

# Perception of Interaural Phase Differences With Envelope and Fine Structure Coding Strategies in Bilateral Cochlear Implant Users

Stefan Zirn<sup>1,2</sup>, Susan Arndt<sup>1</sup>, Antje Aschendorff<sup>1</sup>, Roland Laszig<sup>1</sup>, and Thomas Wesarg<sup>1</sup>

Trends in Hearing  
2016, Vol. 20: 1–12  
© The Author(s) 2016  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/2331216516665608  
tia.sagepub.com  


## Abstract

The ability to detect a target signal masked by noise is improved in normal-hearing listeners when interaural phase differences (IPDs) between the ear signals exist either in the masker or in the signal. To improve binaural hearing in bilaterally implanted cochlear implant (BiCI) users, a coding strategy providing the best possible access to IPD is highly desirable. In this study, we compared two coding strategies in BiCI users provided with CI systems from MED-EL (Innsbruck, Austria). The CI systems were bilaterally programmed either with the fine structure processing strategy FS4 or with the constant rate strategy high definition continuous interleaved sampling (HDCIS). Familiarization periods between 6 and 12 weeks were considered. The effect of IPD was measured in two types of experiments: (a) IPD detection thresholds with tonal signals addressing mainly one apical interaural electrode pair and (b) with speech in noise in terms of binaural speech intelligibility level differences (BILD) addressing multiple electrodes bilaterally. The results in (a) showed improved IPD detection thresholds with FS4 compared with HDCIS in four out of the seven BiCI users. In contrast, 12 BiCI users in (b) showed similar BILD with FS4 ( $0.6 \pm 1.9$  dB) and HDCIS ( $0.5 \pm 2.0$  dB). However, no correlation between results in (a) and (b) both obtained with FS4 was found. In conclusion, the degree of IPD sensitivity determined on an apical interaural electrode pair was not an indicator for BILD based on bilateral multielectrode stimulation.

## Keywords

interaural phase differences, binaural speech unmasking, bilateral cochlear implants, fine structure processing

Date received: 15 December 2015; revised: 29 July 2016; accepted: 29 July 2016

## Introduction

Listening with two ears enables normal-hearing (NH) listeners to perceive interaural level differences (ILDs) and interaural time differences (ITDs). In amplitude-modulated stimuli, ITD can be assessed by looking at either the waveform fine structure, which yields  $ITD_{FS}$ , or at the waveform envelope, which yields  $ITD_{ENV}$ . ITD occurring in a periodic stimulus, such as a pure tone, can be referred to as the interaural phase difference (IPD) and is typically indicated in degrees. Analysis of these interaural cues in the auditory system is crucial for sound-source localization and for binaural unmasking. An overview of the role of these signal cues in spatial and binaural hearing can be found in Blauert (1997).

In binaural listening conditions with a target signal and a masker presented simultaneously to both ears, ILD and IPD facilitate target detection provided that the ILD and IPD of the target at the two ears are not the same as those of the masker (Bronkhorst & Plomp, 1988). Disparate ILD and IPD between target and masker occur in (free-field) situations where the sound source of the target and the sound source of the masker

<sup>1</sup>Department of Oto-Rhino-Laryngology of the Medical Center, University of Freiburg, Germany

<sup>2</sup>University of Applied Sciences Offenburg, Germany

### Corresponding author:

Stefan Zirn, University of Applied Sciences Offenburg, Badstraße 24, 77652 Offenburg, Germany.

Email: stefan.zirn@hs-offenburg.de



are not coincident in space, leading to improved detection and discrimination of target signals, including speech, relative to situations where target and masker originate from the identical position in space.

To investigate this effect with headphones in NH listeners, IPD can be implemented by phase inversion of either the target or the masker in one ear. In the condition with IPD, lower detection thresholds (for e.g., pure tones in noise) or better speech reception thresholds (SRTs; for e.g., speech in noise) are obtained than in the condition without IPD (Licklider, 1948). This type of binaural unmasking is often called binaural masking level differences (BMLD—Colburn, Shinn-Cunningham, Kidd, & Durlach, 2006; Moore, 2012).

Sound source localization and binaural unmasking of speech of bilaterally implanted cochlear implant (BiCI) users are impaired compared with those of NH listeners. The main limitation may arise from the limited availability of interaural timing cues in terms of IPD, whereas BiCI users can perceive ILD with a considerably high precision when using their clinical devices (Kerber & Seeber, 2012; Seeber & Fastl, 2008; van Hoesel et al., 2008).

However, Laback, Majdak, and Baumgartner (2007) showed that IPD are perceivable at low pulse rates by some BiCI users using synchronized electric pulse trains with and without onset or offset differences presented on pitch-matched interaural electrode pairs. Furthermore, BMLD have been successfully demonstrated in BiCI users. BMLD refers to signal detection as opposed to speech recognition. Bilateral stimulation was usually realized using constant high-rate pulse trains. BMLD values of 5 to 11 dB were found using one interaural electrode pair for stimulation (Goupell & Litovsky, 2015; Long, Carlyon, Litovsky, & Downs, 2006) and around 3 dB for three simultaneously activated interaural electrode pairs (Van Deun et al., 2011).

The studies described earlier give rise to optimism for effective interaural cue coding which could result in real-life benefits in terms of binaural unmasking of speech for BiCI users. Most current CI stimulation strategies solely encode the envelope into amplitude variations of biphasic current pulses transferred to the assigned intracochlear electrode contacts. However, the interaural differences underlying binaural unmasking are present in the envelope and in the fine structure of sound signals. Thus, successful transmission of both envelope and fine structure would be required within a clinical signal coding strategy to enable normal levels of binaural unmasking in BiCI users.

Despite the potentially high relevance of this topic, to our knowledge, only one study has investigated binaural unmasking of speech in noise in BiCI users: van Hoesel et al. (2008), who measured SRTs in four BiCI users, provided with Nucleus 24 devices bilaterally. SRTs in

diotic noise were compared in listening conditions with and without a 700- $\mu$ s interaural delay in the speech signals. They investigated binaural unmasking using three different coding strategies implemented on a research processor with a familiarization period of 4 weeks. In one of their coding strategies, they implemented an approach to process the temporal fine structure. This peak-derived *timing* strategy located positive peaks in the fine timing of signals at each filter-band output and then stimulated the associated electrode for each band at times corresponding to those peaks (van Hoesel, 2004). The authors found no binaural unmasking with bilateral synchronized stimulation, not even with the peak-derived timing strategy.

In the present study, the effect of encoding temporal fine structure on binaural unmasking was investigated in BiCI users bilaterally provided with CI systems of MED-EL. The coding strategy FS4, clinically available on those devices, processes envelope and fine structure cues. Laback, Egger, and Majdak (2015) termed FS4 as a promising candidate for a clinical strategy providing temporal fine structure cues. With FS4, pulse timing is triggered by zero-crossings of the bandpass filter outputs assigned to the four apical (low frequency) CI channels resulting in a nonconstant interpulse interval. The amplitude of the pulses is determined by the envelope of the bandpass filter output. Thus, both fine structure and envelope are present at the (typically four) apical electrodes. CI Channels 5 to 12 typically work in a constant rate (equal interpulse interval) mode. In contrast to FS4, high definition continuous interleaved sampling (HDCIS), which is another clinically available coding strategy in current MED-EL systems, works in a constant rate (equal interpulse interval) mode on all 12 electrodes.

