope (Lat Units)
BCI1	-4.79	4.09	-0.19	27.95	100	66.85	0.21
BCI3	-6.93	0.71	-0.01	33.23	100	18.37	0.04
BCI4	-7.76	5.17	0.05	1.19	67.46	16.68	1.81
BCI6	-7.98	7.71	1.11	0.66	24.55	41.24	9.86
BCI7	-3.81	3.24	-0.76	0.67	32.95	24.71	2.61
BCI8	-7.2	10.5	0.2	1.43	72.25	19.76	2.44

Note. P_l and P_r denote the maximum extent of lateralization for left and right, respectively. ITD_c denotes the ITD in msec resulting in a centralized percept and P_{slope} is the slope in lateralization units per msec at ITD_c . LatRatio describes the percentage of lateralization range occurring within the largest naturally occurring ITD range ($\pm 600 \mu\text{sec}$), SDev/Range describes the mean standard deviation across the range ($P_l - P_r$). SDev/Slope is the JND in lateralization units for a noticeable difference in lateralization percept. JND = just noticeable difference; BCI = bilateral cochlear implant.

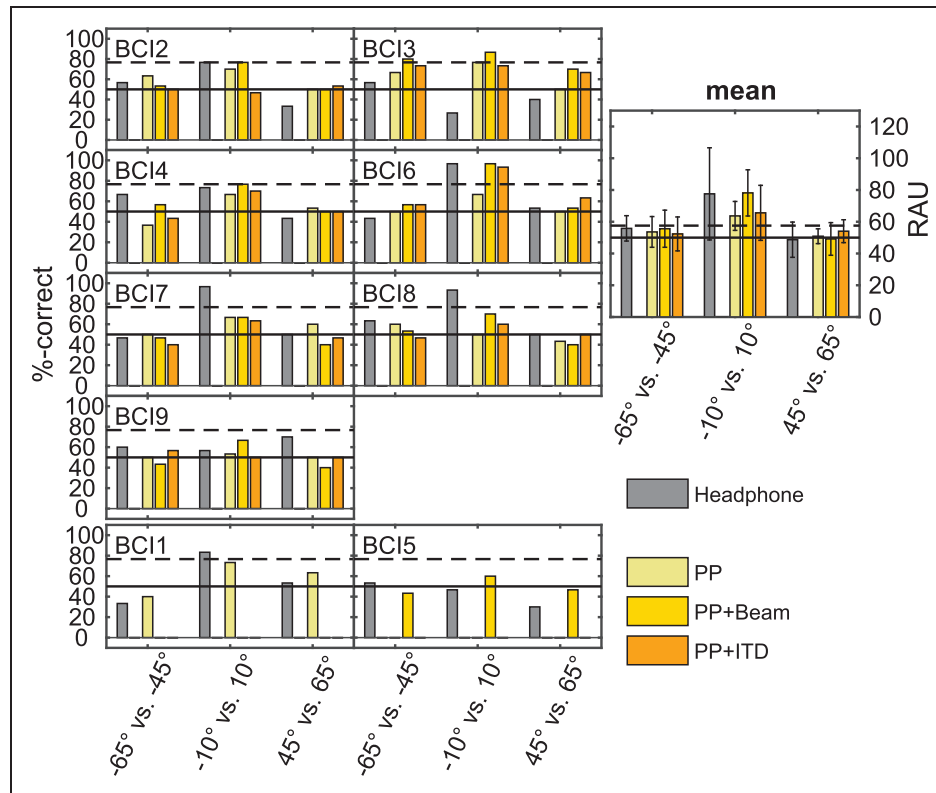


Figure 5. Left panels: Accuracy (percent correct) for the angle discrimination experiment with varying access to ITDs. Each panel denotes the results from one subject. The panel on the right is the RAU-corrected mean and standard deviation across the five subjects who completed all conditions (all except BCI1 and BCI5). Within each panel there are three sets of bars corresponding to the three sound incidents (-65° vs. -45° , -10° vs. 10° , and 45° vs. 65°). For each sound incidence, there are up to four bars indicating the different algorithms. From left to right: Headphone, PP, PP + Beam, and PP + ITD. The solid black line indicates the 50% chance level, and data above the dashed black line are significantly better from chance performance ($\alpha = 0.05/12$). Subjects BCI1 and BCI5 did not complete all conditions (see text for details). ITD = interaural time difference; BCI = bilateral cochlear implant; RAU = rationalized arcsine units.

variance (ANOVA) was conducted on the rationalized arcsine units (RAU, Studebaker, 1985, used here to avoid ceiling effects) transformed percent correct, with the factors being processing variant (headphone, PP, PP + Beam, PP + ITD) and azimuthal direction.

There was a tendency that “PP + Beam” provided higher correct rates than “PP” or “PP + ITD.” However, the ANOVA revealed that the factor processing variant was not significant (Greenhouse-Geisser corrected, $F[1.421, 8.52] = 0.505$, $p = .56$). The factor direction was significant, ($F[2, 12] = 12.506$, $p = .001$), showing higher scores for frontal direction than the two lateral directions. The interaction processing variant \times direction was not significant ($F[6, 36] = 1.744$, $p = .139$).

