Spatial Release From Masking in Simulated Cochlear Implant Users With and Without Access to Low-Frequency Acoustic Hearing

Ben Williges[1], Mathias Dietz[1], Volker Hohmann[1],
and Tim Jürgens[1]

Abstract

For normal-hearing listeners, speech intelligibility improves if speech and noise are spatially separated. While this spatial release from masking has already been quantified in normal-hearing listeners in many studies, it is less clear how spatial release from masking changes in cochlear implant listeners with and without access to low-frequency acoustic hearing. Spatial release from masking depends on differences in access to speech cues due to hearing status and hearing device. To investigate the influence of these factors on speech intelligibility, the present study measured speech reception thresholds in spatially separated speech and noise for 10 different listener types. A vocoder was used to simulate cochlear implant processing and low-frequency filtering was used to simulate residual low-frequency hearing. These forms of processing were combined to simulate cochlear implant listening, listening based on low-frequency residual hearing, and combinations thereof. Simulated cochlear implant users with additional low-frequency acoustic hearing showed better speech intelligibility in noise than simulated cochlear implant users without acoustic hearing and had access to more spatial speech cues (e.g., higher binaural squelch). Cochlear implant listener types showed higher spatial release from masking with bilateral access to low-frequency acoustic hearing than without. A binaural speech intelligibility model with normal binaural processing showed overall good agreement with measured speech reception thresholds, spatial release from masking, and spatial speech cues. This indicates that differences in speech cues available to listener types are sufficient to explain the changes of spatial release from masking across these simulated listener types.

Keywords

cochlear implant, speech perception, auditory model, electroacoustic stimulation, bimodal

Introduction

Compared with normal-hearing (NH) listeners, users of cochlear implants (CIs) often have greater difficulties understanding speech in noisy environments (e.g., Haumann, Lenarz, & Büchner, 2010). This difficulty can be quantified using a speech-in-noise test, for example, by the difference in speech reception thresholds (SRTs, i.e. the signal-to-noise ratio (SNR) at 50% recognition score) with respect to a NH listener. Depending on the types of noise and speech material used, differences in SRTs can be considerable, with NH listeners typically showing SRTs at negative SNRs and CI users at positive SNRs (e.g., Cullington & Zeng, 2008). However, differences in SRTs can also depend on the specific hearing device a patient uses, or is able to use. For instance, improved performance in speech intelligibility tasks is well documented for CI users with preserved low-frequency acoustic hearing (LF) in the implanted ear, when they use a hearing aid (HA) in addition to their CI (Büchner et al., 2009; Ching, van Wanrooy, & Dillon, 2007; Gantz & Turner, 2003; James et al., 2006). This electroacoustic benefit (EA-benefit) can be expressed in terms of the percentage gain in quiet or in noise at a fixed SNR, with respect to performance using the CI alone. The EA-benefit is reported to be in the range of 5% to 30% in quiet (Gantz & Turner, 2003) and 26% on average in noise (Büchner et al., 2009).

[1]Medizinische Physik, Cluster of Excellence "Hearing4all," and Forschungszentrum Neurosensorik, Carl-von-Ossietzky Universität Oldenburg, Germany

Corresponding author:
Ben Williges, Carl-von-Ossietzky Universität Oldenburg, Carl-von-Ossietzky Straße 9-11, Oldenburg 26129, Germany.
Email: ben.williges@uni-oldenburg.de

The EA-benefit can also be expressed in terms of an SRT improvement (measured in dB) and is reported to be 7 dB on average in a German sentence test (Büchner et al., 2009). Bimodal listeners, i.e., CI users who use a HA in the contralateral ear in addition to their CI, also show an EA-benefit, both with (Gifford, Dorman, Sheffield, Teece, & Olund, 2014) and without preserved acoustic hearing in the implanted ear (for an overview, see Schafer, Amlani, Seibold, & Shattuck, 2007).

One factor influencing the speech intelligibility of any listener in everyday life is spatial release from masking (SRM). SRM is defined as the improvement in speech intelligibility (scores or SRT) when speech and noise are spatially separated, compared with colocated speech and noise (for an overview, see Litovsky, 2012). SRM can be attributed to monaural effects, such as listening with the better ear (a consequence of the head-shadow (HS), or binaural effects that manifest, for example, in binaural squelch (BSq) or binaural summation (BS). SRM in NH listeners depends on the room acoustics and can be as high as 10 to 12 dB (Beutelmann & Brand, 2006) under anechoic conditions. SRM is often substantially lower for listeners with symmetric hearing losses (Beutelmann & Brand, 2006) and for bilateral CI users (Gifford et al., 2014) compared with NH listeners. Listeners with asymmetric hearing loss, i.e., those who use acoustic, electric stimulation, or a combination thereof in either ear (e.g., unilateral electroacoustically stimulated EAS users or bimodal CI users) also show reduced SRM compared with NH listeners, although performance is highly variable across subjects (Gifford et al., 2014).

Significant intersubject variability in speech intelligibility performance, EA-benefit, and SRM is prevalent across different listener types (e.g., when comparing unilaterally implanted listeners with preserved ipsilateral vs. contralateral acoustic hearing), but also within types of listeners with the same hearing/device configuration. For instance, some unilateral EAS users show a large EA-benefit, and others very little or none (James et al., 2006; Lenarz et al., 2013). Possible sources of variability within a listener type may be of peripheral origin (e.g., different CI signal processing strategies, different spread of the electrical field generated by the implant, different preservation of spiral ganglion cells, or different frequency ranges of residual acoustic hearing), or originating in differences in more central processing (e.g., to combine speech information from electric and acoustic stimulation). This high inter-individual variability impedes the investigation of possible interactions between the EA-benefit and other factors that influence speech intelligibility (such as SRM) and makes it difficult to determine mechanisms or cues employed by different listener types to improve their speech-in-noise performance. The current study therefore aims to minimize interindividual variability (within a listener type) by systematically simulating these

different listener types. Like Schoof, Green, Faulkner, and Rosen (2013) or Jones, Kan, and Litovsky (2014), we assume that NH listeners are a more homogenous group than hearing-impaired or CI listeners. Thus, data obtained from simulated listeners should show less inter-individual variability within a simulated listener type. Unilateral and bilateral CI, uni- and bilateral electroacoustic, and bimodal listener types are simulated using NH listeners presented with different combinations of vocoded and low-frequency narrow-band speech. The vocoder (Bräcker, Hohmann, Kollmeier, & Schulte, 2009) realistically mimics the most important signal processing principles also available in CIs and, additionally, takes into account physiological details such as spatial spread of the electric field in the cochlea of CI users. Low-frequency narrow-band speech is used to simulate ipsi- or contralateral residual acoustic hearing. Differences in SRTs across listener types should thus be free from inter-individual factors specific to the hearing loss that may be present in clinical studies. However, differences in SRTs between simulated listener types and clinical studies (and also differences in SRM) may still be attributable to either the different signal processing used to simulate listener types or to general inabilities of the binaural auditory system to extract the present spatial speech cues.

Speech intelligibility models can be used to investigate if the results obtained in measurements can be fully attributed to the peripheral processing (due to different availability of speech cues for the different listener types) or if further central processing deficits need to be assumed to predict the data. The binaural speech intelligibility model (BSIM; Beutelmann & Brand, 2006; Beutelmann, Brand, & Kollmeier, 2010), for instance, shows high correlations with measured data from NH and hearing-impaired listeners in spatial speech-in-noise situations (overall correlation coefficient .95). In BSIM, hearing impairment is modeled as changes due to the periphery (i.e., only taking the audibility of the listeners into account). The back-end (i.e., binaural processing of speech and noise in combination with the speech intelligibility index (SII); ANSI, 1997) is left unchanged. Therefore, differences between measured and modeled SRTs can then be attributed to central processing deficits, especially if the model performs better than the listeners. This approach of only changing the peripheral processing and not changing the back-end will also be used in the current study for comparisons between SRTs of simulated listener types and the model.

The first goal of the current study is to systematically investigate how SRTs, SRM, and spatial cues of speech perception change across different simulated listener types. Of special interest is the interaction between EA-benefit and SRM. The second goal of the current study is to investigate whether the measured SRTs and especially the changes in SRM across simulated listener types can be predicted using a binaural speech

intelligibility model without changes in the binaural stage of the model.