Using a clinically available stimulation strategy allows for adaptation to new signal cues. Thus with FS4, long-term familiarization periods to fine structure coding are possible. In van Hoesel et al. (2008), comparatively short familiarization periods could be realized using their research processor running a custom-made stimulation strategy. In contrast to FS4, the clinically available HDCIS strategy only processes envelope information. Within this framework, binaural unmasking in BiCI users using multichannel coding strategies with and without fine structure processing was evaluated.

The hypothesis of the present study was that BiCI users would show larger binaural intelligibility level differences (BILD) when bilaterally programmed with FS4 than they would with HDCIS. In this way, the signal processing strategies FS4 and HDCIS can be evaluated against the background of effective interaural cue coding. It should be noted that even though FS4 processes fine structure cues technically, it remains questionable whether these cues are accessible to the

**Table 1.** Demographic Information of the BiCI Users who Participated in the Broadband Experiment.

ID	Age (years)	Etiology	Bilateral CI experience (years)	Age at implantation left/right ear (years)	Implant type (left/right ear)
CI01	57	Progressive	2.2	52/55	SONATA/CONCERTO
CI02	38	Mumps	1.6	37/37	CONCERTO/CONCERTO
CI03 <sup>a</sup>	59	Progressive	2.6	57/55	CONCERTO/CONCERTO
CI04	68	Progressive	3.3	65/63	CONCERTO/CONCERTO
CI05	60	progressive	2.3	58/54	SONATA/CONCERTO
CI06	59	Progressive	5.4	51/54	SONATA/SONATA
CI07 <sup>a</sup>	58	Congenital	1.6	56/57	CONCERTO/CONCERTO
CI08 <sup>a</sup>	54	Unknown	7.3	47/44	SONATA/PULSAR
CI09 <sup>a</sup>	45	Congenital	7.6	38/36	SONATA/PULSAR
CI10 <sup>a</sup>	47	Unknown	3.1	44/38	PULSAR/CONCERTO
CI11 <sup>a</sup>	53	Presumably sudden	2.3	48/51	CONCERTO/CONCERTO
CI12 <sup>a</sup>	46	Presumably measles	0.5	44/46	CONCERTO/SYNCHRONY

Note. BiCI = bilaterally implanted cochlear implants.

<sup>a</sup>Users who also participated in the narrowband experiment.

listeners. With the approach presented in this work, this question can be addressed. This set of experiments is further termed as *broadband experiments*.

In addition, the relationship between IPD discrimination using narrowband signals and BILD with FS4 was determined. This was done to evaluate whether narrowband IPD performance can predict broadband BILD. This set of experiments is further termed as *narrowband experiments*.

Both broadband and narrowband experiments were also conducted with NH listeners as a reference. In the broadband experiment, the stimuli were presented to NH listeners either unprocessed or processed using a vocoder which removed interaural fine structure differences but preserved interaural envelope differences.

## Methods

### Participants

Twelve BiCI users (age  $54 \pm 8$  years [mean  $\pm$  standard deviation]) and 8 NH listeners (age  $44 \pm 8$  years) participated in the broadband experiments. Seven of the 12 BiCI users and 3 of the 8 NH listeners also participated in the narrowband experiments.

The NH listeners had audiometric thresholds of 20 dB HL or less for frequencies from 250 to 8000 Hz measured in octave steps. All BiCI users had bilateral postlingual deafness, were provided bilaterally with MED-CI systems, and had used their respective OPUS 2 (MED-EL) processors for at least 6 months. Additionally, all BiCI users had at least 10 active electrode contacts on each side at their last clinical fitting prior to the start of the study. Finally, all BiCI users achieved SRTs below

0 dB signal-to-noise ratio (SNR) on the German Oldenburg Sentence Test (Oldenburger Satztest [OLSA], see Stimuli section). For details about the BiCI users, see Table 1.

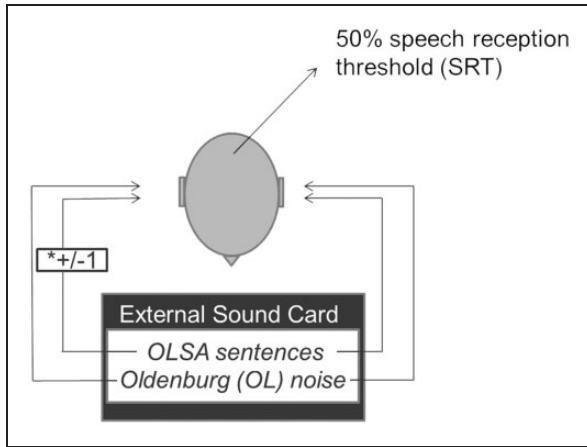
All BiCI users were already familiar with the speech material and procedure of the OLSA, therefore only one list in noise (consisting of 20 sentences) was taken for practising immediately prior to testing. In contrast, five of eight NH listeners were unfamiliar with the OLSA and were therefore trained using two lists (consisting of 20 sentences each) immediately prior to testing.

Study procedures were carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans. Informed consent was obtained from all test subjects.

### Stimuli

**Broadband experiments.** Binaural unmasking was used to assess the influence of IPD in broadband speech signals on speech understanding of BiCI users and NH listeners. Binaural unmasking was determined as according to Licklider 1948): the difference in SRT between two listening conditions: a diotic condition (speech in noise on both ears and no difference between ear signals) and a dichotic condition (speech in noise on both ears and the phase of the speech signal was inverted on one ear). The SRT difference between conditions will be referred to as the BILD.

In NH listeners, BILD was determined using headphone presentation. As in other BILD studies (e.g., Goverts & Houtgast, 2010), a design with an  $N_0S_0$  versus  $N_0S_\pi$  presentation was used: Noise was presented



**Figure 1.** Stimulus presentation in the broadband experiment. In the diotic condition, no difference between the ear signals existed (marked by “+1”). In the dichotic condition, the speech signal was inverted in one ear (marked by “-1”).

homophasically and the target signal was presented either homophasically or antiphasically (see Figure 1). In the antiphasic condition ( $N_0S_\pi$ ), a phase reversal was applied to the speech signal for one ear.

In the dichotic (antiphasic) condition, imposed IPD introduced two interaural cues: interaural envelope differences and interaural fine structure differences.

To compare the BILD with and without fine structure cues in NH listeners, a special condition was implemented. In this condition, the ear signals (speech in noise) were processed using a 12-channel vocoder based on finite impulse response filters implemented in MATLAB to mimic the bandpass filter bank of the MED-EL CI system (details about this type of filter bank can be found in Zirn, Arndt, Aschendorff, & Wesarg, 2015). After filtering, the signal envelope was extracted in every frequency band using the Hilbert transform (Hilbert, 1912). This was done for both ear signals separately. The channel- and side-specific envelopes were then used to modulate the amplitude of 12 narrow bands of noise which were obtained from white noise by applying the same filter bank as described earlier. By taking the same narrow bands of noise for the left and the right ear signals (diotically) imposed with side-specific amplitude fluctuations, interaural fine structure differences were eliminated in the vocoded condition. In contrast, interaural envelope differences were still present in the vocoded stimuli.

