Investigating Bilateral Cochlear Implant Users' Localization of Amplitude- and Time-Panned Stimuli Produced Over a Limited Loudspeaker Arrangement

Fernandez, Janani

2022-03

Fernandez , J , Sivonen , V & Pulkki , V 2022 , ' Investigating Bilateral Cochlear Implant Users' Localization of Amplitude- and Time-Panned Stimuli Produced Over a Limited Loudspeaker Arrangement ' , American journal of audiology , vol. 31 , no. 1 , pp. 143-154 . https://doi.org/10.1044/20

http://hdl.handle.net/10138/342570
https://doi.org/10.1044/2021_AJA-21-00083

Research Article

Investigating Bilateral Cochlear Implant Users' Localization of Amplitude- and Time-Panned Stimuli Produced Over a Limited Loudspeaker Arrangement

Janani Fernandez,[a] Ville Sivonen,[b] and Ville Pulkki[a]

[a] Department of Signal Processing and Acoustics, Aalto University, Espoo, Finland [b] Head and Neck Center, Department of Otorhinolaryngology—Head and Neck Surgery, Helsinki University Hospital and University of Helsinki, Finland

ARTICLE INFO

Article History:
Received April 21, 2021
Revision received July 23, 2021
Accepted November 2, 2021

Editor-in-Chief: Ryan W. McCreery
Editor: Andrea Warner-Czyz

https://doi.org/10.1044/2021_AJA-21-00083

ABSTRACT

Objective: The objective of this study was to investigate the localization ability of bilateral cochlear implant (BiCI) users for virtual sound sources produced over a limited loudspeaker arrangement.
Design: Ten BiCI users and 10 normal-hearing subjects participated in listening tests in which amplitude- and time-panned virtual sound sources were produced over a limited loudspeaker setup with varying azimuth angles. Three stimuli were utilized: speech, bandpassed pink noise between 20 Hz and 1 kHz, and bandpassed pink noise between 1 kHz and 8 kHz. The data were collected via a two-alternative forced-choice procedure and used to calculate the minimum audible angle (MAA) of each subject, which was subsequently compared to the results of previous studies in which real sound sources were employed.
Result: The median MAAs of the amplitude-panned speech, low-frequency pink noise, and high-frequency pink noise stimuli for the BiCI group were calculated to be 20°, 38°, and 12°, respectively. For the time-panned stimuli, the MAAs of the BiCI group for all three stimuli were calculated to be close to the upper limit of the listening test.
Conclusions: The computed MAAs of the BiCI group for amplitude-panned speech were marginally larger than BiCI users' previously reported MAAs for real sound sources, whereas their computed MAAs for the time-panned stimuli were significantly larger. Subsequent statistical analysis indicated a statistically significant difference in the performances of the BiCI group in localizing the amplitude-panned sources and the time-panned sources. It follows that time-panning over limited loudspeaker arrangements may not be a useful clinical tool, whereas amplitude-panning utilizing such a setup may be further explored as such. Additionally, a comparison with the patient demographics indicated correlations between the results and the patients' age at time of diagnoses and the time passed between date of diagnosis and their implant surgeries.

Over the last decade, bilateral cochlear implantation has become increasingly common in many countries. This procedure provides various advantages over unilateral cochlear implantation, including an increase in overall speech comprehension in both quiet and noisy environments and improved sound localization abilities (Dunn et al., 2008; Litovsky et al., 2004; Tyler et al., 2006).