Methods

Participants

Ten NH listeners with a median age of 27.5 years (min: 23 years, max: 34 years, 5 women, 5 men) participated in the study. Listeners' absolute threshold did not exceed 20 dB HL for frequencies between 125 Hz and 8 kHz, as measured using standard pure-tone audiometry. All listeners were paid for their participation and had some listening experience with the speech test used in the current study. The listeners signed informed consent prior to their participation. Ethical consent was granted by the Ethics Committee of the University of Oldenburg.

Apparatus and Calibration

The listeners were seated in a sound-attenuating booth, listening to stimuli via Sennheiser HDA 200 circumaural headphones, which were D/A converted by a RME Fireface UC soundcard. The stimuli were generated on a standard PC with MATLAB using customized scripts to simulate different listener types. Free-field equalization was carried out (according to DIN EN ISO 389–8, 2004) using a finite impulse response-filter with 801 coefficients. Stimulus levels were calibrated to dB SPL using a Bruel & Kjaer (B&K) 4153 artificial ear with a B&K ½ in. microphone type 4134, which was attached to a B&K 2610 measurement amplifier.

Simulation of Different Listener Types

Electric hearing. The current study used a vocoder (Bräcker et al., 2009) to simulate electric hearing. This vocoder uses a sequence of processing steps that mimic realistically the technical and physiological steps of signal processing in CI listeners (Figure 1). A continuous interleaved sampling (CIS)-like strategy (Wilson et al., 1991) is used to simulate the processing from acoustic signal to electric signal. The audio signal (at a sampling frequency of 55.1 kHz) is decomposed into 12 frequency channels (corresponding to the number of electrodes in Med-EL CI devices) using a third-order gammatone filterbank (Hohmann, 2002, here termed CI-simulation filterbank) with center frequencies between 120 and 7410 Hz (Table 1). The bandwidth of each channel was set to one equivalent rectangular bandwidth (ERB). The Hilbert envelope of the output of each channel is then sampled at a frequency of 4583 Hz (i.e., 1/12 of the audio signal's sampling frequency, which is close to the maximum total stimulation rate available in Med-EL CI devices, as also used in Bräcker et al., 2009), producing
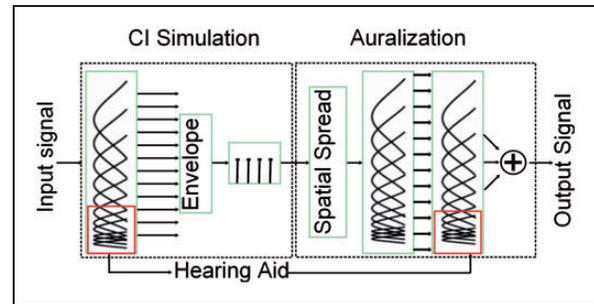
Figure 1. Processing scheme of the vocoder that was used to simulate cochlear implanted and electroacoustic listeners. The red area shows the processing path of the LF condition. CI = cochlear implant; LF = low-frequency acoustic hearing.

Table 1. Parameter of the Vocoder Used for the Three Simulated Listening Types: Cochlea Implant, Low-Frequency Acoustic, and Electroacoustic. All Other Simulated Listening Types Are Combinations of These Three.

Parameter	Simulation	CI Auralisation	LF Auralisation	EAS Auralisation
Center frequency (Hz)	120	390	120	120
	235	550	235	235
	384	759	384	384
	579	1031		1031
	836	1384		1384
	1175	1843		1843
	1624	2440		2440
	2222	3216		3216
	3019	4225		4225
	4084	5537		5537
	5507	7243		7243
	7410	9460		9460

Note. CI = cochlear implant; LF = low-frequency acoustic hearing; EAS = electroacoustic stimulation.

an envelope-weighted pulse train-like signal. The timing of the envelope sampling across channels can either be performed sequentially using a fixed order across channels in a round-robin manner, or by randomizing the sequence of pulses across channels within a time frame. The current study uses the randomized sequence because sequential stimulation results in a strong pitch percept (corresponding to the pulse frequency), which is unpleasant to listen to. Such a high-frequency rate pitch percept is unlikely to be present in actual CI listeners (cf. Zeng, 2002). Thus, the randomized pulse order is also more realistic from a perceptual point of view. After this pulsatile sampling, spatial spread of the electric field in the perilymph is simulated by multiplying each pulse at a given time with a two-sided exponentially decaying

Table 2. Ten Simulated Listening Types With Abbreviation and Description.

Listener type number	Condition name	Description
1	mon NH	Monaural normal hearing
2	bin NH	Binaural normal hearing
3	uni CI	Unilateral CI (using the vocoder)
4	bil CI	Bilateral CI (using the vocoder with same settings for each ear, whereas the pulsatile sampling was done independently in each ear)
5	bim CI	Bimodal CI (vocoder on one ear and low-frequency acoustic hearing on the other ear)
6	uni LF	Unilateral LF (using low-frequency acoustic hearing on one ear)
7	bil LF	Bilateral LF (using low-frequency acoustic hearing on both ears)
8	uni EAS	Unilateral EAS (replacing the three channels with lowest center frequencies by the processing done for unilateral LF)
9	bim EAS	Bimodal electroacoustic (as unilateral EAS, but with additional low-frequency acoustic hearing on the opposite ear)
10	bil EAS	Bilateral electroacoustic (as unilateral EAS, but stimulating both ears, whereas the pulsatile sampling in the vocoder parts was done in each ear independently)

Note. NH = normal hearing; CI = cochlear implant; LF = low-frequency acoustic hearing; EAS = electroacoustic stimulation.

function a across channels across channels (*Equation 1*, thus in the frequency domain). This step transfers stimulation to adjacent channels with |d| being the absolute distance in mm from the stimulating electrode with the decaying constant of 3.6 mm. For computational reasons, d was limited to an interval of ±14 mm distance from the stimulating electrode. For larger distances, no spatial spread was assumed.

$$a = e^{\left(-\frac{|d|}{3.6\,mm}\right)} \tag{1}$$

In the following auralization step, the signal of each channel is filtered using two third-order one-ERB-wide gammatone filterbanks (Hohmann, 2002) with slightly different center frequencies than the CI-simulation filterbank (here: 390 to 9460 Hz, see Table 2). The center frequencies chosen here correspond to the place-frequencies of the 12 equidistant positions of a Med-EL electrode array (insertion depth of 24 mm) within a human cochlea of 32 mm length according to Greenwood's (1990) frequency-to-place map for humans. Thus, the first auralization filterbank simulates the transfer of the signal in each channel to the respective position along the cochlear partition. Hence, this step provides the carrier to the sampled envelopes, which has the same narrow-band characteristic as the respective gammatone channel. The second auralization filterbank (same center frequencies as the first auralization filterbank) compensates for filter group delays (Hohmann, 2002), reduces side-band modulations, and sums the frequency channels to create a mono audio signal. RMS-scaling of the vocoder's (broadband) output to the same RMS as the input signal is done to compensate for level changes that are due to pulsatile sampling or band pass filtering (described later).

At the end of the vocoder processing, the mono signal consists of narrow-bandpass filtered carriers (with a stochastic, noise-like fine structure), with envelopes corresponding to the sampled envelopes that are spectrally smeared and with carriers that are slightly shifted in frequency according to the electrode-to-best-frequency mapping.

Acoustic hearing. Normal hearing is simulated by using the (input) audio signal without any processing. LF (simulating, e.g., residual acoustic hearing after atraumatic CI surgery) is simulated by using the three channels with lowest center frequencies of the vocoder's CI-simulation filterbank. Each of these three filter outputs is provided directly to the synthesis filterbank without additional processing. Thus, the acoustic signal is band limited between 120 Hz (simulating the low-frequency cutoff of HA tube/receiver arrangement) and 390 Hz (simulating a profound high-frequency hearing loss). No amplification or compression was performed in this step, and no supra-threshold deficits of hearing impairment (such as temporal or spectral processing impairments, e.g., as discussed in Léger, Moore, & Lorenzi, 2012) were simulated.