NH listeners were tested in two conditions: unprocessed and vocoded. The BiCI users received the unprocessed stimuli through the auxiliary audio input of their processors, both of which were programmed with either FS4 or HDCIS. HDCIS is an envelope-based strategy stimulating at a constant stimulation rate similar

to the original continuous interleaved sampling (CIS) strategy proposed by Wilson et al. (1991). In contrast to the original implementation of CIS, HDCIS uses another type of envelope detection based on the Hilbert transform and a higher stimulation rate (typically 800–1,600 pps, see Muller et al., 2012). As a consequence of the constant stimulation rate, the inter-pulse intervals are constant with HDCIS. In contrast, the FS4 strategy considers the fine structure in the signal in four apical electrodes by stimulating at the zero-crossings of the corresponding bandpass-filtered signals (Riss et al., 2014). This procedure results in nonconstant inter-pulse intervals.

Speech material was taken from the OLSA (Wagener, Brand, & Kollmeier, 1999). The masker was Oldenburg noise, a steady-state noise with a speech-shaped spectrum and was also taken from the OLSA. The masker always started to play 1 s before the onset of the target sentence. The typical duration of the OLSA sentences was 2 s. The duration of the masker was 4 s.

**Narrowband experiments.** The approach of the narrowband experiments was to exclusively examine the effect of interaural fine structure differences. Interaural envelope differences were deliberately minimized. For this purpose, IPD sensitivity was measured using narrowband signals. A 150-Hz pure tone (corresponding to the center frequency of CI Channel 1 of the MED-EL CI system) was used with a total duration of 0.5 s, ramped up and down with the rising and falling slopes (100 ms each) of a Hanning window (see Figure 2). 150 Hz corresponds to the center frequency of CI Channel 1 in the typical FS4 map with all 12 CI channels activated. In the right ear, the pure tone was presented without any phase variations according to Equation 1.

$$y_r(t) = A_c \sin(\omega_c t) \quad (1)$$

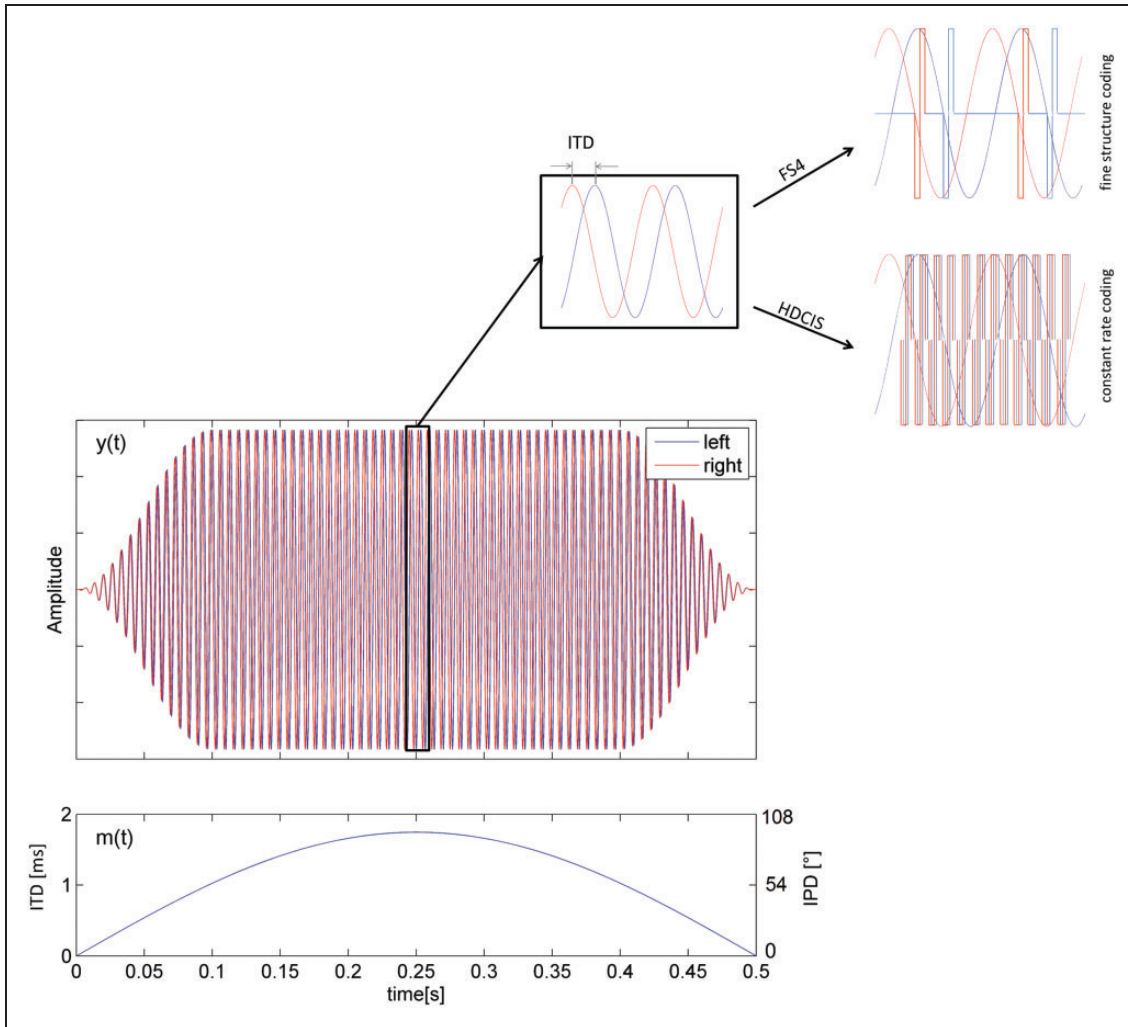
The phase of the left pure tone, however, was modulated according to Equation 2.

$$y_l(t) = A_c \sin(\omega_c t + m(t)) \quad (2)$$

with  $\omega_c = 2\pi f_c$ ,  $f_c$ : carrier frequency 150 Hz, and  $m(t)$ : sinusoidal modulator of the phase of  $y(t)$ .

The frequency of the sinusoidal modulation signal was 1 Hz.

According to Equation 2 and Figure 2, the phase shift between the left and right ear signal was faded smoothly in and out. The IPD was  $0^\circ$  at the onset and offset; at the temporal center of the binaural stimulus, the IPD was largest. The upper IPD limit determined by the test design was  $180^\circ$ . With Equation 3, the ITD in seconds can be determined from the IPD in degrees.



**Figure 2.** Time signals  $y(t)$  of narrowband stimuli with IPD are shown in the center—two 150 Hz sinusoidal tones with an IPD in the temporal center of the stimulus of  $100^\circ$  corresponding to an ITD of 1.9 ms. The left ear signal (blue) was phase modulated and lagged behind the right ear signal (red). The modulator  $m(t)$  is shown at the bottom. On the upper right, the resulting stimulation pattern is shown for HDCIS and FS4 schematically.