The improvement in localization may be attributed to the fact that the auditory system of a bilateral cochlear implant (BiCI) user receives more information in comparison to a unilateral cochlear implant user. Humans typically determine the location of sound sources by analyzing localization cues, such as the interaural time difference

---

Correspondence to Janani Fernandez: janani.fernandez@aalto.fi. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

(ITD) and the interaural level difference (ILD). The ILD cue is dependent on the head-shadowing effect, which is more pronounced at frequencies above 1 kHz (Shaw, 1974; Yost & Dye Jr, 1988). For frequencies up to 1.5 kHz, the ITD tends to be the more dominant localization cue; whereas at higher frequencies, temporal difference cues become increasingly ambiguous (Buell & Hafter, 1991; Moore, 2012). Note, however, that ITD cues are significantly altered by the signal encoding technique employed by cochlear implants (CIs), and previous research indicated that BiCI users primarily depend on ILD cues for localization (Seeber & Fastl, 2008; van Hoesel & Tyler, 2003). BiCI users also show limited ability to detect envelope ITDs (Senn et al., 2005; Seeber & Fastl, 2004; van Hoesel et al., 2009); although it is generally considered to be a weak localization cue (Blauert, 1997). Other localization cues, such as monaural cues and dynamic cues, are also altered in the case of BiCI users either due to microphone placement or the current technology implemented within CIs, with such devices primarily optimized for enhancing speech intelligibility. This is largely due to the limitations of electrical stimulation with current CI systems.

It can be noted from these limitations of the current CI system that although bilateral implantation allows BiCI users to localize sound, their localization skills are poorer in comparison to those of normal-hearing (NH) subjects. In the horizontal plane, it has been reported that BiCI users localize speech sound sources with a root-mean-square (RMS) error of 21.5°; on the other hand, their localization of noise sound sources was poorer, with an RMS error ranging from 8.1° to 43.4° (Grantham et al., 2007). Similarly, Neuman et al. (2007) utilized both speech and pink noise stimuli and reported the mean RMS error as 29°. In a more recent study, the RMS error of BiCI users was also 29° for noise stimuli, whereas for the NH control group, it was 6° for young participants and 5.4° for older participants (Dorman et al., 2016). Other studies of localization have reported their results using the minimum audible angle (MAA), that is, the smallest angle required between two sound sources for the subject to be able to perceive them as distinct sources. MAAs for NH listeners have been reported as low as 0.97° in the horizontal plane (Perrott & Saberi, 1990). For BiCI users, the MAA in the horizontal plane for young children with BiCIs ranged between 5° and 26°, whereas the NH control group attained MAAs ranging between 1° and 2° (Litovsky et al., 2006). It should be noted that BiCI subjects who had received their second implant before the age of 1 year performed in a manner similar to the NH control group. Senn et al. (2005) concluded that the MAAs of BiCI users ranged between 3° and 8°; however, this study had only five participants while Litovsky et al. (2006) had 13 participants in their study.

Studies of the localization abilities in BiCI users have primarily focused on the localization of real sound sources (Dorman et al., 2016; Nopp et al., 2004; van Hoesel & Tyler, 2003). Such studies are, therefore, limited by the number and positions of loudspeakers in the test setup. Typical solutions to this problem would require increasing the number of loudspeakers in the arrangement or investing in expensive hardware setups to dynamically change the loudspeaker positions; both of these solutions demand a large budget and sufficient space for the setup.

Another solution would be the utilization of virtual sound sources, which are sound sources perceived by the listener to be at a position that does not physically align with a loudspeaker direction. Such sources are typically synthesized by the application of a monophonic audio signal to a set of loudspeakers or a pair of headphones after suitable processing (Blauert, 1997). There have been several studies of BiCI users' localization of such sound sources produced over headphones. One such study utilized head-related transfer functions (HRTFs; Blauert, 1997) to create virtual sound sources by connecting binaural signals to the auxiliary inputs of the CI sound processors (in a manner similar to headphone playback with NH listeners) and reported that BiCI users could identify the correct source position in the horizontal plane with reasonable accuracy; whereas in the median plane, the subjects could only identify the correct hemisphere (Majdak, et al., 2011). Sechler et al. (2017) also employed HRTFs to create virtual sound sources, which were played to subjects via a virtual reality headset and reported that the RMS error of BiCI subjects was frequency dependent and had an average value of 23.81°, whereas the RMS error of the control group was 12.4° and was not azimuth angle dependent. Additionally, a case study found that the BiCI users had similar error values for real sound sources played over loudspeakers and for virtual sounds that were created by manipulating localization cues and directly transmitted to the sound processor via a cable (Seeber & Fastl, 2004). Finally, Moua et al. (2019) recorded a loudspeaker arrangement using a binaural dummy head in the listener position, with the recorded binaural signals subsequently transmitted to the BiCI users' device, in order to evaluate BiCI users' ability to localize stationary and moving virtual sound sources.