To test, if the subjects were able to discriminate the directions, hypothesis tests were run on the RAU-converted percent correct against a binomial distributed guessing-strategy, where the subject would randomly select “right” or “left.” The threshold to be significantly different from guessing after Bonferroni correction for multiple testing is given by the dashed line in Figure 5, right panel. All algorithms only showed significant

discrimination ability in the frontal direction, and for lateral sound incident angles, no significant discrimination ability was found.

BCI1 did not complete all angle discrimination measurements due to an illness preventing her from further participating in the study. However, BCI1’s performance is comparable to that of the other subjects for those conditions that were tested, with a fair score at frontal directions and close to guessing performance for lateral sound incident sets. BCI5 did not finish Experiment 2, as he was unable to achieve more than 63% correct with the highest scoring PP condition measured. BCI5 was neither able to discriminate sound coming from 20° apart nor able to lateralize stimuli differently as a function of ITD. Therefore, both tests were discontinued.

Experiment 3: Angle Discrimination With Varying Access to ILD Cues

The results of the angle discrimination experiment with varying access to ILD cues are displayed in Figure 6. Nine panels in the first two columns show individual

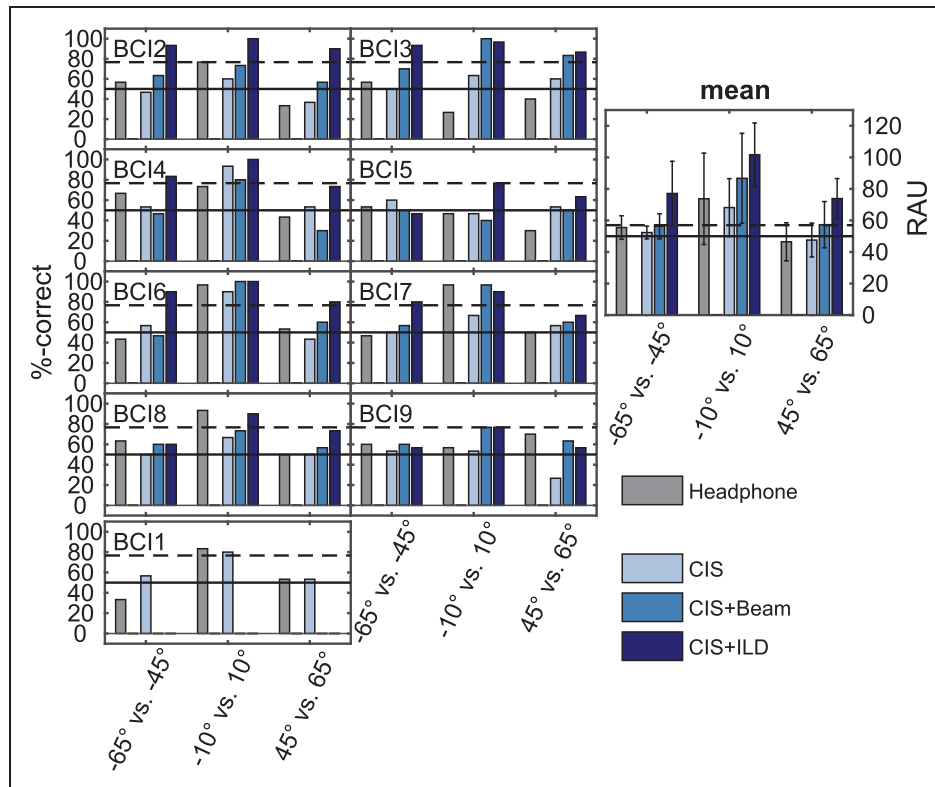


Figure 6. Left panels: Accuracy (percent correct) for the angle discrimination experiment with varying access to ILDs. Each panel denotes the results from one subject. The panel on the right is the RAU-corrected mean and standard deviation across the eight subjects who completed all conditions (all except BCI1). Within each panel there are three sets of bars corresponding to the three sound incidents (-65° vs. -45° , -10° vs. 10° , and 45° vs. 65°). For each sound incidence, there are up to four bars indicating the different algorithms. From left to right: Headphone, CIS, CIS + Beam, and CIS + ILD. The solid black line indicates the 50% chance level, data above the dashed black line is significantly better than chance performance ($\alpha = 0.05/12$). Subject BCI1 did not complete all conditions (see text for details). ILD = interaural level difference; BCI = bilateral cochlear implant; RAU = rationalized arcsine units; CIS = continuous interleaved sampling.

percent correct for each sound incident (-65° vs. -45° , -10° vs. 10° , or 45° vs. 65°) and processing variant. Analogous to Experiment 2, these CIS based processing variants were natural cues (“CIS”), beamformer with target preserved cues (“CIS + Beam”), and modified ILD cues (“CIS + ILD”).