Note. IPD = interaural phase difference; ITD = interaural time difference; HDCIS = high definition continuous interleaved sampling; FS4 = Fine structure processing strategy.

The nature of the binaural cue presented by the phase-modulated stimuli involves a dynamic interaural timing cue. The sensation is a dynamically moving sound source on the imaginary line joining the ears; it begins in the middle, moves to the left, and then moves back to the middle.

$$\text{ITD} = \frac{1}{f} \times \frac{\text{IPD}}{360^\circ} \quad (3)$$

The largest IPD used in the test was  $180^\circ$ , which corresponds to an ITD of 3.3 ms in the temporal center of the stimulus occurring at 250 ms (entire stimulus duration was 500 ms). This value approximately corresponds to the IPD obtained by phase inversion in the broadband

experiment in CI Channel 1 at large SNR, described later.

The stimulation patterns at the electrodes were controlled using two OPUS 2 processors, each connected with an implant detector board (a PULSAR implant in a box). The stimulation patterns were visualized via a multichannel oscilloscope which was connected to the apical electrode contacts of the implant detector board. The stimulation patterns gained in this way confirmed the mutual activation of CI Channel 1 predominantly using the stimuli described earlier (150 Hz pure tones). CI Channel 2 showed a similar activation pattern with an attenuation of 10 to 12 dB compared with CI Channel 1. CI Channel 3 was further attenuated compared with CI Channel 2.

We therefore consider the excitation achieved in the narrowband experiment as stimulation on an interaural apical electrode pair. The analysis of stimulation using this method further allowed looking into details of phase coding using FS4. The measurements confirmed that the phase of the 150 Hz pure tones was coded into pulse timing on CI Channel 1. The temporal precision of zero-crossing determination implemented in FS4 is inversely related to the sampling rate in the respective CI channel. On CI Channel 1, the sampling rate is typically situated between 3,000 and 10,000 pps. Only zero-crossings of either a positive or negative flank of the bandpass filter output are considered to lead to a stimulation rate equal (for single pulses) or proportional (for double pulses) to the instantaneous sound frequency. The sampling rate of 3 to 10 kHz leads to a temporal accuracy of zero-crossing determination of 0 to 100  $\mu$ s for 10 kHz and 0 to 333  $\mu$ s for 3 kHz. In BiCI users, this procedure results in an interaural jitter in the order of  $\pm 100$  to  $\pm 333$   $\mu$ s.

Similar measurements with two OPUS 2 processors programmed with HDCIS confirmed constant-rate coding at a fixed rate (typically between 1,200 and 1,600 pps) with temporal envelope fluctuations proportional to the amplitude of the acoustic input.

**Stimulus generation (and presentation).** In both experiments, acoustic stimuli with 44.1 kHz sampling frequency and 16 bit quantization depth were generated on a PC running MATLAB. The applied soundcard type was RME Fireface UC. Stimuli were presented to BiCI users via a stereo audio cable connected to the auxiliary inputs of the OPUS 2 processors, and to NH listeners, via Sennheiser HD 280 Pro headphones.

## Procedure

**Fitting.** Prior to the experiment, all 12 BiCI users were bilaterally fitted with FS4. For the initial test, this fitting was taken over and controlled in terms of interaural loudness balancing. All BiCI users had at least 10, usually 12, active electrode contacts on each side. The four apical electrodes were bilaterally active in all BiCI users. Furthermore, all BiCI users had four fine structure channels on each side. The settings of the automatic gain control (AGC) were the same on both sides. After completion of the first set of (broadband and narrowband) tests, the BiCI users were programmed with HDCIS preserving the channel number and frequency table from FS4. Also, the AGC and the loudness growth function remained unchanged by the new fitting. Only the stimulation levels (threshold and most comfortable levels) on both sides were adjusted to the needs of each subject. This fitting was then maintained for a familiarization period of 6 to 12 weeks. Upon completion of this familiarization period, the broadband and narrowband

tests were repeated with HDCIS bilaterally. After test completion, the subjects could decide which coding strategies they preferred. Eight out of the 12 subjects wanted to use two bilateral program pairs, one with FS4 and other with HDCIS.

After fitting, but prior to all experiments, each subject's most comfortable presentation level for each stimulus was determined before both the narrowband and the broadband experiments. For NH listeners, this was done via the headphones and for the BiCI users via auxiliary inputs of the CI processors.

For the broadband experiment, the individual most comfortable presentation level was determined by means of individual loudness scaling for different presentation levels, resulting in individual loudness growth functions. For this purpose, 2 s segments of OL noise were presented to the subjects in a fixed procedure starting at 55 dB SPL and were then increased in 5 dB increments, one trial per level. While listening to the stimuli, subjects had to indicate the subjective loudness for each level on a scale ranging from "0" (*inaudible*) to "50" (*extremely loud*). If a subject had problems rating a given stimulus, the stimulus with the corresponding level was repeated.

Finally, for the following broadband experiment, the level of OL noise was held constant at the presentation level corresponding to the value of "30" (comfortably loud) on the subjective loudness scale. The level of the speech signal was varied according to the SNR.

The sound pressure level of the fixed starting stimulus of 55 dB SPL was calibrated using an ear simulator type G:R.A.S. IEC 60318-1 and a sound level meter type Norsonic Nor140. The type of headphones was Sennheiser HD 280 Pro. In BiCI users, the same electrical signal corresponding to 55 dB SPL via headphones was presented to the auxiliary inputs of the CI processors as the starting value. The level was then increased in 5 dB increments, similar to the procedure described earlier. Rating was done on the same subjective loudness scale as described earlier.

For the narrowband experiments, a similar loudness scaling procedure was applied: The sound pressure level for acoustic stimulation was increased from a very low level stepwise up to the sensation level "loud." For each stimulus, a rating on the same scale as described earlier was conducted. The presentation level corresponding to "30" was chosen for the following experiment.

**Broadband experiments.** BILD were tested using the OLSA sentence test (HörTech, 2011). SRTs at 50% correct performance in noise were measured using an adaptive procedure (Hagerman & Kinnefors, 1995). The SRTs of both the diotic and dichotic conditions were measured two times. The presentation level of OL noise was fixed, whereas the presentation level of speech was variable

according to the actual SNR. If the subject could correctly identify more than two words of a presented sentence, the SNR was then lowered; if the subject could identify two or fewer words, the SNR was raised. This procedure corresponds to that described in the OLSA documentation (HörTech, 2011). A major difference between the BILD procedures described in the HörTech documentation and those of the present study was that although the document suggests free-field stimulus presentation, in the present study, signals were presented via cable to the auxiliary inputs of the speech processors to BiCI users and via headphones to NH listeners.

BILD was defined similarly by Licklider (1948) and Goverts and Houtgast (2010) as the difference between the speech-reception thresholds in noise in diotic presentation mode ( $SRT N_0S_0$ ) and in dichotic presentation mode with antiphasic speech ( $SRT N_0S_\pi$ ).