It should be noted that the aforementioned headphone-based localization studies typically either omit the effect of the individual HRTF or utilize generic HRTFs in place of individual HRTFs which may deviate from the true HRTFs of the listener and lead to poorer localization ability (Bronkhorst, 1995). Furthermore, dynamic localization cues cannot be studied with such a setup. Producing virtual sound sources instead over a loudspeaker arrangement should mitigate these disadvantages and could prove to be a useful clinical tool. However, to the best of the authors' knowledge, there have not been any studies investigating the BiCI users' localization ability of virtual sound sources produced directly over loudspeakers.

This study investigates the feasibility of utilizing virtual sound sources produced over a limited loudspeaker arrangement, as such a technique would be useful in routinely assessing localization abilities during clinic visits. The MAAs in the horizontal plane were evaluated for a BiCI group and a control group of NH participants. These results were subsequently compared to the MAAs previously reported in the literature in order to determine whether the BiCI group could localize virtual sound sources with a reasonable degree of accuracy. The production of virtual sound sources employed an arrangement of only two loudspeakers, which is a requirement easily fulfilled by existing a standard clinical setup for a speech-in-noise test. The computed MAA results were further analyzed by computing the cues present in time- and amplitude-panned sound sources by employing an auditory model. In addition, the patient demographic data were analyzed in order to determine if there were any statistically relevant correlation between the data and the results of the listening test.

Method

Listeners

Ten BiCI users participated in the listening tests. They were a mix of prelingually and postlingually deafened subjects. At the time the test was administered, each BiCI subject had a minimum of 2 years of BiCI hearing experience. Table 1 lists the relevant patient demographic data for the BiCI subjects, labeled CI1–CI10. All BiCI subjects utilized Med-El brand devices, with the exception of subjects CI7, CI8, and CI10, who utilized Cochlear devices. Furthermore, all BiCI subjects were female, with the exception of subject CI6.

The control group consisted of 10 NH subjects, whose ages were matched as closely as possible to that of the BiCI group. The mean age of the control group was 39.9 years, while that of the BiCI group was 40.9 years. The control group consisted of female participants, with the exception of subjects NH5 and NH8.

Equipment

The listening tests were conducted in a sound-treated booth of dimensions 2.54 m × 2.64 m × 2.10 m. There were eight (Genelec 8050A) loudspeakers present in the room, two of which were utilized in the tests. The base angle (θ_0) between the loudspeakers utilized in the tests was 45° in the horizontal plane. The chair upon which the test subject sat was placed at a distance of approximately 1 m from the loudspeakers. The test software was run in MATLAB Version 2016b, hosted on a mid-2012 model MacBook Pro, with a MOTU UltraLite mk3 Hybrid as the soundcard. The listener's responses were collected via the left and right buttons on a computer mouse. The test setup is depicted in Figure 1.

Stimuli

There were three different stimuli used in the tests: two pink noise stimuli and one speech stimulus. The speech consisted of two words chosen randomly from a database of phonetically balanced bisyllabic Finnish words (Jauhiainen, 1974), recorded by a female speaker at a sampling rate of 44.1 kHz. The noise stimuli were pink noise samples of a sampling rate of 48 kHz, filtered through two bandpass filters (n = 400 samples): one with a bandpass range of 20 Hz to 1 kHz, the other with a bandpass range of 1 kHz to 8 kHz. The cutoff frequencies of the noise stimuli were selected based on the frequencies at which the ITD and ILD cues are the more dominant cues in localization. The duration of the noise samples was 0.2 s, while that of the speech samples varied depending on the words chosen. The stimuli were presented at an average sound pressure level of 65 dB(A).