A two-factor repeated-measures ANOVA was conducted on the RAU transformed % correct values with factors processing variant (Headphone, CIS, CIS + Beam, CIS + ILD) and azimuthal direction. Both factors, processing variant, ($F[2, 6] = 10.699$, $p < .011$), and azimuthal direction (Greenhouse-Geisser corrected, $F[1.128, 7.893] = 16.593$, $p = .003$) were significant, whereas the interaction term ($F[6, 42] = 0.739$, $p = .621$), was not significant.

Post hoc testing with pairwise t tests (Bonferroni corrected) revealed that the “CIS + ILD” algorithm was significantly better than the natural cues (“CIS”) condition ($p = .007$) showing improved angle discrimination when ILD cues were being enhanced due to the interaural cue enhancement algorithm. No significant difference was found between “CIS + ILD” and “CIS + Beam”

($p = .115$), or between CIS and “CIS + Beam” ($p = .575$). No significant difference was found between “Headphone” and CIS ($p = 1.00$) or “Headphone” and “CIS + Beam” ($p = 1.00$). “Headphone” compared to “CIS + ILD” approaches significance after Bonferroni correction ($p = .053$). In addition, all subjects showed statistically significant better angle discrimination performance when the sound was presented frontally than laterally (post hoc testing of the factor direction: frontal vs. left: $p = .020$, frontal vs. right: $p = .009$). Presenting from the right or the left side had no significant difference (left vs. right: $p = .155$).

The same discrimination analysis as in Experiment 2 was also done for this experiment. The algorithm “CIS + ILD” showed significant discrimination ability in all directions. The other algorithms only showed significant discrimination ability in the frontal direction.

Discussion

This proof-of-concept study tested lateralization in quiet and angle discrimination performance in noise with different variants of a novel BCI algorithm.

The three hypotheses that inspired the BCI algorithm design were (1) ITDs applied to the output of a low-rate speech coding strategy should cause a lateralized percept; (2) using (quasi) stationary interaural differences corresponding to the dominant source direction is expected to be preferential to most CI listeners in most complex listening situations; and (3) modified—generally enlarged—interaural differences should improve sound localization compared to natural cues, at least after acclimatization.

Experiment 1 directly tested the first hypothesis. Six out of nine BCI subjects were able to lateralize speech solely based on ITD and overall lateralization is similar to the five-electrode stimulation from Churchill et al. (2014). Lateralization range and variance are slightly worse than 100 pps single channel data from Baumgärtel et al. (2017). In contrast, Egger, Majdak, and Laback (2016) report improved ITD sensitivity, when moving from single-electrode to dual-electrode stimulation. There are two possible reasons for the absent improvement in this study: First, the pulse rates differ across electrodes, which may cause more ITD interference than the regularly interleaved constant rate pulse trains in Egger et al. (2016). They are also on average > 100 pps which is expected to reduce performance. The second reason speaks to a major limitation of all bilateral multichannel studies: There is no systematic compensation for interaural electrode mismatches, despite evidence that this is particularly important to provide perceptually exploitable ITDs to the auditory system (e.g., Hu & Dietz, 2015; Kan et al., 2013). To date, no practical method is available to identify all electrode pairs in a reasonable amount of time, nor is it straightforward how to reprogram the frequency allocation tables, once a nonconstant mismatch is identified. We expect that for the speech stimuli employed here, an exact interaural matching is even more critical than previously demonstrated in the single electrode pair studies (Hu & Dietz, 2015; Kan et al., 2013). Even more so, with the employed gammatone filterbank, frequency-dependent latencies are inevitable, which can result in up to 1 msec latency difference for neighboring channels. The combination of interaural mismatch and a gammatone filterbank will induce large ITD offsets corresponding to the latency difference. Only when further combined with a low-rate speech coding strategy which allows for ITD sensitivity, this effect becomes perceptually relevant. We only became aware of this complex interplay after analyzing the unsuccessful attempts of BCI5 and the strong offset of BCI6. In the following subject BCI7, we again encountered a bias and had prepared a constant one-electrode shift to crudely compensate for the presumed interaural mismatch. Indeed the shifted version provided both a slightly reduced bias (in line with the gammatone-induced ITD hypothesis) and a small but systematically

ITD-dependent lateralization. The remaining bias indicates that the mismatch may have been >1 electrode but we are not yet prepared enough to measure and compensate any mismatch across the array, despite its increasing importance and apparently high prevalence (see Aronoff, Stelmach, Padilla, & Landsberger, 2016; Hu & Dietz, 2015). Of the three participants (BCI1, BCI 3, and BCI4) which previously participated in a study identifying electrode pairs with optimal ITD sensitivity (Hu & Dietz, 2015) only BCI4 had no mismatch in the previous study and is also the best performing subject of these three in the extent of lateralization experiment (Figure 4). Switched-off electrodes or not fully inserted electrode arrays (in the case of BCI5, two extra-cochlea electrodes on the right side) also can lead to a large interaural mismatch, and offer a possible explanation for BCI5's poor performance.