**Narrowband experiments.** For determination of IPD thresholds, a three alternative forced choice paradigm was implemented. Three stimuli (two diotic; one dichotic, i.e., containing IPD) were presented to the subjects in each trial. The position of the stimulus with IPD was randomized. Subjects' task was to determine the position of the stimulus with IPD.

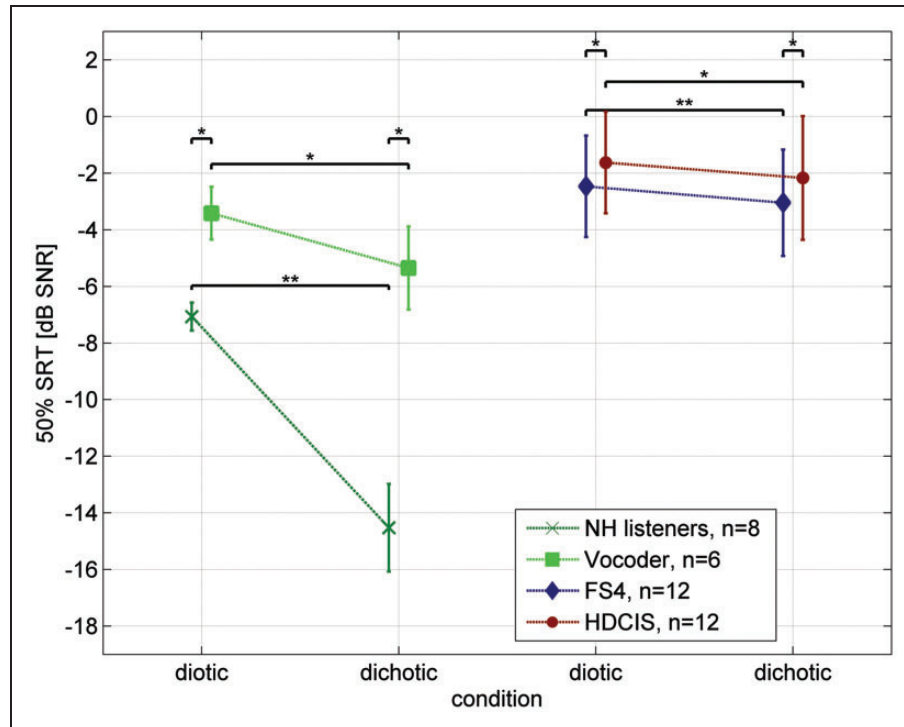
The IPD were adjusted according to a two-down, one-up rule to estimate the 70.7% point of the psychometric function (Levitt, 1971), referred to as IPD thresholds. The initial IPD was  $40^\circ$  corresponding to 0.74 ms (see Equation 3). The IPD was halved after every second reversal until a step size of  $5^\circ$  was reached. After 12 reversals, the run was completed and the mean from the last four reversals was taken as the IPD threshold. The procedure was repeated once (test and retest) and the mean of the thresholds obtained with both tests was taken as the final individual IPD threshold.

**Statistical analysis.** Statistical analysis was done using Wilcoxon signed-rank tests for paired data. For correlation analysis, Pearson product-moment correlation coefficients were derived. The level of significance was defined as  $\alpha = 5\%$ ; significant  $p$ -values ( $<5\%$ ) were marked with \*, and highly significant  $p$ -values ( $<1\%$ ) with \*\*.

## Results

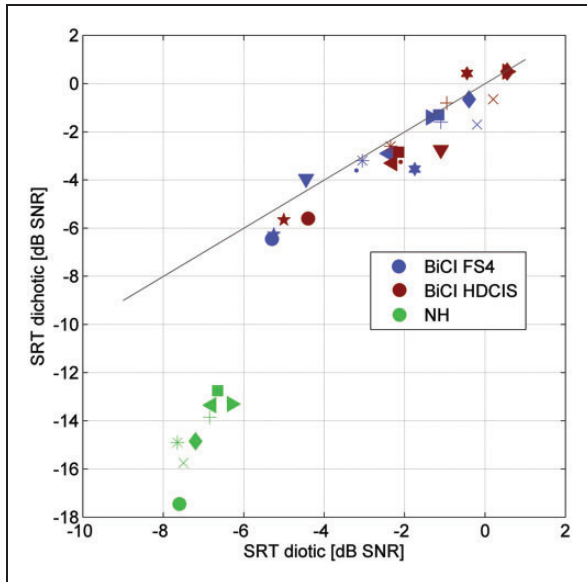
### Broadband Experiments

The BILD results of NH listeners and BiCI users are shown in Figure 3 (averaged data) and Figure 4 (individual results). The SRT of NH listeners was  $-7.1 \pm 0.6$  dB



**Figure 3.** 50% SRTs  $\pm$  standard deviation in the diotic and dichotic condition for NH listeners (left) and BiCI users (right). The SRTs of NH listeners are shown for the unprocessed and vocoder processed stimuli. The SRTs of BiCI users are shown for two coding strategies, HDCIS and FS4.

Note. SRTs = speech reception threshold; HDCIS = high definition continuous interleaved sampling; NH = normal hearing; BiCI = bilaterally implanted cochlear implants; FS4 = Fine structure processing strategy.



**Figure 4.** Scatter plot of diotic versus dichotic SRTs for NH listeners (green), BiCI users programmed with HDCIS (brown) and with FS4 (blue). Every point below the angle bisector represents a positive BILD.

Note. SRTs = speech reception threshold; BiCI = bilaterally implanted cochlear implants; HDCIS = high definition continuous interleaved sampling; BILD = binaural speech intelligibility level difference; FS4 = Fine structure processing strategy.

SNR (mean  $\pm$  standard deviation) in the unprocessed diotic condition and  $-14.6 \pm 1.7$  dB SNR in the unprocessed dichotic condition resulting in a BILD of  $7.5 \pm 1.2$  dB. This BILD was highly significant (two-sided Wilcoxon signed-rank test,  $p = 0.008$ ).

Six of the eight NH listeners participated in an additional experiment applying vocoder-processed stimuli. Compared with unprocessed stimuli, speech reception was impaired using the vocoder in both the diotic and dichotic conditions. Thus, the SRTs in both, the diotic ( $-3.3 \pm 1.0$  dB SNR,  $p = .03$ ) and the dichotic ( $-5.3 \pm 0.25$  dB SNR,  $p = .03$ ) conditions were elevated compared with the SRTs with unprocessed stimuli. Nevertheless, a strongly reduced but still significant BILD of  $2.0 \pm 0.6$  dB SNR ( $p = .03$ ) remained.

In BiCI users programmed with FS4, the mean SRT in the diotic condition was  $-2.5 \pm 1.9$  dB SNR and in the dichotic condition was  $-3.0 \pm 1.9$  dB SNR. Subtraction of the individual SRTs in the dichotic condition from them in the diotic condition and averaging led to the resulting mean BILD of  $0.6 \pm 1.9$  dB (two-sided Wilcoxon signed-rank test,  $p = .01$ ).

The same BiCI users programmed with HDCIS showed a mean SRT of  $-1.6 \pm 1.8$  dB SNR in the diotic condition and  $-2.2 \pm 2.2$  dB SNR in the dichotic condition. The resulting mean BILD was  $0.5 \pm 2.0$  dB and also significant (two-sided Wilcoxon signed-rank test,  $p = .03$ ).