Production of the Virtual Sound Sources

Virtual sound sources may be broadly categorized based on which property of the sound signal is manipulated to create the sound source. Amplitude-panned sound sources are created by applying the same signal to two or more loudspeakers, but with different amplitudes (Pulkki, 1997); time-panned sound sources are instead produced by altering the delay between the output of two or more loudspeakers. This study investigates the effect of both types of virtual sound sources.

The method of amplitude panning implemented in the test was vector-base amplitude-panning, which may be formulated as follows:

$$[g_1\, g_2] = [p_1\, p_2]\begin{pmatrix} l_{l1}\, l_{l2} \\ l_{r1}\, l_{r2} \end{pmatrix}^{-1}, \tag{1}$$

where $[g_1\, g_2]$ are the gains applied to each loudspeaker signal, $[p_1\, p_2]$ is the panning direction expressed in vector form, and $[l_{l1}\, l_{l2}]$ and $[l_{r1}\, l_{r2}]$ are the position vectors of the left and right loudspeaker, respectively (Pulkki, 1997).[1]

The time-panning method was implemented by introducing an interchannel time delay (ICTD) between the loudspeakers, the value for which was computed using

---

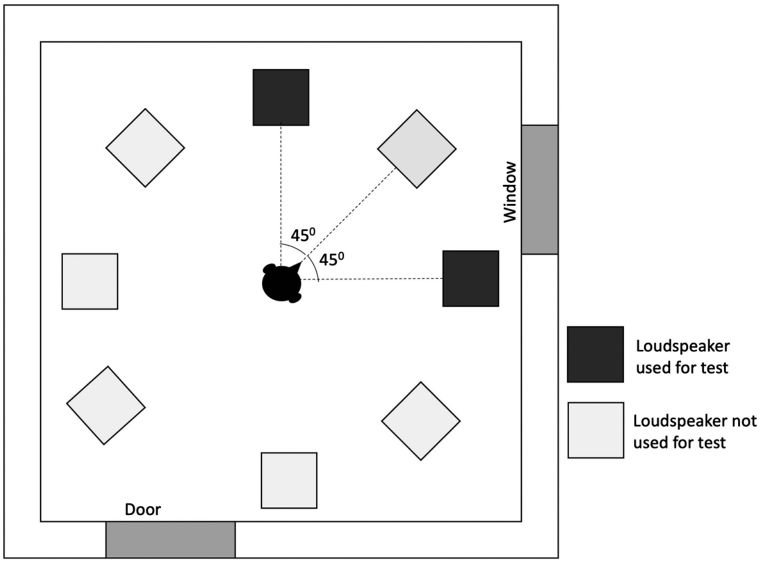
[1] Implemented utilizing the MATLAB library from https://github.com/polarch/Vector-Base-Amplitude-Panning.

Table 1. Patient demographic data: The subjects in the CI group are labeled as CI1–CI10.

Subject	Age (years)	Age when diagnosed (years)	Implant side	Time deafened (years)		Hearing aid use (years)		BiCI experience (years)	Implant type		Sound processor	
			CI1	Left	Right	Left	Right		Left	Right	Left	Right
CI1	41.4	30.5	Right	6.2	8.5	> 20	> 20	2.4	Concerto Flex 28	Synchrony Flex 24	Sonnet	Opus2
CI2	37.8	6.1	Left	27.9	29.0	30	31	2.6	Concerto Flex 28	Concerto Flex 28	Rondo	Rondo
CI3	37.4	32.9	Left	1.1	2.6	0	18	1.9	Concerto Flex 28	Synchrony Flex 28	Opus2	Sonnet
CI4	37.9	21.3	Right	12.9	14.1	> 20	> 20	2.6	Concerto Flex 28	Synchrony Flex 24	Sonnet	Opus2
CI5	61.4	34.3	Left	10.4	24.3	24	26	2.7	C40+ Standard	Concerto Flex 28	Opus2	Sonnet
CI6	55.2	12.2	Right	27.3	40.2	46	0	2.7	C40+ Standard	Concerto Flex 28	Sonnet	Opus2
CI7	43.8	5.3	Left	30.5	35.8	31	35	2.7	CI24RE (CA)	CI512	CP910	CP910
CI8	42.1	1.7	Left	32.7	37.6	35	39	2.8	CI512	CI512	CP810	CP910
CI9	50.5	0	Left	43.0	47.8	33	37	2.6	Sonata ti 100 Standard	Synchrony Flex 28	Opus2	Sonnet
CI10	21.7	9.2	Left	4.4	10.3	11	17	2.2	CI24RE (CA)	CI522	CP810	CP910