Different Listener Types

Ten different listener types were simulated, consisting of different combinations of electric and acoustic hearing (see Figure 2 and Table 2). These included monaural and binaural normal hearing, CI, LF, and EAS listener types. In addition, bimodal EAS (LF in one ear and EAS in the
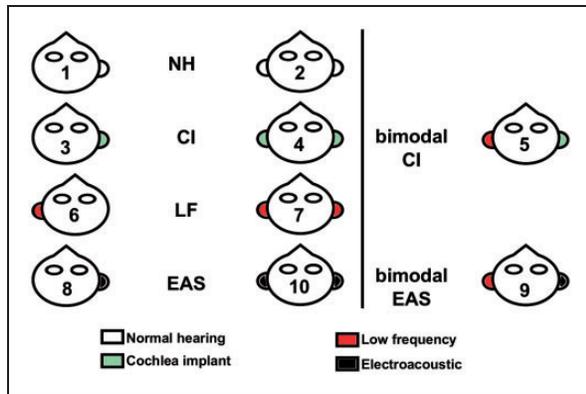
Figure 2. Overview of the 10 simulated listener types, namely normal hearing (NH) monaural, NH binaural, unilateral CI, bilateral CI, unilateral LF, bilateral LF, unilateral EAS, bilateral EAS, bimodal EAS, and bimodal CI. Numbers inside the heads correspond to Table 2.

CI = cochlear implant; LF = low-frequency acoustic hearing; EAS = electroacoustic stimulation.

other) and bimodal CI (LF in one ear and CI-vocoder in the other) listener types were simulated. All monaural or unilateral conditions were measured with the right ear, except for monaural LF, which was simulated using the left ear. This was done to be able to measure speech intelligibility for acoustic hearing in the same ear (with the same head-related impulse response (HRIR) and low-pass filtering) as used for bimodal listener types.

Speech Intelligibility Measurements

A German sentence test (Oldenburger Satztest; Wagener, Kühnel, & Kollmeier, 1999) was used to measure SRTs in the presence of stationary noise, i.e., the SNR corresponding to 50% speech intelligibility score. This speech test consists of sentences with five words each in the fixed order: subject, verb, numeral, adjective, object (e.g., "Peter kauft drei nasse Sessel", "Peter buys three wet chairs"). The spectrum of the noise equals the long-term-averaged spectrum of all sentences of the sentence test. During the course of the measurement, the noise level was held constant at 65 dB SPL and the speech level – and thus the SNR – was varied to measure the 50% SRT using an adaptive procedure (Brand & Kollmeier, 2002). Lists of 20 sentences were used, and 0 dB SNR was chosen as a starting value in all conditions. If a measurement was going to exceed 25 dB SNR (corresponding to 90 dB SPL speech level) three times, this measurement was aborted. Listeners' responses were obtained using a touch screen that showed all possible response alternatives of each word of the sentence test (closed set response, cf. Warzybok, Rennies, Brand, Doclo, & Kollmeier, 2013).

Virtual Acoustics

Spatial rendering of noise and speech was accomplished by convolving the stimuli with Kemar-HRIRs recorded in an anechoic room with the sound source placed 1 m from the microphone. These Kemar-HRIRs were taken from a publicly available database (Algazi, Duda, Thompson, & Avendano, 2001). Speech was presented from the front (i.e., with 0° incident angle) and noise from either −90° (facing the left ear), 0°, or 90° (facing the right ear), generating three conditions: S_0N_{-90}, S_0N_0, and S_0N_{90}. Note that the HRIR at −90° for one ear is not identical to the HRIR at 90° for the contralateral ear and vice versa. Speech and noise signal were convolved with the desired HRIR, scaled to the desired SNR, added together, and then processed to simulate the listener types.

Familiarization and Measuring Process

Prior to the measurements, each participant completed four lists of 20 sentences each for familiarization with the speech material and with the signal processing used in the current study. Listening conditions (different simulated listener types) used during familiarization were binaural NH, bilateral CI, bilateral EAS, and bilateral LF. After familiarization, listening conditions were pseudorandomized across subjects. For each listening condition, the noise-incident angle was also pseudorandomized. Measurements were conducted during two sessions of maximum 2 hr duration. Listeners were offered regular breaks, and were given the opportunity for additional breaks at any point in the experiment.

Calculation of Spatial Cues for Speech Perception

Based on the measured SRTs the *SRM*, head-shadow[1] (HS_{mon}, HS_{bin}), binaural squelch (*BSq*), and binaural summation (*BS*) were calculated per subject for each listening condition using the following formulae to investigate spatial cues of speech intelligibility:

$$SRM = SRT(S_0N_0) - SRT(S_0N_{-90}). \quad [2] \tag{2}$$

$$HS_{mon} = \min(SRT(Mon., S_0N_{contra}) \\ - SRT(Mon., S_0N_{ipsi})). \tag{3}$$

$$HS_{bin} = \min(SRT(Bin., S_0N_{contra}) \\ - SRT(Mon., S_0N_{ipsi})). \tag{4}$$

$$BSq = \min(SRT(Bin., S_0N_{ipsi}) \\ - SRT(Mon., S_0N_{contra})). \tag{5}$$

$$BS = \min(SRT(Bin., S_0N_0) \\ - SRT(Mon., S_0N_0)). \tag{6}$$

In these equations, "*Mon.*" denotes the monaural (unilateral) and "*Bin.*" the binaural (bilateral) listening condition; "N_{contra}" denotes a condition with noise contralateral to the listening ear (in monaural listening condition) and "N_{ipsi}" ipsilateral.

Model

Binaural speech intelligibility model. The BSIM (Beutelmann & Brand, 2006; Beutelmann et al., 2010) was used to predict SRTs for different simulated listener types and noise-incident angles. BSIM can accurately predict SRTs for a wide range of listeners with different audiometric thresholds and for different rooms (Beutelmann & Brand, 2006). BSIM uses speech (s_R and s_L) and noise (n_R and n_L, where R and L denotes right and left, see Figure 3) as separate input signals and filters them into 30 channels equally spaced on an ERB-scale with center frequencies ranging from 146 to 8346 Hz. In each band, an equalization-cancellation (EC) mechanism (Durlach, 1963) is used to find (a) the optimal time delay and (b) the optimal gain adjustment between right and left signals to minimize the noise and maximize the speech signal, thus improving the SNR in each channel. Of the two monaural SNRs (mon_R and mon_L), and the binaurally enhanced SNR, the highest SNR is chosen as input to the SII. Thus, both effects—listening with the better ear and binaural processing—are taken into account. The absolute hearing threshold is also considered by employing a hearing threshold noise (Beutelmann & Brand, 2006), which is spectrally shaped to match the audiometric thresholds. Hearing threshold noise is
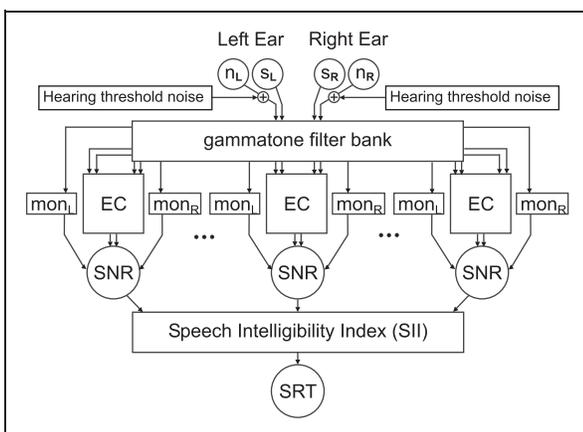
added to the left and right noise signal such that it contributes to the band-wise monaural and binaural SNR calculation. Note, however, that the EC-mechanism calculates the optimal time delay and gain estimation without this hearing threshold noise. Binaural processing errors are included in BSIM as Gaussian distributed errors of time delay and gain in the EC-mechanism. In the current study, the processing errors are those of NH listeners, as described in Beutelmann et al. (2010). The SII is calculated using the highest SNRs (out of binaural and two monaural SNRs) in each frequency channel according to ANSI S3.5–1997 with one third octave band procedure and speech-in-noise band-importance function. The input SNR that results in a specific SII-reference value is then taken as the predicted SRT. This transformation from SII to SRT is described in more detail below.