A two-factor, within-subjects ANOVA was calculated based on the SRT data, with factors of diotic/dichotic and FS4/HDCIS. The analysis revealed main effects of diotic/dichotic,  $F(1, 11) = 25.4$ ,  $p = .0004$ , and of FS4/HDCIS,  $F(1, 11) = 7.9$ ,  $p = .017$ , and no interaction between the factors,  $F(1, 11) = 0.012$ ,  $p = .92$ .

Looking at the SRTs in the conditions separately, a significant difference between either HDCIS or FS4 occurred. This was true for both, the diotic (two-sided Wilcoxon signed-rank test,  $p = .01$ ) and the dichotic conditions (two-sided Wilcoxon signed-rank test,  $p = .03$ ).

### Narrowband Experiments

Figure 5 shows IPD thresholds obtained in NH listeners and BiCI users. The mean IPD threshold in the three NH listeners included in this experiment was  $20.8 \pm 10.0^\circ$ , which corresponds to an ITD of  $0.37 \pm 0.19$  ms (see Equation 3).

The results of the seven BiCI users included in this experiment were very heterogeneous and dependent on the coding strategy used. The constant rate coding strategy HDCIS led to the worst IPD threshold of  $176.4^\circ$  in average, which closely corresponds to an ITD in the temporal center of the stimulus of 3.3 ms representing the upper threshold limit determined by test design. With FS4, the same BiCI users had significantly lower IPD thresholds ( $p = .02$ ). The mean IPD threshold with FS4 was  $117.4^\circ$  which corresponds to an ITD in the temporal center of the stimulus of 2.2 ms.

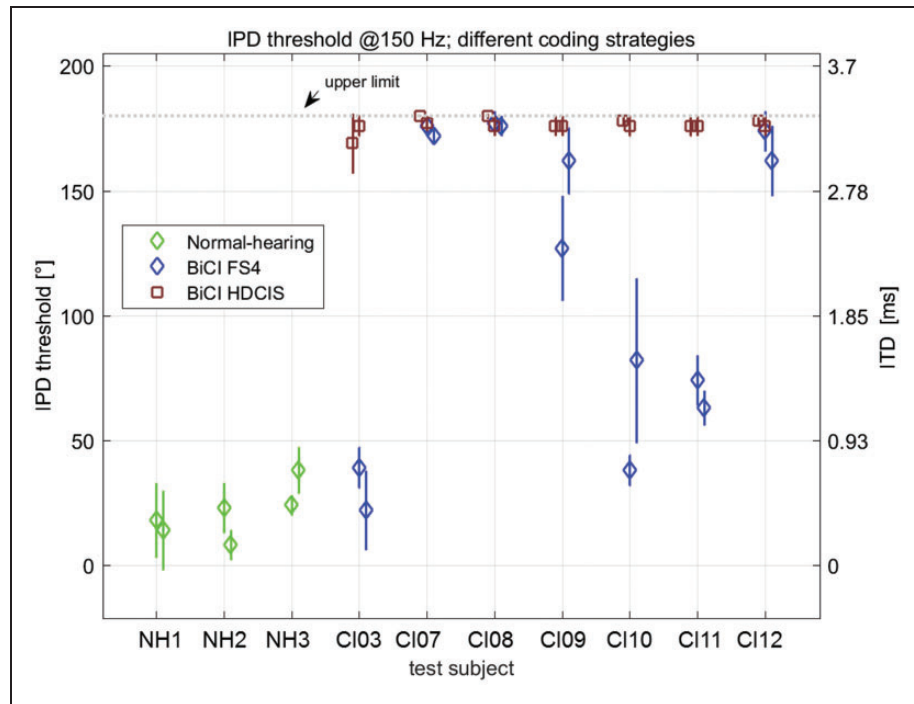
Using FS4, four of the seven BiCI subjects reached IPD thresholds considerably lower than the 3.3 ms upper limit. Two of the seven BiCI users (CI03 and CI10) even reached IPD thresholds close to those of NH listeners.

For the seven BiCI users participating in both experiments, no significant correlation was found (Pearson product-moment correlation coefficient,  $r = -0.42$ ,  $p = .34$ ) between the IPD thresholds in the narrowband experiment and BILD in the broadband experiment both obtained with FS4 (see Figure 6). The correlation was calculated based on seven data pairs (one per subject). It has to be noted that seven subjects is a low sample size with which to do a correlation.

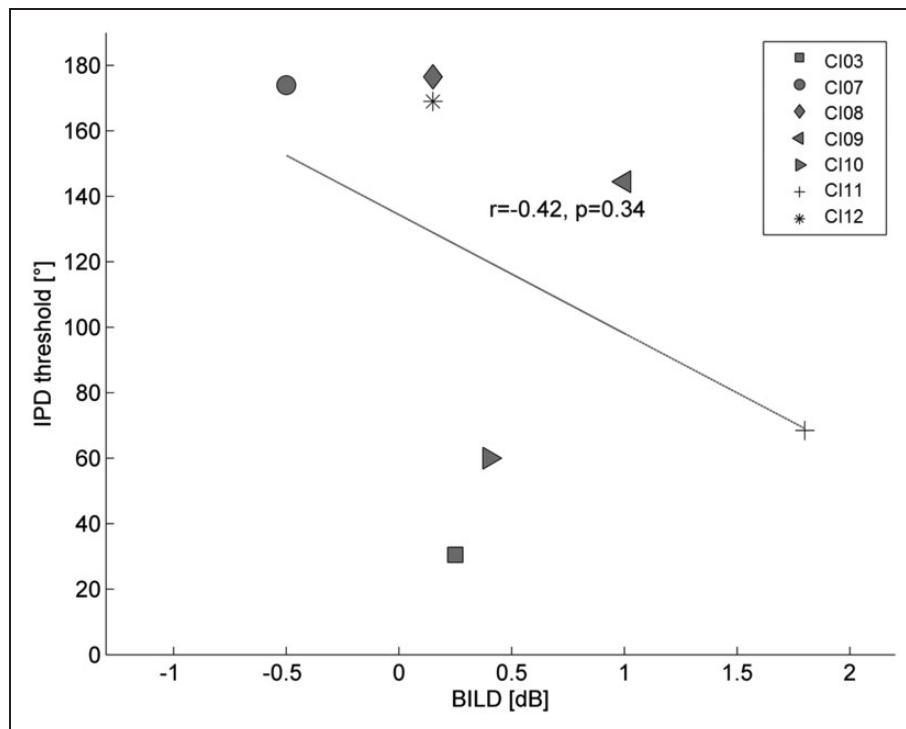
### Discussion

This study investigated the ability of BiCI users who had been sequentially programmed with either FS4 or HDCIS to binaurally unmask speech in the presence of speech-shaped noise. For this purpose, a binaural speech-intelligibility-based test measuring BILDs was conducted. BILD, which was first described by Licklider (1948), is a manifestation of binaural unmasking with highest speech intelligibility with noise plus speech in one ear and noise plus inverted speech in the





**Figure 5.** IPD thresholds  $\pm$  standard deviation obtained in the narrowband experiment. The results of NH listeners are shown on the left (green diamonds), the results of BiCI users programmed either with HDCIS (brown squares) or FS4 (blue diamonds) are shown on the right. Note. IPD = interaural phase difference; NH = normal hearing; BiCI = bilaterally implanted cochlear implants; HDCIS = high definition continuous interleaved sampling; FS4 = Fine structure processing strategy.