The question arises if this improvement in lateralization comes at the cost of speech intelligibility, as it is the case when pulse rate is reduced in CIS (Churchill et al., 2014). We therefore tested speech intelligibility in noise with CIS and the low-rate PP coding strategy without any further pre- or postprocessing (Figure 7).

Speech reception thresholds (SRTs) of the PP strategy were significantly poorer by 9.8 dB (i.e., higher) than SRTs measured using headphone or CIS (Wilcoxon

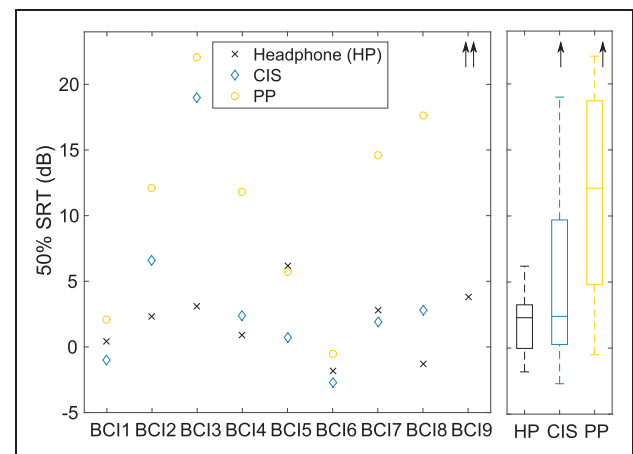


Figure 7. Left panel: Individual 50% SRTs in dB for the German matrix sentence test (OISa, Wagener, Kühnel, & Kollmeier, 1999) in stationary noise for each subject for listening with Headphone over their clinical devices (HP), or with research coils and our selfimplemented CIS and PP coding strategies. The panel on the right shows the same data as boxplots across subjects. Speech level was held constant at 65 dB SPL and the noise was varied adaptively, starting at 45 dB SPL (+20 dB SNR). For familiarization, the HP control condition was always measured first. Arrows indicate not measurable thresholds (>40 dB SNR) for BCI9 for CIS and PP coding strategy. BCI = bilateral cochlear implant; SNR = signal-to-noise ratio; SRTs = speech reception thresholds; PP = peak picking; CIS = continuous interleaved sampling.

signed rank test with Bonferroni correction applied to p values: $p = .023$ for both comparisons). SRTs with headphones and with CIS were not significantly different from each other (Headphone vs. CIS: $p = .86$). The variability across subjects is relatively high, for example, BCI3 shows an SRT of 3 dB using the headphones, but shows SRTs of around 20 dB for the two other coding strategies, whereas BCI1 shows an SRT-range from 4 dB across all tested coding strategies. A possible reason for the poorer SRT with PP is that it is unfamiliar to the subjects. Based on reports about CI users equipped with the similar FAST strategy (Smith, Parkinson, & Krishnamoorthi, 2013), it seems reasonable that subjects can achieve a similar SRT in these tests as with CIS or any other strategy after a few weeks of training. The PP strategy may offer the additional advantage over FAST that, similar to FS4, it provides temporal fine structure-like information in the apical canals. Another possible reason for the performance difference between CIS and PP may also be that all subjects tested in this study also participated in another study that tested speech intelligibility task using the same CIS implementation. Thus the BCI users got more training for the CIS coding strategy than for the PP strategy. Last, as discussed earlier, for PP, interaural electrode matching is more important, which was not included in this study. In an optimistic outlook with acclimatization and electrode pairing, most listeners may perform similar with PP as with CIS similar to a previous comparison of FAST and ACE (Croghan & Smith, 2015).

At frontal locations, delivering coherent binaural cues using the steered binaural beamformer improved angle discrimination with CIS, confirming hypothesis 2. The DOA is working robustly and precisely down to -10 dB SNR (Adiloglu et al., 2015). Therefore, localization performance with a beamformer is expected to be SNR-independent and even possible at negative SNRs. Consequently, effects are expected to be even larger at worse SNR. However, in a real setting the DOA is going to be a source of errors. It remains to be tested in which conditions and for which subject groups such an algorithm is the preferred choice. The lack of improvement with the beamformer at lateral angles is arguably due to the almost absent discrimination ability at those angles even in quiet performance (inferred from, e.g., Kerber & Seeber, 2012). The quiet performance imposes the upper performance limit but was not tested here because of time constraints.

As with hypothesis 1, testing hypothesis 3 was limited by acute testing. It is also addressed a bit indirectly, because of testing time constraints we did not include a control condition with the full BCI algorithm and with nonenhanced binaural cues. For Experiment 3 (CIS), this control condition would have been virtually identical to the CIS + Beam variant: natural but coherent ILDs.

The minor differences are that CIS + Beam contains envelope ITDs and that its output coherence is not a perfect one. These differences are not expected to be of perceptual relevance and led us to use CIS + Beam as the ILD control condition for hypothesis 3. In line with Francart et al. (2011), we found a significant improvement with linearized ILD mapping. This was expected at frontal angles where the generally enlarged ILDs obviously improve discrimination. Fortunately, the enlargement was not at the cost of lateral performance where a trend of improved performance is still observed in both hemispheres. Further improvements are expected with acclimatization to the new mapping (Keating et al., 2016) or with individualized optimization (Brown, 2018). Again, a certain trade-off with speech intelligibility remains: The more DR is used for ILD, the less remains for the signals amplitude modulation. That said, the linear ILD maps appear to make optimal use of their attributed DR.