Deviations from the standard BSIM in the current study. The current study involves speech intelligibility of both simulated electric (i.e., vocoded) and acoustic hearing. Because vocoded speech is less intelligible than nonvocoded speech, two different SII-reference values (CI_{SII} and NH_{SII}) were chosen for the transformation of SII to SRT:

- CI_{SII} was found as that SII-value matching the SRT in the binaural CI S_0N_0 listening condition.
- NH_{SII} was found as that SII-value matching the SRT in the binaural NH S_0N_0 listening condition.
- For the (mixed) EA listening conditions, the SII-channels containing the acoustic signals (SII-channels 1 to 6, 146 to 414 Hz) were assumed to contribute to the SII-reference value EAS_{SII} with weight NH_{SII}, whereas the SII-channels containing the simulated electric signals (7 to 30, 487 to 8346 Hz) were assumed to contribute with CI_{SII}. Thus, the EAS-SII value is:

$$EAS_{SII} = \frac{6}{30}NH_{SII} + \frac{24}{30}CI_{SII} \qquad (7)$$

This weighted sum represents a linear combination of NH and CI speech intelligibility, but the effect on speech intelligibility in EA listener types is nonlinear because the SII-to-SRT-transfer is a nonlinear process.

One further adjustment was made for the LF listening condition: The acoustic signal was scaled such that only the low-frequency part (SII-channels 1 to 6) was used by the EC-mechanism, and the high-frequency parts were masked by the hearing threshold noise. In all other listening conditions, the hearing threshold noise was set to 0 dB HL.

Figure 3. Processing steps in the binaural speech intelligibility model (BSIM). Adapted with permission from J. Acoust. Soc. Am. 127, 2479 (2010). Copyright 2010 by Acoustical Society of America.
EC = equalization-cancellation; SNR = signal-to-noise ratio; SII = speech intelligibility index; SRT = speech reception threshold.

Results

This section presents speech intelligibility measurements in noise assessed in 10 different simulated listener types to systematically investigate changes in SRTs, SRM, and spatial cues of speech perception across these listener types. Subsequently, these measurements are compared with predictions of a binaural speech intelligibility model (BSIM).

General Trends in the Measurements

Figure 4 shows measured SRTs for 10 different simulated listener types (as indicated by the 10 head symbols and numbers) as box-and-whisker plots. Boxes denote 25th to 75th percentiles and whiskers ±1.5 interquartile range between 25th to 75th percentiles, medians are given as horizontal lines within the boxes. Each listening condition is shown with three different noise-incident angles (−90°, 0°, 90°). Generally, the NH-listener types show lowest SRTs (−20 to −7.6 dB SNR), followed by the mixed EA listener types (−12 to 0 dB SNR), simulated CI (−10 to 2.9 dB SNR), and LF (11.7 to 18 dB SNR) in ascending order. Spatially separating speech and noise (i.e., noise-incident angles −90° and 90°) lowered the SRT in all symmetric binaural conditions (i.e., with the same processing at the left and right ear) with respect to spatially colocating speech and noise (noise-incident

angle 0°). SRTs at −90° are 2 to 3 dB lower than SRTs at 90° because the HRIR of the right ear for −90° is not identical to the HRIR of the left ear for 90° and vice versa. In unilateral conditions, spatially separating speech and noise resulted in lower SRTs when the noise was contralateral to the (listening) ear and in higher SRTs when the noise was ipsilateral to the (listening) ear. A two-factor repeated-measures analysis of variance (ANOVA) with within-subject factors "simulated listener group" and "noise-incident angle" was significant for each factor (noise-incident angle: degrees of freedom, $df = 2$, $F = 384.2$, $p < .001$; simulated listener group: $df = 1.4$ (Greenhouse-Geisser correction), $F = 83.4$, $p < .001$) and revealed significant interaction between both factors ($df = 18$, $F = 39.9$, $p < .001$). This statistics and all following were conducted using IBM SPSS 23.

Interaction Between EA-Benefit and SRM

In order to investigate a possible interaction between SRM and EA-benefit, an ANOVA and pairwise comparisons between SRTs were performed on the measured data. A two-factor repeated-measures ANOVA on the SRTs for the unilateral EAS and unilateral CI listener type and noise-incident angle as within-subject-factors showed a significant effect of noise-incident angle ($df = 2$,
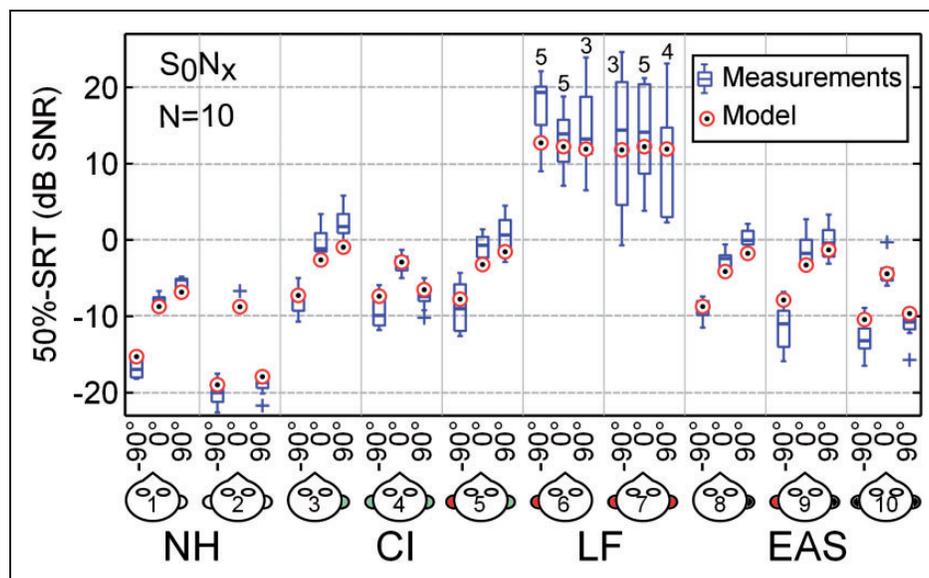

Figure 4. Speech reception thresholds (box-and-whisker plots of 10 participants) and model predictions (red circles) for different simulated listener types. The different simulated listener types on the abscissa are from left to right: normal hearing monaural, normal hearing binaural, unilateral CI, bilateral CI, bimodal, unilateral LF, bilateral LF, unilateral EAS, bimodal EAS, and bilateral EAS. For each listener group, there are three different noise-incident angles (−90°, 0°, 90°). The numbers above the LF boxplots denote number of participants showing SRTs above the maximum allowed 25 dB SNR.
SNR = signal-to-noise ratio; SRT = speech reception threshold; NH = normal hearing; CI = cochlear implant; LF = low-frequency acoustic hearing; EAS = electroacoustically stimulated.

$F = 561.9$, $p < .001$), a significant effect of listener type ($df = 1$, $F = 23.8$, $p = .001$), and a significant interaction between the two factors ($df = 2$, $F = 4.6$, $p = .026$). This indicates that EA-benefit (defined as the difference in SRTs between CI and EAS listener types for each noise-incident angle) and SRM interact in this unilateral condition. For the binaural EAS and binaural CI listener types, a two-way ANOVA showed a significant effect of noise-incident angle ($df = 2$, $F = 191.8$, $p < .001$), a significant effect of listener type ($df = 1$, $F = 66.6$, $p < .001$), and a significant interaction between these two factors ($df = 2, F = 22.9$, $p < .001$). This result indicates that EA-benefit and SRM also interact in these binaural conditions.

Pairwise t-test comparisons with Bonferroni correction ($n = 12$, $p_{corr} = 0.05/12 = .0042$) were carried out, and a one-sided t-test showed a 1.3 dB EA-benefit ($t(9) = 4.93$, $p = .001$), for frontal noise (S_0N_0) and a 1.8 dB EA-benefit for noise azimuth $90°$ (S_0N_{90}, $t(9) = 3.67$, $p = .0025$) for unilateral CI compared with unilateral EAS. A 1.8 dB EA-benefit was also found for binaural CI compared with binaural EAS for S_0N_0 ($t(9) = 2.36$, $p = .0215$), significant without Bonferroni correction, extending to 3.4 dB for spatially separated speech and noise in the conditions S_0N_{90}, ($t(9) = 9.08$, $p < .001$) and S_0N_{-90} ($t(9) = 10.19$, $p < .001$). This confirms that bilateral access to LF acoustic hearing improves SRTs in spatially separated speech and noise.