**Figure 6.** Correlation analysis of IPD thresholds versus BILD both obtained with FS4 in the narrowband versus broadband experiment in the seven BiCI users who participated in both experiments. Note. IPD = interaural phase difference; BILD = binaural speech intelligibility level difference; BiCI = bilaterally implanted cochlear implants; HDCIS = high definition continuous interleaved sampling; FS4 = Fine structure processing strategy.

other ear. We named that experiment the broadband experiment. Furthermore, the IPD sensitivity was evaluated for narrowband stimuli in what we named the narrowband experiment.

The broadband experiment showed that BILD obtained with both stimulation strategies, HDCIS and FS4, was significant but small compared with NH listeners. A two-factor, within-subjects ANOVA revealed a significant influence of the coding strategy favoring FS4. This outcome stands in contrast to that of van Hoesel et al. (2008), who found no binaural speech unmasking in their four BiCI subjects. There are two potential reasons for the different outcomes. First, the limited number of four BiCI users participating in van Hoesel et al.'s study limits the conclusions that can be drawn. Second, the paradigms differed: van Hoesel et al. applied an ITD of 700  $\mu$ s to speech in diotic noise independent of the SNR, whereas in the present study, phase inversion of the target speech signal in diotic noise was applied. Carhart, Tillman, and Johnson (1967) compared the effect of phase inversion (as done in the present study) and interaural delay (as done in van Hoesel et al.) and found that phase inversion led to an improvement of 4 to 5 dB under continuous noise, whereas a 0.8-ms interaural delay resulted in a 3-dB improvement. Levitt and Rabiner (1967) also compared the effect of phase inversion and interaural delay. They found a 6-dB improvement of speech intelligibility due to phase inversion, but only a 3-dB improvement due to 0.5 to 10 ms of interaural delay. The larger effects obtained with phase inversion on speech intelligibility were the motivation for the present study to use this approach.

BiCI users obtained better absolute SRTs and IPD thresholds when bilaterally fitted with FS4 than with HDCIS. Nevertheless, BILD were similar with FS4 and HDCIS, despite the additional temporal fine-structure information present in the FS4 scheme.

A reduced but still significant BILD which was exclusively based on interaural envelope differences was measurable in our NH subjects in the "processed" (vocoder) condition compared with the unprocessed condition. This indicates that solely interaural envelope differences enable BILD on a reduced level compared with the unprocessed condition.

This outcome is in general agreement with that of van de Par and Kohlrausch (1997) who measured BMLDs at 250 Hz and at 4000 Hz. The stimuli at 4000 Hz were designed to preserve the temporal *fine structure* information available at 250 Hz within the envelope. The stimuli centered at 4000 Hz generally produced smaller BMLDs than did the stimuli at 250 Hz, although the pattern of results as a function of masker bandwidth was the same.

An important question to be addressed in this context was whether the BiCI users with postlingual deafness are generally able to perceive IPD. Laback et al. (2007)

included only four of their eight BiCI users in their study on lateralization discrimination. The four omitted subjects did not fulfill the selection criterion, as defined by the ability to reproducibly perform left/right discrimination on the basis of 600  $\mu$ s interaural delay in a sequence of four pulses at a pulse rate of 100 pps. This fact already indicates that not every postlingually deafened BiCI user is able to perceive IPD in a physiologic range.

The BiCI users in the present study were all postlingually deaf, which reduces the likelihood that they possessed unusually poor IPD sensitivity (Litovsky, Jones, Agrawal, & van Hoesel, 2010). Further, Laback et al. (2015) found no clear relation between onset of profound bilateral hearing loss and sensitivity to ITDs. To quantify the individual sensitivity to IPD for a subset of the BiCI users in the present study, IPD thresholds were determined in 7 of the 12 BiCI users. Large interindividual variability was found. However, at least four of the seven BiCI users reached measurable IPD thresholds lower than the upper limit of the test (180° phase shift in 150 Hz pure tones) with FS4. Two of the seven BiCI users actually reached IPD thresholds close to the IPD thresholds of our NH listeners.

For comparison, van Hoesel (2007) investigated the sensitivity of three BiCI users to time-varying interaural delays using a binaural-beat task in which a pulse train with diotic onsets followed by increasing interaural delays was to be distinguished from one that remained diotic. They used pulse trains with 300 ms duration and three different pulse rates (100, 200, and 300 pps). At 100 pps, the three BiCI users reached detection thresholds between 0.1 and 0.5% change in the stimulation rate corresponding to a maximum interaural delay of 300 to 1500  $\mu$ s. At 200 pps, the detection thresholds increased to 0.8 to 11% corresponding to 2.4 to 33 ms. The IPD threshold found in the present study for 150 Hz was 2.2 ms with FS4, thus in between the outcomes at 100 and 200 pps of the study of van Hoesel.

van Hoesel et al. (2008) concluded, as did Carlyon, Long, and Deeks (2008), that the sensitivity of BiCI users to time-varying interaural delays is inversely related to the stimulation rate with lowest (best) thresholds for lowest pulse rates tested (in this case 100 pps).

To compare the IPD thresholds obtained in NH listeners in the present study with previous findings, publications on binaural beat perception in NH listeners are helpful. In the present study, the IPD generated interaural frequency differences in the rising and falling flank of the modulator in the order of  $\pm 1.5$  Hz for 90° IPD and  $\pm 3$  Hz for 180° IPD.

The frequency difference of 2 Hz of two tones presented each to one ear is critical for binaural beats: Below, the sound appears to move right and left across

the head. For higher frequencies, the perception becomes rougher and appears to fluctuate in loudness (Moore, 2012). Binaural beats are best perceivable in NH listeners for tone frequencies around 400 to 500 Hz with reduced sensitivity for higher and lower frequencies (Licklider, 1950; Perrott & Nelson, 1969).

Interestingly, no significant correlation between IPD thresholds and BILD occurred in the present study. Thus, lower IPD thresholds were not associated with larger BILD with FS4 in these BiCI users. In other words, a reasonably good IPD sensitivity on an apical interaural electrode pair obtained with FS4 was not a sufficient precondition for BILD based on multielectrode stimulation on both sides also obtained with FS4. A reason for this finding might be channel interactions (Nelson, Donaldson, & Kreft, 2008), which can distort interaural timing cues if more than one electrode is stimulated at each ear. Egger, Majdak, and Laback (2016) investigated the sensitivity to interaural delays in BiCI users with stimuli which were presented to either one or two electrode pairs. They found that presenting consistent interaural delays at two electrode pairs resulted in improved sensitivity only if the tonotopic separation was large and if they were stimulated with the same levels as the corresponding single pairs. As a result, the double pair was louder than the respective single pairs. The authors concluded that channel interaction is critical to the process of combining interaural timing information across electrodes.