The ITD part of hypothesis 3 was more difficult to incorporate in this proof-of-concept. It is clear that setting all ILDs to zero will not be the best choice for an application of the BCI algorithm. Nevertheless, we chose this variant, only to find out if pure ITD-based discrimination is possible for both zero and far-from-zero references in noise. Despite all the above-mentioned limitations and challenges, subjects were performing above chance at frontal angles. We did not include a control condition with natural ITDs only (full coherence but no enhancement and no ILDs). The tested discrimination of $\pm 10^\circ$ with enhancement is virtually identical to discriminating $\pm 20^\circ$ without enhancement; thus, at frontal angles, the enhancement factor should translate into a d' factor. However, such a control condition or a systematic testing of the decline of ITD discrimination as a function of baseline ITD would be required to study the optimal ITD enhancement factor.

Future studies have to show if acclimatization, individualization, and a smart combination of both ITD and ILD will allow to further improve both speech intelligibility and localization discrimination. In case of ITD and ILD combination, the mapping slopes can likely be a bit shallower at central angles, leaving more “spatial dynamic range” to lateral angles if this is desired.

Overall, the proposed combination of beamformer and artificial enhancement of binaural cues from a target sound may be more likely acceptable for BCI users than binaural beamformers are for bilateral hearing aid users. The concentration toward only one sound source is not recommended as a standard setting, and in particular not for children who often require a more complete input of the surrounding auditory space. A future refinement of the localization enhancement algorithm may be able to provide more than one direction for those BCI users who can exploit switching interaural

cues and who can localize two or more concurrent sounds (Bernstein, Goupell, Schuchman, Rivera, & Brungart, 2016). Irrespective of the actual algorithm employed in this study, identifying the subject-desired target direction is its own field of research. Among the different approaches are signal-driven DOA estimators (e.g., Adiloglu et al., 2015), brain-computer interfaces (Mirkovic, Bleichner, De Vos, & Debener, 2016), gesture control to select the target (Grimm, Luberadzka, Müller, & Hohmann, 2016), and eye tracking to employ the visual look-direction (Wendt, Brand, & Kollmeier, 2014). Any of these evolving strategies can be used in combination with the localization enhancement algorithm used here to steer the beamformer and to apply the target azimuth corresponding ILDs and ITDs.

With respect to the signal processing, the presented approach is fundamentally different from all other algorithms in the way that interaural cues are imposed on a common electrodogram shared in each ear. Previous algorithms manipulate interaural cues by altering the acoustic signals at the two processors (Brown, 2018; Francart et al., 2011; Wouters et al., 2013). As part of a relatively long chain of signal processing stages, some of the acoustic cues are omitted, others are distorted, jittered, or compressed (see Dietz, 2016, for a review). At any stage, ILDs or ITDs have always been encoded implicitly through differences between two channels. In contrast, the proposed algorithm extracts the interaural cues from the temporally and spectrally rich input. It then continues with processing only a single acoustic signal and finally applies the desired interaural cues on the single signal at the latest possible stage, bypassing the above-mentioned issues. A similar concept is also used for multichannel low-bitrate sound coding, where instead of storing two channels as needed for a binaural signal, only one master channel is stored together with sparse information about the interaural differences (Faller & Baumgarte, 2003).

Summary and Conclusions

This study investigated how ITDs and ILDs that were delivered using speech signals processed by a standard CIS or a low pulse rate speech coding strategy can contribute to sound lateralization and localization in bilateral CI users with binaurally linked devices. These investigations were accomplished by testing different variants of a novel BCI algorithm. The following summarizes the results:

1. Six out of nine BCI subjects were able to lateralize speech sounds in quiet with the low-rate coding strategy using only ITD cues in the pulse timing. Four out of six BCI subjects needed larger ITDs than the physiological range to achieve maximum lateralization.

2. The low-rate strategy with beamforming, ITD-enhancement and ILD omission allowed for above-chance angle discrimination in noise for frontal direction. As expected, it was not as good as the natural combination of ILDs and ITDs with beamforming.
3. A binaural beamformer steered toward the dominant sound source in combination with ILD-enhancement improved angle discrimination in noise for both frontal and lateral azimuthal directions, when using the standard CIS strategy.

Going beyond beamforming and providing interaural cues with better interaural coherence was possible by converting a monophonic acoustic signal to an electrodogram and then applying (quasi) constant ITDs or ILDs on the pulses. This may provide multiple advantages over acoustic signal manipulations and the interaurally fully coherent output may be beneficial for sound localization with CIs.