The SRTs for unilateral EAS were significantly lower relative to bimodal CI in the S_0N_0 condition ($t(9) = 3.90$, $p < .0018$), but not significantly lower for the other noise directions (S_0N_{-90}, $t(9) = 0.08$, $p = .469$, S_0N_{90}, $t(9) = 1.4$, $p = .0977$). SRTs for bimodal CI was not significantly lower than for unilateral CI in any noise direction (S_0N_{-90}, $t(9) = 0.97$, $p = .178$, S_0N_0, $t(9) = 1.09$, $p = .151$, S_0N_{90}, $t(9) = 1.52$, $p = .0814$). SRTs for bimodal EAS were significantly lower relative to unilateral EAS in the S_0N_0 ($t(9) = 3.44$, $p = .0037$), and the S_0N_{-90} ($t(9) = 3.73$, $p = .0024$) condition, but not significantly lower in the S_0N_{90} condition ($t(9) = 0.33$, $p = .3731$). Thus, replacing contralateral with ipsilateral LF acoustic hearing improved speech intelligibility for frontal noise and adding an additional LF acoustic hearing contralaterally to EAS improved speech intelligibility for frontal noise even further.

Box-and-whiskers plots in Figure 5 show individual SRM extracted from the SRT data in Figure 4, according to *Eq. 2*, for all 10 simulated listener types. With respect to the binaural conditions, SRM is highest for the NH types, followed by EAS, CI, and LF[3]. Two-sided t-tests showed that SRM for unilateral CI is not significantly different from bilateral CI ($t(9) = 2.48$, $p = .0351$). Also SRM for unilateral CI is not significantly different from bimodal CI ($t(9) = 0.23$, $p = .8256$), but significantly different from unilateral EAS ($t(9) = 4.12$, $p = .0026$). In
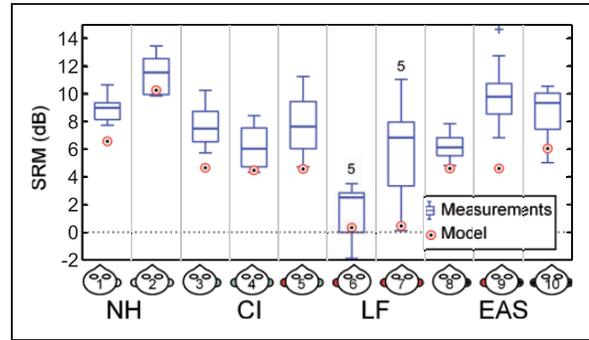

Figure 5. Boxplots of individual SRM (S_0N_{-90} compared with S_0N_0) for each simulated listener type. Red circles show model predictions. The dotted line shows the zero-dB reference. Numbers on top of boxplots denote numbers of SRT-values above 25-dB SNR, which prevented calculation of individual cues. The conditions are the same as those shown in Figure 4.
SRM = spatial release from masking; SNR = signal-to-noise ratio; SRT = speech reception threshold; NH = normal hearing; CI = cochlear implant; LF = low-frequency acoustic hearing; EAS = electroacoustically stimulated.

other words, SRM changes when adding acoustic hearing to one-sided electric hearing ipsilaterally but not contralaterally. SRM is significantly different for bilateral EAS compared with bilateral CI ($t(9) = 5.12$, $p < .001$), and for bimodal EAS compared with bilateral CI ($t(9) = 5.63$, $p < .001$). This confirms that SRM and EA-benefit interact in these conditions, i.e., that the EA-benefit depends on the direction of the noise relative to the speech.

Spatial Speech Cues

Monaural and binaural HS, BSq, and BS were extracted from the SRT data of each participant of the study according to *Eq. 3* to *6* to investigate their contributions to SRM. The results are displayed in Figure 6. Positive values indicate an improvement in (i.e., a lower) SRT in the binaural listening condition compared with the unilateral listening condition. HS was highest (11.4 dB monaural, 13.4 dB binaural) for NH listener types, followed by EAS, CI, bimodal CI, bimodal EAS, and LF in descending order. BSq is generally smaller than HS in absolute dB values. Highest BSq is found in EAS (4.1 dB), followed by NH (3.2 dB) and bimodal EAS (2.6 dB). The listener types CI, bimodal CI, and LF showed BSq close to 0 dB. Two-sided t-tests showed that the BSq for CI is significantly different from the BSq for EAS, ($t(9) = 8.45$, $p < .001$), but not significantly different from bimodal EAS ($t(9) = 1.76$, $p = .112$). The NH BSq is not significantly different from EAS ($t(9) = 1.11$, $p = .2957$), or from bimodal EAS ($t(9) = 0.92$, $p = .3812$). BS effects are even smaller than BSq effects in absolute terms. Here, the highest values were found in simulated CI
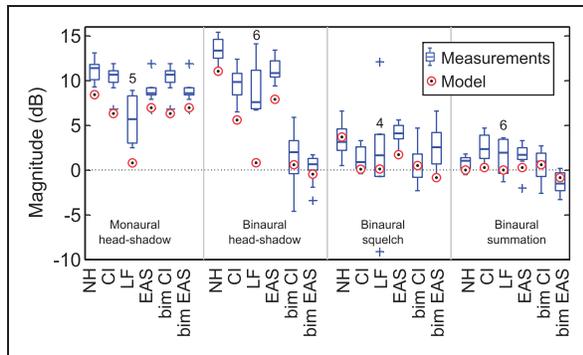
Figure 6. Magnitude in dB of monaural and binaural head-shadow, binaural squelch, and binaural summation for six different simulated listening types: normal hearing, CI, low-frequency acoustic hearing (LF), electroacoustic (EAS), bimodal CI (bim CI), and bimodal EAS (bim EAS). Red circles show model predictions; the dotted line is the zero-dB baseline. All cues are calculated for each subject ($N = 10$) individually. Numbers on top of boxplots denote numbers of SRT values above 25 dB SNR, which prevented calculation of individual cues for some subjects.
CI = cochlear implant; SNR = signal-to-noise ratio; SRT = speech reception threshold.

(up to 2.4 dB), smaller in NH (1.1 dB), and even negative (i.e., the bilateral condition showed poorer SRTs than the unilateral condition) for bimodal EAS. No significant differences were found, when comparing BS for NH with CI ($t(9) = 3.35$, $p = .0086$), or with EAS ($t(9) = 1.37$, $p = .2044$).

Modeling Results

Measured SRTs, SRM, and spatial cues are compared with predictions by the speech intelligibility model BSIM. Red circles in Figure 4 show predicted SRTs using BSIM for all simulated listener types and noise-incident angles. By definition, the two fitted conditions, NH binaural (S_0N_0) and CI binaural (S_0N_0), exactly reproduce the median measured SRTs. The model predictions for all other simulated listener types follow the pattern of the measured results qualitatively very well and also match quantitatively the measured results in many cases. An EA-benefit is predicted by the model in unilateral conditions (compare unilateral CI and unilateral EAS) and binaural symmetric conditions (compare bilateral CI and bilateral EAS) at all noise azimuths. The bimodal EA-benefit is predicted to be 0.8 dB (in comparison with the unilateral EA-benefit, which is 1.5 dB). Note that a lower SRT is predicted for all bilateral symmetric listener types in spatially separated speech and noise (S_0N_{90} or S_0N_{-90}) compared with colocated speech and noise (S_0N_0), confirming that the model is able to predict the typical pattern of SRM.

Reproducing the trends in the measurements, the SRM (see red circles in Figure 5) is predicted to be largest in NH listeners (10.2 dB), second largest in bilateral EAS (6.0 dB), followed by unilateral CI, bilateral CI, bimodal CI, and bimodal EAS (all around 4.5 dB), and almost nonexistent in bilateral LF (0.3 dB). This is in agreement with the measurements obtained, which showed no significant difference between unilateral CI, bimodal CI, and bilateral CI. However, the model underestimates the SRM in bimodal EAS listening condition by 4.8 dB.