Further, Laback et al. (2015) mentioned two other electric/physiologic factors which potentially affect BiCI users' ITD sensitivity. First, the auditory nerve responses to electric stimulation at low stimulation rates show stronger phase locking than they do to acoustic stimulation. Second, in electric hearing, even with the apical-most CI electrodes, the low-frequency pathway is often not adequately stimulated. Thus, precise ITD processing of the medial superior olive may not be fully utilized. These factors, separately or in conjunction with the channel interactions described earlier, may account for the fact that, in the present study, BILD was in a similar range with FS4 and HDCIS.

### Acknowledgments

We want to thank Dr. Andrew Oxenham, two anonymous reviewers, Dr. Peter Nopp, and Dirk Meister for their extensive and helpful comments. Further, we want to thank Michael Todd for proofreading a version of the manuscript.

### Declaration of Conflicting Interests

The authors declared 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 work was supported by MED-EL Elektromedizinische Geraete Gesellschaft m. b. H., Innsbruck, Austria and MED-EL Deutschland GmbH.

### References

- Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization* (revised ed.). Cambridge, MA: MIT.
- Bronkhorst, A. W., & Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *The Journal of the Acoustical Society of America*, *83*, 1508–1516.
- Carhart, R., Tillman, T. W., & Johnson, K. R. (1967). Release of masking for speech through interaural time delay. *The Journal of the Acoustical Society of America*, *42*, 124–138.
- Carlyon, R. P., Long, C. J., & Deeks, J. M. (2008). Pulse-rate discrimination by cochlear-implant and normal-hearing listeners with and without binaural cues. *The Journal of the Acoustical Society of America*, *123*, 2276–2286.
- Colburn, S. H., Shinn-Cunningham, B., Kidd, G. Jr., & Durlach, N. (2006). The perceptual consequences of binaural hearing. *International Journal of Audiology*, *45*(Suppl 1): S34–S44.
- 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 for Research in Otolaryngology: JARO*, *17*, 55–67.
- Goupell, M. J., & Litovsky, R. Y. (2015). Sensitivity to interaural envelope correlation changes in bilateral cochlear-implant users. *The Journal of the Acoustical Society of America*, *137*, 335–349.
- Goverts, S. T., & Houtgast, T. (2010). The binaural intelligibility level difference in hearing-impaired listeners: The role of supra-threshold deficits. *The Journal of the Acoustical Society of America*, *127*, 3073–3084.
- Hagerman, B., & Kinnefors, C. (1995). Efficient adaptive methods for measuring speech reception threshold in quiet and in noise. *Scandinavian Audiology*, *24*, 71–77.
- Hilbert, D. (1912). *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*. Leipzig, Germany: Teubner.
- HörTech. (2011). *OLSA handbuch*. Retrieved from [http://www.hoertech.de/web/dateien/HT.OLSA\\_Handbuch\\_Rev01.0\\_mitUmschlag.pdf](http://www.hoertech.de/web/dateien/HT.OLSA_Handbuch_Rev01.0_mitUmschlag.pdf).
- Kerber, S., & Seeber, B. U. (2012). Sound localization in noise by normal-hearing listeners and cochlear implant users. *Ear and Hearing*, *33*, 445–457.
- Laback, B., Egger, K., & Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hearing Research*, *322*, 138–150.
- Laback, B., Majdak, P., & Baumgartner, W. D. (2007). Lateralization discrimination of interaural time delays in four-pulse sequences in electric and acoustic hearing. *The Journal of the Acoustical Society of America*, *121*, 2182–2191.

- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(Suppl 2): 467+.
- Levitt, H., & Rabiner, L. R. (1967). Binaural release from masking for speech and gain in intelligibility. *The Journal of the Acoustical Society of America*, 42, 601–608.
- Licklider, J. (1948). The influence of interaural phase relations upon the masking speech by white noise. *The Journal of the Acoustical Society of America*, 20, 150–159.
- Licklider, J. C. R. (1950). On the frequency limits of binaural beats. *The Journal of the Acoustical Society of America*, 22, 468.
- Litovsky, R. Y., Jones, G. L., Agrawal, S., & van Hoesel, R. (2010). Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans. *The Journal of the Acoustical Society of America*, 127, 400–414.
- Long, C. J., Carlyon, R. P., Litovsky, R. Y., & Downs, D. H. (2006). Binaural unmasking with bilateral cochlear implants. *Journal of the Association for Research in Otolaryngology: JARO*, 7, 352–360.
- Moore, B. C. J. (2012). *An introduction to the psychology of hearing* (6th ed.). Leiden, Netherlands: Brill.
- Muller, J., Brill, S., Hagen, R., Moeltner, A., Brockmeier, S. J., Stark, T., . . . Anderson, I. (2012). Clinical trial results with the MED-EL fine structure processing coding strategy in experienced cochlear implant users. *ORL; Journal for Otorhino-Laryngology and Its Related Specialties*, 74, 185–198.
- Nelson, D. A., Donaldson, G. S., & Kreft, H. (2008). Forward-masked spatial tuning curves in cochlear implant users. *The Journal of the Acoustical Society of America*, 123, 1522–1543.
- Perrott, D. R., & Nelson, M. A. (1969). Limits for the detection of binaural beats. *The Journal of the Acoustical Society of America*, 46, 1477–1481.
- Riss, D., Hamzavi, J. S., Blineder, M., Honeder, C., Ehrenreich, I., Kaider, A., . . . Arnoldner, C. (2014). FS4, FS4-p, and FSP: A 4-month crossover study of 3 fine structure sound-coding strategies. *Ear and Hearing*, 35, e272–e281.
- Seeber, B. U., & Fastl, H. (2008). Localization cues with bilateral cochlear implants. *The Journal of the Acoustical Society of America*, 123, 1030–1042.
- van de Par, S., & Kohlrausch, A. (1997). A new approach to comparing binaural masking level differences at low and high frequencies. *The Journal of the Acoustical Society of America*, 101, 1671–1680.
- Van Deun, L., van Wieringen, A., Francart, T., Buchner, A., Lenarz, T., & Wouters, J. (2011). Binaural unmasking of multi-channel stimuli in bilateral cochlear implant users. *Journal of the Association for Research in Otolaryngology: JARO*, 12, 659–670.
- van Hoesel, R. J. (2004). Exploring the benefits of bilateral cochlear implants. *Audiology & Neuro-otology*, 9, 234–246.
- van Hoesel, R. J. (2007). Sensitivity to binaural timing in bilateral cochlear implant users. *The Journal of the Acoustical Society of America*, 121, 2192–2206.
- van Hoesel, R., Bohm, M., Pesch, J., Vandali, A., Battmer, R. D., & Lenarz, T. (2008). Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies. *The Journal of the Acoustical Society of America*, 123, 2249–2263.
- Wagener, K., Brand, T., & Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache Teil 11: Optimierung des Oldenburger Satztests. *Z Audiol.*, 38.
- Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., & Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature*, 352, 236–238.
- Zirn, S., Arndt, S., Aschendorff, A., & Wesarg, T. (2015). Interaural stimulation timing in single sided deaf cochlear implant users. *Hearing Research*, 328, 148–156.