While it is too early for suggesting an ideal BCI strategy, the beamformer with linear ILD mapping is a promising candidate for a binaurally linked CI system. In the long term, if interaural electrode matching and an individual performance characterization are available, a PP-like speech coding with modestly enhanced ITD and linearized ILDs can be applied to some or all channels. Together with the inherent beamformer, such a strategy may provide both improved speech intelligibility and improved localization abilities, especially at low SNRs.

Authors' Note

Portions of this study were presented at the 40th annual mid-winter meeting of the Association for Research in Otolaryngology (ARO), Baltimore, Maryland, USA, and at the 2017 Conference on implantable prostheses, Granlibakken, California, USA.

Acknowledgments

The authors thank the research participants for their time and valuable feedback, Stefan Strahl for providing support with the RIB II, Anja Eichenauer for helping to program and adjust the coding strategies used here, and Regina Baumgärtel for support with the lateralization experiment. They also thank the associate editor Christopher Plack and two anonymous reviewers for their invaluable feedback on earlier drafts.

Declaration of Conflicting Interests

The authors report no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this

article: This study was supported by the European Union's Seventh Framework Programme FP7/2007–2013 under the Advancing Binaural Cochlear Implant Technology (ABCIT) grant agreement no. 304912, the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 716800 (ERC Starting Grant to Mathias Dietz), DFG JU2858/2-1, the Canada Research Chairs program, and the European Regional Development Fund jointly with the Federal State of Lower Saxony under VIBHear grant agreement no. ZW6-85003643.

References

- Adiloglu, K., Kayser, H., Baumgärtel, R. M., Rennebeck, S., Dietz, M., & Hohmann, V. (2015). A binaural steering beamformer system for enhancing a moving speech source. *Trends in Hearing, 19*, 1–13. doi:10.1177/2331216515618903
- Aronoff, J. M., Stelmach, J., Padilla, M., & Landsberger, D. M. (2016). Interleaved processors improve cochlear implant patients' spectral resolution. *Ear and Hearing, 37*, e85–e90. doi:10.1097/AUD.0000000000000249
- Baumgärtel, R. M., Hu, H., Kollmeier, B., & Dietz, M. (2017). Extent of lateralization at large interaural time differences in simulated electric hearing and bilateral cochlear implant users. *The Journal of the Acoustical Society of America, 141*(4), 2338–2352. doi:10.1121/1.4979114
- Baumgärtel, R. M., Krawczyk-Becker, M., Marquardt, D., Völker, C., Hu, H., Herzke, T., ... Dietz, M. (2015). Comparing binaural pre-processing strategies I: Instrumental evaluation. *Trends in Hearing, 19*, 1–16. doi:10.1177/2331216515617916
- Bernstein, J. G., Goupell, M. J., Schuchman, G. I., Rivera, A. L., & Brungart, D. S. (2016). Having two ears facilitates the perceptual separation of concurrent talkers for bilateral and single-sided deaf cochlear implantees. *Ear and Hearing, 37*, 289–302. doi:10.1097/AUD.0000000000000284
- Best, V., Roverud, E., Streeter, T., Mason, C. R., & Kidd, G. (2017). The benefit of a visually guided beamformer in a dynamic speech task. *Trends in Hearing, 21*, 1–11. doi:10.1177/2331216517722304
- Bregman, A. S. (1999). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Brown, C. A. (2016). System and method for enhancing the binaural representation for hearing-impaired subjects. *Patent Number 20160330551*. Retrieved from <http://www.freepatentsonline.com/y2016/0330551.html>
- Brown, C. A. (2018). Corrective binaural processing for bilateral cochlear implant patients. *PLOS ONE, 13*, e0187965. doi:10.1371/journal.pone.0187965