Figure 6 plots model predictions (red circles) for HS, BSq, and BS. The model reproduces the trend for monaural and binaural HS effects, but underestimates the effect size by 3 to 5 dB in all listening conditions. Note that the model also underestimates the monaural HS in NH condition by 4 dB. This means that this underestimation is a general bias of BSIM when regarding unilateral situations, because all calculations of HS involve the SRTs of at least one unilateral condition. The model predictions for BSq agree with the measured data, except for bimodal EAS, where the model underestimates the BSq. The model does not show a BS effect, i.e., the BS magnitude is equal to 0 dB, in contrast to the measurement data in NH, CI, and EAS. However, the model predicts BS for bimodal CI in agreement with the measurements, and even negative BS for bimodal EAS. Bimodal CI as an asymmetric listening condition provides the model with nonredundant data, that is, different SNRs across frequency bands and across the two ears, which the model can use to produce slightly better SRTs.

Discussion

This study systematically investigated SRTs, SRM, and spatial speech cues in ten different simulated listener types. The listener types included unilateral and bilateral NH listeners, CI users, listeners having access to low-frequency acoustic information only, and listeners using a combination of electric and acoustic hearing. SRM and EA-benefit showed interaction in EAS users with unilateral and bilateral access to LF acoustic hearing. SRM was found to be larger in simulated bilateral EAS users than in simulated bilateral CI users.

Mechanisms of the Interaction Between SRM and EA-Benefit

The vocoder used in the current study contains important elements to simulate both the signal processing in actual CIs (e.g., number of channels and pulse rate) and physiological features of electrical hearing (e.g., the assignment of electrodes to cochlear locations corresponding to best frequencies, and spatial spread of the electrode signals along the tonotopical arrangement on the auditory nerve). As in actual bilateral CI listeners (cf.

Laback, Egger, & Majdak, 2015), a simulated bilateral CI user has access to interaural level difference (ILD) cues, as well as limited access to interaural time difference (ITD) cues conveyed in the modulated envelopes of the signal (ITD_{env}), cues that may contribute to SRM (Ihlefeld & Litovsky, 2012). ITD cues present in the temporal fine structure (ITD_{fine}) are not available for CI users. Bilateral EAS listeners, in contrast, have ITD_{fine} cues available by virtue of acoustic hearing at low frequencies. These ITD_{fine} cues are beneficial for speech understanding when speech and noise are spatially separated, and thus improve SRTs in the S_0N_{90} and S_0N_{-90} conditions. This indicates that binaural processing (in addition to better ear listening) helps to segregate speech and noise in listeners with access to bilateral acoustic hearing at low frequencies, providing larger SRM. Thus, SRM and EA-benefit do interact in bilateral listening conditions. Of course, unilateral listeners (uni CI, uni EAS) cannot make use of binaural cues but can still benefit from better SNR at low frequencies, which might explain why SRM and EA-benefit do interact in the S_0N_0 and S_0N_{90} listening conditions. In the S_0N_{-90} conditions (i.e., noise contralateral to the listening ear), speech intelligibility is dominated by high frequencies due to the HS effect, and overall SNRs are low. Due to these low SNRs (around $-10\,dB$), low-frequency acoustic information does not provide any EA-benefit.

Modeling the Interaction Between SRM and EA-Benefit

The general pattern of SRTs obtained from the measurements, and particularly the interaction of SRM and EA-benefit in binaural listening conditions, is in line with model predictions using the BSIM (Beutelmann & Brand, 2006). BSIM employs an EC mechanism in each frequency band and chooses the best SNR out of (a) the two monaural SNRs and (b) the binaurally processed SNR. BSIM automatically exploits ILD and ITD_{env} cues, if these are the only cues available (as in simulated bilateral CI listeners). However, if ITD_{fine} cues are also available (as in simulated bilateral EA listeners) BSIM also uses these cues. The agreement between measured and modeled SRTs indicates that the NH binaural system can make optimal use of all available speech cues for SRM. It is not necessary to assume central processing errors in the model in addition to the degraded peripheral processing, to obtain reduced SRM in simulated bilateral CI users (with respect to binaural NH listeners) and enlarged SRM in simulated bilateral EAS users (with respect to bilateral CI users). In fact, the SRM of the model is in agreement with the smallest SRM across the individual participant results within the respective simulated listener types (see Figure 5). Introducing

additional central processing deficits in any of the listener types would have reduced the predicted SRM even further (cf. Beutelmann & Brand, 2006), making the model less predictive of the data.

Modeling the EA-Benefit

A 2 to 3 dB SNR benefit from EA hearing was found in simulated bilateral EA users compared with bilateral CI users for the S_0N_0 condition. The model's ability to correctly predict this EA-benefit is due to the assignment of two different SII-reference values, one for undistorted (acoustic) speech and one for vocoded speech, matching the binaural NH listeners' and simulated bilateral CI listeners' SRTs, respectively. This resulted in a larger SII-reference value for electric signals (due to their higher SRTs) than for acoustic signals. Replacing larger SII-reference values (electric, CI_{SII}) by smaller SII-reference values (normal hearing, NH_{SII}) in low-frequency channels, for example, for bilateral EA listeners, resulted in the predicted EA-benefit. Although this way of modeling EA-benefit matches well the measured SRT data, the possible mechanism underlying EA-benefit is not modeled. Instead, previously suggested mechanisms, such as availability of F0-information (Brown & Bacon, 2009) or low-frequency glimpsing (Li & Loizou, 2008), are implicitly attributed to the acoustic frequency channels having richer speech information and thus higher intelligibility.

Differences Between Measurements and Model Predictions

There are some differences between measured and predicted SRM, HS, and BS results. As stated in the introduction, we assume that differences between model predictions and measurements involve different central processing or binaural processing.

With regard to SRM (in binaural conditions) and BS, BSIM can make optimal use of all available binaural information in the signals to improve the SNR via an EC-process. However, BSIM underestimates SRM for all listener types. Introducing smaller EC-processing errors in BSIM would increase SRM only in binaural conditions (binaural NH, bilateral CI, bilateral EAS and to a small extent bimodal CI and bimodal EAS), but not in monaural conditions like monaural NH or unilateral CI. There are also no obvious reasons for introducing smaller EC-processing errors than those values used by Beutelmann & Brand (2006), who inferred these values from psychoacoustic experiments with NH listeners. Because the SRM is also underestimated in monaural conditions (no EC processing active), this SRM discrepancy points more to effects, that the backend of BSIM, the SII, is not able to model the full

contribution of different speech and noise locations on speech intelligibility. This, unfortunately, leads subsequently to underestimation also of the HS. Large parts of the monaural SRM certainly arise from the frequency-dependent shadowing of sound due to the head. However, also other mechanisms may contribute to the actual size of the monaural SRM, which are not incorporated in BSIM, such as uncommonness of the human listeners to this monaural situation, especially in the difficult condition if the noise faces the listening ear.

One deficit of BSIM is its inability to model BS when the left and right signals are very similar (i.e., when left and right signal contain the same SNR in each frequency band and the same speech cues), such as in binaural NH listeners or simulated bilateral CI users. In contrast to human listeners, BSIM cannot make use of two very similar signals (left and right) to segregate speech and noise better than in monaural listening conditions. Thus, a general mismatch in BS between measurements and model predictions in symmetrical conditions (such as bilateral CI and bilateral EAS) – as found in the current study – does not require the involvement of any central deficit in neural coding of speech. In asymmetric conditions, such as bimodal CI and bimodal EAS, simulated and predicted BS are well matched, because BSIM can utilize the complementary frequency-dependent SNRs at the right and left ears.

Because BSIM uses the SII as a back-end, BSIM by definition can only regard spectral masking effects, that is, energetic masking aspects from the noise on the speech. The EC-process in BSIM has proven to be an effective model in stationary noise (Beutelmann & Brand, 2006) but can, in principle, also be applied to fluctuating noises. Stationary portions in the noise allow the EC process to use ITDs and ILDs to extract a higher SNR for spatially separated signals and noise as compared with conditions when the spatial separation does not exist or is small. However, in its current form, BSIM should not be used with maskers that are assumed to additionally consist of informational masking, such as a concurrent talker. BSIM cannot be used for sound source localization and separation of speech and masker, which are mechanisms alleviating speech intelligibility in such maskers. However, BSIM also cannot get distracted from concurring speech information in the background, as would be the case with a concurrent talker as masker. Thus, the model predictions reported here cannot be transferred to background noises that consist of additional informational masking.