Note. Their age (in years) at the time of the test is listed in the age column, whereas the age when diagnosed column lists their age (in years) at which they were diagnosed as deaf, that is, when their pure-tone average (PTA) between 500 Hz and 4 kHz in both ears was more than 80 dB hearing level (HL). The hearing aid use column lists the number of years the subjects wore a hearing aid in their left and right ears prior to cochlear implant (CI) surgery. The implant side lists the side of the first implant (CI1). The time deafened column is the ear-specific time between a PTA of > 80 dB HL and CI surgery in years. The bilateral cochlear implant (BiCI) experience column lists the time between the second implantation surgery and the date of the listening test in years.

Figure 1. Diagram of the test setup.



an empirical formula that was extrapolated from a study by Lee and Rumsey (2013), with minor corrections from Lee (2017). The formula utilized for this study was:

$$\tau_{\mathrm{ICTD}} = 0.025\frac{30\theta}{\theta_0}, \mathit{for}\ \theta \leq \frac{2\theta_0}{3}, \quad (2)$$

$$\tau_{\mathrm{ICTD}} = 0.05\left(\frac{30\theta}{\theta_0} - 20\right) + 0.5, \mathit{for}\ \theta > \frac{2\theta_0}{3}, \quad (3)$$

where τ_{ICTD} is the delay in milliseconds, θ is the intended source position or panning angle, and θ_0 is the loudspeaker base angle. In both panning techniques, the maximum absolute value of θ was 45°, whereas the minimum absolute value was 1°.

Test Procedure

The administered listening test was a two-alternative forced-choice (2AFC) task. A panning technique and a stimulus were chosen at random. In the case of the speech stimulus, a set of two words was chosen at random from the database. Two sequential virtual sound sources of short duration were then synthesized in directions {θ, −θ} or {−θ, θ}. In this manner, they appear to emanate from the left and right hemispheres (with respect to each other), consecutively. The precise location of the panned sound sources is determined by the panning angle value θ. The two chosen stimuli were played in a random order with a 0.1-s pause between them. The test subjects were asked to answer the question "In which direction is the sound going?", and were to provide their response by clicking either the left or right mouse button to indicate "left" or "right," respectively.

To avoid listening fatigue, several measures were taken. First, the subjects had the option of taking a break every 10 min. Second, six adaptive tracks were used, that is, one for each combination of panning technique and stimulus type. Each track followed a "3-down 1-up" transformed staircase procedure according to Levitt (1971), corresponding to the 79.4% of the psychometric function. The initial value of θ was either −45° or 45°. The initial step size was 10°. After one reversal, the step size was reduced to 5°, and after one more reversal, the step size was finally set to 2°. After a total of 10 reversals, the adaptive track was terminated.

Data Analysis

Due to the modification of the adaptive track, the calculation of the results was completed as follows to obtain the 70% convergence point. The data from each listening test were used to calculate a vector of the percentage of correct answers for each value of θ. This vector was then fitted to a curve by employing the "interp1" built-in function of MATLAB. From this curve, the angle corresponding to 70% correct responses was extrapolated, which is considered to be the MAA for the subject for the corresponding panning technique and stimulus.