- Busby, P. A., Whitford, L. A., Blamey, P. J., Richardson, L. M., & Clark, G. M. (1994). Pitch perception for different modes of stimulation using the Cochlear multiple-electrode prosthesis. *The Journal of the Acoustical Society of America, 95*(5), 2658–2669. doi:10.1121/1.409835
- Churchill, T. H., Kan, A., Goupell, M. J., & Litovsky, R. Y. (2014). Spatial hearing benefits demonstrated with presentation of acoustic temporal fine structure cues in bilateral cochlear implant listeners. *The Journal of the Acoustical Society of America, 136*(3), 1246–1256. doi:10.1121/1.4892764
- Croghan, N. B. H., & Smith, Z. M. (2015). *Cognitive factors and adaptation to a novel cochlear-implant coding strategy*. Presented at the 42 Annual Scientific and Technology Conference of the American Auditory Society, Scottsdale, AZ.
- Dietz, M. (2016). Models of the electrically stimulated binaural system: A review. *Network: Computation in Neural Systems, 27*(2–3), 186–211. doi:10.1080/0954898X.2016.1219411
- Dietz, M., & Backus, B. (2015, 22 June). Sound processing for a bilateral cochlear implant system. *EU Patent Application No. 15173203.9*.
- Dietz, M., Ewert, S. D., & Hohmann, V. (2011). Auditory model based direction estimation of concurrent speakers from binaural signals. *Speech Communication, 53*, 592–605. doi:10.1016/j.specom.2010.05.006
- Egger, K., Majdak, P., & Laback, B. (2016). Channel interaction and current level affect across-electrode integration of interaural time differences in bilateral cochlear-implant listeners. *Journal of the Association of Research in Otolaryngology, 17*, 55–67. doi:10.1007/s10162-015-0542-8
- Faller, C., & Baumgarte, F. (2003). Binaural cue coding-part II: Schemes and applications. *IEEE Transactions on Speech and Audio Processing, 11*, 520–531. doi:10.1109/TSA.2003.818108
- Francart, T., Lensen, A., & Wouters, J. (2011). Enhancement of interaural level differences improves sound localization in bimodal hearing. *The Journal of the Acoustical Society of America, 130*(5), 2817–2826. doi:10.1121/1.3641414
- Grimm, G., Luberadzka, J., Müller, J., & Hohmann, V. (2016). *A simple algorithm for real-time decomposition of first order Ambisonics signals into sound objects controlled by eye gestures*. Presented at the Interactive Audio Systems Symposium, University of York, UK.
- Harczos, T. (2015). *Cochlear implant electrode stimulation strategy based on a human auditory model* (PhD thesis). Technische Universität Ilmenau, DE.
- Hartmann, W. M., & Rakerd, B. (1989). On the minimum audible angle: A decision theory approach. *Journal of the American Academy of Audiology, 85*(5), 2031–2041. doi:10.1121/1.397855
- Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., Schosser, H., Garnham, C., & Anderson, I. (2006). MED-EL cochlear implants: State of the art and a glimpse into the future. *Trends in Amplification, 10*(4), 201–219. doi:10.1177/1084713806296720
- Hohmann, V. (2002). Frequency analysis and synthesis using a gammatone filterbank. *Acta Acustica United with Acustica, 88*, 433–442.
- Hu, H., & Dietz, M. (2015). Comparison of interaural electrode pairing methods for bilateral cochlear implants. *Trends in Hearing, 19*, 1–22. doi:10.1177/2331216515617143
- Hu, H., Dietz, M., Williges, B., & Ewert, S. D. (2018). Better-ear glimpsing with symmetrically-placed interferers in bilateral cochlear implant users. *The Journal of the Acoustical Society of America, 143*, 2128–2141. doi:10.1121/1.5030918
- Kan, A., Stoelb, C., Litovsky, R. Y., & Goupell, M. J. (2013). Effect of mismatched place-of-stimulation on binaural fusion and lateralization in bilateral cochlear-implant