Comparing Measured Data With Data From Actual Patients

A comparison with clinical data obtained from actual patients of different listener types is difficult, at least in absolute terms (i.e., comparing SRT values) because (a) the different speech and noise material and different room acoustics used in other studies will result in different SRTs even for NH listeners (cf. Kollmeier et al., 2015) and (b) there is a large variability across actual listeners of a specific listener type (e.g., Gifford et al., 2014; James et al., 2006). Instead of absolute SRT comparisons, relative comparisons, in terms of spatial speech cues such as SRM, HS, BSq, BS, can be made.

Gifford et al. (2014) investigated spatial speech-in-noise performance in bilateral CI users and bimodal listeners with and without ipsilaterally preserved acoustic hearing. For bilateral CI users, they found 3 to 8 dB SRM, 5 to 10 dB (monaural) HS, −2 to 2 dB BSq, and 0 to 2 dB BS. This is generally in line with the results of simulated bilateral CI users in the current study (averages of 6 dB SRM, 10 dB HS, 1 dB BSq, and 2 dB BS). Further confirmation comes from Schoof et al. (2013), who simulated bilateral CI users using a noise vocoder and reported a 7.5 dB HS effect and 3 dB BS. However, the SRM they reported was only 1.6 dB and thus smaller than both the clinical data and our vocoder-simulation data. This might be explained by differences in the signal processing used in the different vocoders. As Schoof et al. (2013) discussed, their noise-excited vocoder obscured ITD_{env} cues, possibly due to random fluctuations inherent in the noise carrier. The vocoder used in the present study (Bräcker et al., 2009) samples the Hilbert-envelope at a relatively high rate and can thus preserve ITD_{env} cues. This may be more realistic with respect to the signal processing in real bilateral CI users (cf. Laback et al., 2015).

Spatial cues for speech perception extracted from measured data for simulated bimodal listeners with preserved hearing in the implanted ear differed from those found in the study of Gifford et al. (2014). In their study, SRM was 3 to 6 dB, HS 3 to 7 dB, BSq 2 to 12 dB, and BS 1 to 3 dB (compared with median values of 8 dB, 8 dB, 3 dB, and −1 dB, respectively, in the current study).

Limitations With Respect to Actual Patients

A possible reason for differences in SRM and spatial speech cues to actual patients is that some factors evident in measurements with patients are not simulated in the current study. These include, for example, different degrees of residual low-frequency hearing, the usage of directional microphones within HAs or CI speech processors (cf. Gifford et al., 2014), no consideration of compression in either the HA or the CI, no timing differences between the transmitted signals in the ears, and no head movements, which actual patients are likely to utilize in their everyday environments.

Compression in either the HA or CI will reduce ILD cues and leave ITD_{fine} and ITD_{env} cues essentially intact (Wiggins & Seeber, 2011). The reduction in ILD range decreases localization performance (e.g., Kelvasa & Dietz, 2015). This likely reduces SRM especially in those listener groups largely relying on ILD cues, such as bilateral CI or bilateral EAS. If the compression is not too strong, the SNR at one ear would not change much, leaving SRM for the unilateral listener types largely as that in the simulated listeners. Disregarding timing differences between HA and CI in this study is not likely to cause much difference in SRM in monolateral EAS listeners: According to Geißler, Büchner, Chalupper, and Battmer (2015), the timing between acoustic and electric stimulation in the same ear using timing ranges typical for commercial devices (up to 6 ms) has no influence on speech perception. The same holds for bilateral EAS, which is assumed here to consist of the same two HAs and CIs on each ear. The effect of device timing for bimodal and bimodal EAS listeners on speech intelligibility is less clear. Regarding head movements, Grange & Culling (2013) showed that making use of even small head movements can improve SRTs for CI users in spatially separated conditions. However, Grange & Culling (2013) also showed that not all CI patients utilize head movement strategies to improve SRT. Thus, with respect to compression, timing differences between devices, and head movements, the processing used to simulate listeners in this study can be viewed as an "optimal" setting for comparisons with actual listener types.

Another difference to actual patients is the ability of CI patients to make use of ITD_{env} and ILD cues for the improvement of their speech intelligibility. Especially ITD cue utilization is susceptible if deafness or profound hearing loss was present prelingually (Kan & Litovsky, 2015; Laback et al., 2015). In contrast, the NH listeners used here to simulate different listener types can access all these cues. Therefore, the results of this study are not transferable to those obtained from prelingually deaf CI subjects.

Missing Bimodal Benefit

The missing EA-benefit in simulated bimodal CI listeners in the current study (which was present across a range of clinical studies reviewed by Ching et al., 2007) may arise from the very limited overlap between the evaluated frequency regions in simulated CI and contralateral LF acoustic hearing. Such an arrangement might only apply for those listeners in clinical studies with poor audiometric thresholds contralateral to their implanted ear. In contrast, for actual bimodal listeners, the EA-benefit was found not to correlate strongly with audiometric thresholds averaged over 250 and 500 Hz

(Gifford, Dorman, McKarns, & Spahr, 2007). Instead, a large interindividual variability in the EA-benefit (George, Devocht, Chalupper, & Stokroos, 2015) may indicate that some listeners experience problems "fusing" electric and acoustic signals across the two ears. Further studies could usefully quantify spatial cues for speech perception systematically both in actual bimodal patients and in simulated bimodal listeners with different amounts of contralateral acoustic hearing, to determine whether binaural processing is utilized by these listeners to segregate speech and noise. If this is indeed the case, it raises the possibility that binaural processing can be optimally supported using binaural signal processing in CI and HA to alleviate speech perception for these listeners in everyday life.

Conclusions

This study quantified SRTs in different simulated listener types using spatial arrangements of speech and noise. SRM and spatial cues of speech perception were extracted with the aim of systematically investigating their sizes in these listener types. Furthermore, the measured SRTs were compared with SRTs predicted by a binaural speech intelligibility model. The following conclusions can be drawn:

1. SRTs were lower for CI listener types with than without access to LF acoustic hearing, indicating an EA-benefit. SRM was found to be largest for NH listener and was considerably smaller for listener types with simulated electric hearing. SRM was higher for simulated CI listener types with than without access to LF acoustic hearing. This indicates that the EA-benefit and SRM interact.
2. The SRTs predicted by the BSIM with normal binaural processing match the measured SRTs of simulated listener types very well. The model also predicts the trends in SRM and spatial cues for speech perception across listener types. Because the predicted SRM is in agreement with the smallest SRM values extracted across the subjects, a degraded central processing in the model would have introduced a mismatch with measured data and thus does not need to be assumed for simulated listener types. A logical next step would be to investigate whether degraded central processing needs to be assumed in actual CI users to explain SRM differences.
3. Comparison of our results with SRM and spatial cues for speech perception in real patients shows good agreement for the bilateral CI user group (cf. Gifford et al., 2014). However, larger differences between simulated listeners and actual patients are observed for those users having access to LF, possibly due to the large variability in acoustic frequency

ranges accessible in actual listeners. Thus, the vocoder simulations used in the current study enable a systematic investigation of spatial speech intelligibility, which is less influenced by interindividual differences within a listener group than measurements with actual patients.

Notes

1. HS is mostly defined in the literature as a purely monaural effect (cf. Gifford et al., 2014). In the current study both the monaural as well as a binaural definition (cf. Litovsky, 2012; Schafer et al., 2007) are used to be able to disentangle monaural and binaural effects. Note that binaural HS is the opposite of BSq, that is, binaural HS indicates how an additional ear affects speech perception if the noise is contralateral to the added ear (thus, the additional ear consists of a more favorable SNR). BSq, in contrast, indicates speech perception changes if the noise is ipsilateral to the added ear (thus, the additional ear consists of a poorer SNR).
2. For the monaural LF case, $SRT(S_0N_{90})$ was used.
3. Interestingly, no SRM was found in either the unilateral LF or bilateral LF condition when using the median SRTs as basis for calculation (rather than individual data across the ten participants).

References

Algazi, V. R., Duda, R. O., Thompson, D. M., & Avendano, C. (2001). *The CIPIC HRTF database. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics* (pp. 99–102). New Paltz, NY: Mohonk Mountain House.