- users. *The Journal of the Acoustical Society of America*, 134(4), 2923–2936. doi:10.1121/1.4820889
- Kayser, H., & Anemüller, J. (2014). *A discriminative learning approach to probabilistic acoustic source localization*. Presented at 14th International Workshop on the Acoustic Signal Enhancement (IWAENC), IEEE, pp. 99–103. New York, NY: IEEE.
- Kayser, H., Ewert, S. D., Anemüller, J., Rohdenburg, T., Hohmann, V., & Kollmeier, B. (2009). Database of multichannel in-ear and behind-the-ear head-related and binaural room impulse responses. *EURASIP Journal on Advances in Signal Processing*, 2009(1), 298605: 1–10. doi: 10.1155/2009/298605.
- Keating, P., Rosenior-Patten, O., Dahmen, J.C., Bell, O. & King, A.J., 2016. Behavioral training promotes multiple adaptive processes following acute hearing loss. *eLife* 5. doi: 10.7554/eLife.12264
- Kelvasa, D., & Dietz, M. (2015). Auditory model-based sound direction estimation with bilateral cochlear implants. *Trends in Hearing*, 19, 1–16. doi:10.1177/23312165151616378
- Kerber, I. S., & Seeber, I. B. U. (2012). Sound localization in noise by normal-hearing listeners and cochlear implant users. *Ear and Hearing*, 33(4), 445–457. doi:10.1097/AUD.0b013e318257607b
- Kokkinakis, K., Hazrati, O., & Loizou, P. C. (2011). A channel-selection criterion for suppressing reverberation in cochlear implants. *The Journal of the Acoustical Society of America*, 129, 3221–3232. doi:10.1121/1.3559683
- Kollmeier, B., & Peissig, J. (1990). Speech intelligibility enhancement by interaural magnification. *Acta Oto-Laryngologica. Supplementum*, 469, 215–223.
- Laback, B., Egger, K., & Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hearing Research*, 322, 138–150. doi:10.1016/j.heares.2014.10.004
- Licklider, J. C. R. (1948). The influence of interaural phase relations upon the masking of speech by white noise. *The Journal of the Acoustical Society of America*, 20(2), 150–159. doi:10.1121/1.1906358
- Lopez-Poveda, E. A., Eustaquio-Martín, A., Stohl, J. S., Wolford, R. D., Schatzer, R., & Wilson, B. S. (2016). A binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. *Ear and Hearing*, 37, 138–148. doi:10.1097/AUD.0000000000000273
- Luts, H., Eneman, K., Wouters, J., Schulte, M., Vormann, M., Buechler, M., & Spriet, A. (2010). Multicenter evaluation of signal enhancement algorithms for hearing aids. *The Journal of the Acoustical Society of America*, 127(3), 1491–1505. doi:10.1121/1.3299168
- Marquardt, D., Hohmann, V., & Doclo, S. (2015). Interaural coherence preservation in multi-channel wiener filtering-based noise reduction for binaural hearing aids. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(12), 2162–2176. doi:10.1109/TASLP.2015.2471096
- Mirkovic, B., Bleichner, M. G., De Vos, M., & Debener, S. (2016). Target speaker detection with concealed EEG around the ear. *Frontiers in Neuroscience*, 10(349), 1–11. doi:10.3389/fnins.2016.00349
- Monaghan, J. J., & Seeber, B. U. (2016). A method to enhance the use of interaural time differences for cochlear implants in reverberant environments. *The Journal of the Acoustical Society of America*, 140, 1116–1129. doi:10.1121/1.4960572
- Nobbe, A., Schleich, P., Zierhofer, C., & Nopp, P. (2007). Frequency discrimination with sequential or simultaneous stimulation in MED-EL cochlear implants. *Acta Oto-Laryngologica*, 127(12), 1266–1272. doi:10.1080/00016480701253078
- Nogueira, W., Büchner, A., Lenarz, T., & Edler, B. (2005). A psychoacoustic NofM-type speech coding strategy for cochlear implants. *EURASIP Journal on Applied Signal Processing*, 2005, 3044–3059. doi:10.1155/ASP.2005.3044
- Roman, N., Wang, D., & Brown, G. J. (2003). Speech segregation based on sound localization. *The Journal of the Acoustical Society of America*, 114(4), 2236–2252. doi:10.1121/1.1610463
- Rowland, R. C., & Tobias, J. V. (1967). Interaural intensity difference limen. *Journal of Speech and Hearing Research*, 10(4), 745–756. doi:10.1121/1.1939795
- Seeber, B. U., & Fastl, H. (2008). Localization cues with bilateral cochlear implants. *The Journal of the Acoustical Society of America*, 123, 1030–1042. doi:10.1121/1.2821965
- Smith, Z. M. (2014). Hearing better with interaural time differences and bilateral cochlear implants. *The Journal of the Acoustical Society of America*, 135(4), 2190–2191. doi:10.1121/1.4877139
- Smith, Z. M., & Delgutte, B. (2007). Sensitivity to interaural time differences in the inferior colliculus with bilateral cochlear implants. *Journal of Neuroscience*, 27(25), 6740–6750. doi:10.1523/JNEUROSCI.0052-07.2007
- Smith, Z. M., Parkinson, W. S., & Krishnamoorthi, H. (2013). *Efficient coding for auditory prostheses*. Conference on implantable auditory prostheses, Lake Tahoe, CA, 14–19 July.
- Studebaker, G. A. (1985). A “rationalized” arcsine transform. *Journal of Speech and Hearing Research*, 28, 455–462. doi:10.1044/jshr.2803.455
- van Hoesel, R. J. M. (2007). Sensitivity to binaural timing in bilateral cochlear implant users. *The Journal of the Acoustical Society of America*, 121(4), 2192–2206. doi:10.1121/1.2537300
- van Hoesel, R. J. M., & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *The Journal of the Acoustical Society of America*, 113(3), 1617doi:10.1121/1.1539520
- van Schoonhoven, J., Sparreboom, M., van Zanten, B. G. A., Scholten, R. J. P. M., Mylanus, E. A. M., Dreschler, W. A., ... Maat, B. (2013). The effectiveness of bilateral cochlear implants for severe-to-profound deafness in adults: A systematic review. *Otology & Neurotology*, 34(2), 190–198. doi:10.1097/MAO.0b013e318278506d
- Wagener, K. C., Kühnel, V., & Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests [Development and evaluation of a German sentence test I: Design of the Oldenburg sentence test]. *Zeitschrift Für Audiologie*, 38(1), 4–95.
- Wendt, D., Brand, T., & Kollmeier, B. (2014). An eye-tracking paradigm for analyzing the processing time of sentences with different linguistic complexities. *PLoS ONE*, 9, e100186doi:10.1371/journal.pone.0100186

- Wesker, T., Meyer, B., Wagener, K., Anemüller, J., Mertins, A., & Kollmeier, B. (2005). *Oldenburg logatome speech corpus (OLLO) for speech recognition experiments with humans and machines*. Ninth European Conference on Speech Communication and Technology, Barcelona, Spain.
- Wilson, B. S., Finley, C., Lawson, D., Wolford, R., Eddington, D. K., & Rabinowitz, W. (1991). Better speech recognition with cochlear implants. *Nature*, 352, 236–238. doi:10.1038/352236a0
- Wouters, J., Doclo, S., Koning, R., & Francart, T. (2013). Sound processing for better coding of monaural and binaural cues in auditory prostheses. *Proceedings of the IEEE*, 101(9), 1986–1997. doi:10.1109/JPROC.2013.2257635