ANSI. (1997). *Methods for the calculation of the speech intelligibility index. American National Standard S3.5–1997.* Melville, NY: Standards Secretariat, Acoustical Society of America.

Beutelmann, R., & Brand, T. (2006). Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *120*, 331–342.

Beutelmann, R., Brand, T., & Kollmeier, B. (2010). Revision, extension, and evaluation of a binaural speech intelligibility model. *The Journal of the Acoustical Society of America*, *127*, 2479–2497.

Bräcker, T., Hohmann, V., Kollmeier, B., & Schulte, M. (2009). Simulation und Vergleich von Sprachkodierungsstrategien in Cochlea-Implantaten [Simulation and comparison of sound coding strategies in cochlear implants]. *Zeitschrift für Audiologie*, *48*, 158–169.

Brand, T., & Kollmeier, B. (2002). Efficient adaptive procedures for threshold and concurrent slope estimates for psychophysics and speech intelligibility tests. *The Journal of the Acoustical Society of America*, *111*, 2801–2810.

Brown, C. A., & Bacon, S. P. (2009). Low-frequency speech cues and simulated electric-acoustic hearing. *The Journal of the Acoustical Society of America*, *125*, 1658–1665.

Büchner, A., Schüssler, M., Battmer, R. D., Stöver, T., Lesinski-Schiedat, A.Lenarz, T. (2009). Impact of low-frequency hearing. *Audiology and Neurotology*, *14*, 8–13.

Ching, T. Y. C., van Wanrooy, E., & Dillon, H. (2007). Binaural-bimodal fitting or bilateral implantation for managing severe to profound deafness: A review. *Trends in Amplification*, *11*(3), 161–192.

Cullington, H. E., & Zeng, F.-G. (2008). Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects. *The Journal of the Acoustical Society of America*, *123*(1), 450–461.

DIN EN ISO 389 8. (2004). *Acoustics – Reference zero for the calibration of audiometric equipment – Part 8: Reference equivalent threshold sound pressure levels for pure tones and circumaural earphones, standard of the international standards organization (ISO).* Geneva, Switzerland: International Standards Organization.

Durlach, N. I. (1963). Equalization and cancellation theory of binaural masking-level differences. *The Journal of the Acoustical Society of America*, *35*, 1206–1218.

Gantz, B. J., & Turner, C. W. (2003). Combining acoustic and electrical hearing. *The Laryngoscope*, *113*(10), 1726–1730.

Geißler, G., Büchner, A., Chalupper, J., & Battmer, R. D. (2015). *Just noticeable delay between electric and acoustic stimulation.* Paper presented at 18th Annual Meeting of the Deutsche Gesellschaft für Audiologie (DGA), Bochum, Germany.

George, E., Devocht, E., Chalupper, J., & Stokroos, R. (2015). *Key contributors to successful bimodal fitting.* Paper presented at 18th Annual Meeting of the Deutsche Gesellschaft für Audiologie (DGA), Bochum, Germany.

Gifford, R. H., Dorman, M. F., McKarns, S. A., & Spahr, A. J. (2007). Combined electric and contralateral acoustic hearing: Word and sentence recognition with bimodal hearing. *Journal of Speech, Language and Hearing Research*, *50*(4), 835–843.

Gifford, R. H., Dorman, M. F., Sheffield, S. W., Teece, K., & Olund, A. P. (2014). Availability of binaural cues for bilateral implant recipients and bimodal listeners with and without preserved hearing in the implanted ear. *Audiology and Neurotology*, *19*(1), 57–71.

Grange, J. A., & Culling, J. F. (2013). The benefit of cochlear implant users' head movement to speech intelligibility in noise. In T. Dau, S. Santurette, J. C. Dalsgaard, L. Tranebjærg, T. Andersen, & T. Poulsen (Eds.),, *Proceedings of ISAAR 2013: Auditory plasticity – Listening with the brain. 4th International symposium on auditory and audiological research, 2013* (pp. 389–396). Copenhagen, Denmark: The Danavox Jubilee Foundation.

Greenwood, D. (1990). A cochlear frequency-position function for several species—29 years later. *Journal of the American Academy of Audiology*, *87*(6), 2592–2605.

Haumann, S., Lenarz, T., & Büchner, A. (2010). Speech perception with cochlear implants as measured using a roving-level adaptive test method. *ORL; Journal for Otorhinolaryngology and Its Related Species*, *72*(6), 312–318.

Hohmann, V. (2002). Frequency analysis and synthesis using a gammatone filterbank. *Acta Acustica united with Acustica*, *88*(3), 433–442.

Ihlefeld, A., & Litovsky, R. Y. (2012). Interaural level differences do not suffice for restoring spatial release from masking in simulated cochlear implant listening. *PLoS ONE*, *7*(9), e45296, 1–9.

James, C. J., Fraysse, B., Deguine, O., Lenarz, T., Mawman, D., Ramos, A., & Sterkers, O. (2006). Combined electroacoustic stimulation in conventional candidates for cochlear implantation. *Audiology & Neurootology*, *11*(Suppl 1), 57–62.

Jones, H., Kan, A., & Litovsky, R. Y. (2014). Comparing sound localization deficits in bilateral cochlear-implant users and vocoder simulations with normal-hearing listeners. *Trends in Hearing*, *18*, 1–16.

Kan, A., & Litovsky, R. Y. (2015). Binaural hearing with electrical stimulation. *Hearing Research*, *322*, 127–137.

Kelvasa, D., & Dietz, M. (2015). Auditory model-based sound direction estimation with bilateral cochlear implants. *Trends in Hearing*, *19*, 1–16.

Kollmeier, B., Warzybok, A., Hochmuth, S., Zokoll, M. A., Uslar, V., Brand, T., & Wagener, K. C. (2015). The multilingual matrix test: Principles, applications, and comparison across languages: A review. *International Journal of Audiology*, *54*(Suppl 2), 3–16.

Laback, B., Egger, K., & Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hearing Research*, *322*, 138–150.

Léger, A. C., Moore, B. C. J., & Lorenzi, C. (2012). Abnormal speech processing in frequency regions where absolute thresholds are normal for listeners with high-frequency hearing loss. *Hearing Research*, *294*(1–2), 95–103.

Lenarz, T., James, C., Cuda, D., Fitzgerald O'Connor, A., Frachet, B., Frijns, J. H. M., & Uziel, A. (2013). European multi-centre study of the nucleus hybrid L24 cochlear implant. *International Journal of Audiology*, *52*(12), 838–848.

Li, N., & Loizou, P. C. (2008). A glimpsing account for the benefit of simulated combined acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *123*(4), 2287–2294.

Litovsky, R. Y. (2012). Spatial release from masking. *Acoustics Today*, *8*(2), 18–25.

Schafer, E. C., Amlani, A. M., Seibold, A., & Shattuck, P. L. (2007). A meta-analytic comparison of binaural benefits between bilateral cochlear implants and bimodal stimulation. *Journal of the American Academy of Audiology*, *18*(9), 760–776.

Schoof, T., Green, T., Faulkner, A., & Rosen, S. (2013). Advantages from bilateral hearing in speech perception in noise with simulated cochlear implants and residual acoustic hearing. *The Journal of the Acoustical Society of America*, *133*(2), 1017–1030.

Wagener, K. C., Kühnel, V., & Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache Teil I-III: Design, Optimierung und Evaluation des Oldenburger Satztests [Development and evaluation of a sentence test for the German language: Parts I-III: Design, optimization and evaluation of the Oldenburg sentence test]. *Zeitschrift für Audiologie*, *38*, 4–95.

Warzybok, A., Rennies, J., Brand, T., Doclo, S., & Kollmeier, B. (2013). Effects of spatial and temporal integration of a single early reflection on speech intelligibility. *The Journal of the Acoustical Society of America*, *133*(1), 269–282.

Wiggins, I. M., & Seeber, B. U. (2011). Dynamic-range compression affects the lateral position of sounds. *The Journal of the Acoustical Society of America*, *130*(6), 3939–3953.

Wilson, B. S., Finley, C., Lawson, D., Wolford, R., Eddington, D. K., & Rabinowizt, W. (1991). Better speech recognition with cochlear implants. *Nature*, *352*, 236–238.

Zeng, F.-G. (2002). Temporal pitch in electric hearing. *Hearing Research*, *174*(1), 101–106.