Statistical Analysis

Due to the non-Gaussian nature of the results, the data were analyzed with the Friedman test. The statistical analysis was performed with the Statistics and Machine Learning MATLAB Toolbox, Version 11.1. The dependent variable in the model was the calculated MAAs from the listening tests for the two groups, BiCI group and the NH control group, whereas the independent variables were the stimuli and panning method. For the analysis of the patient demographic data, the Spearman correlation coefficient was computed with 95% confidence intervals. The dependent variable in the model was the calculated MAAs from the listening tests for the BiCI, whereas the independent variables were the patient demographic data.

Results

The computed MAAs of the BiCI and the NH control group are presented in Tables 2 and 3, respectively. The tables also report the medians and interquartile ranges of the results for the various stimuli and the two panning methods. Figure 2 depicts a box plot of the MAA results for the amplitude-panned stimuli. The red lines represent the medians of the computed MAAs for the BiCI and the NH groups for each stimulus. It should be noted that some of the results were 2° or 90°, which correspond to the lower and upper limits of the listening test, respectively. It follows, therefore, that the MAAs of these particular listeners lie outside the limits of the test, due to the step size and the loudspeaker positions. These results have been omitted from the calculation of the medians and interquartile ranges reported in Tables 2 and 3 (up to one subject in the case of amplitude-panned stimuli and up to four subjects for the time-panned stimuli, for the BiCI group).

A Friedman test was carried out on the computed MAAs of all 10 members of the BiCI group for the time-panned and amplitude-panned speech stimuli. A statistically significant difference in the performances was identified ($\chi^2 = 10$, $p = .0016$, $df = 19$). Similarly, the Friedman tests of the MAAs of the noise stimuli showed a statistically significant difference in the BiCI group's performances for amplitude-panned and time-panned stimuli ($\chi^2 = 8$, $p = .0047$, $df = 19$ for the low-frequency stimulus, and $\chi^2 = 6.4$, $p = .011$, $df = 19$ for the high-frequency stimulus). No statistically significant difference was found for the control group in the case of the high-frequency noise stimulus ($\chi^2 = 1.6$, $p = .205$, $df = 19$) and the speech stimuli ($\chi^2 = 0.4$, $p = .52$, $df = 19$); however, there was a statistically significant difference in performance for the low-frequency stimulus ($\chi^2 = 6.4$, $p = .011$, $df = 19$).

A Friedman test was also carried out comparing the performances across the three stimuli for each panning method. There was a statistical difference found in the performances for the BiCI group ($\chi^2 = 7.74$, $p = .020$, $df = 29$). Subsequently, a Wilcoxon signed-ranks test was carried out with a Bonferroni correction applied, resulting in a significance level at $p < .017$. The only statistically significant difference was found between the amplitude-panned speech stimuli and the high-frequency noise stimulus ($p = .0059$). There were no statistically significant difference in the localization performances between the amplitude-panned speech stimuli and the low-frequency noise stimulus ($p = .027$) and between the performances for the two noise stimuli ($p = .679$). The MAA results for the control group did not show any statistically significant difference in performance across the amplitude-panned stimuli ($\chi^2 = 0.67$, $p = .716$, $df = 29$).

Table 2. Results of the listening tests for the bilateral cochlear implant group.

Subject	Amplitude panning—pink noise, 20 Hz–1 kHz (deg)	Amplitude panning—pink noise, 1 kHz–8 kHz (deg)	Amplitude panning—speech stimuli (deg)	Time panning—pink noise, 20 Hz–1 kHz (deg)	Time panning–pink noise, 1 kHz–8 kHz (deg)	Time panning—speech stimuli
CI1	8	7	4	> 90	> 90	87
CI2	36	38	14	> 90	68	87
CI3	23	9	8	79	88	73
CI4	9	9	11	78	46	78
CI5	14	18	10	77	> 90	76
CI6	79	77	8	79	> 90	41
CI7	10	53	16	80	74	>90
CI8	89	> 90	47	> 90	77	>90
CI9	> 90	71	39	> 90	85	>90
CI10	20	39	16	85	90	84
Mdn (CI group)	20	38	12.5	79	75.5	79
Interquartile range (CI group)	37	48.5	8	2	17	12.5

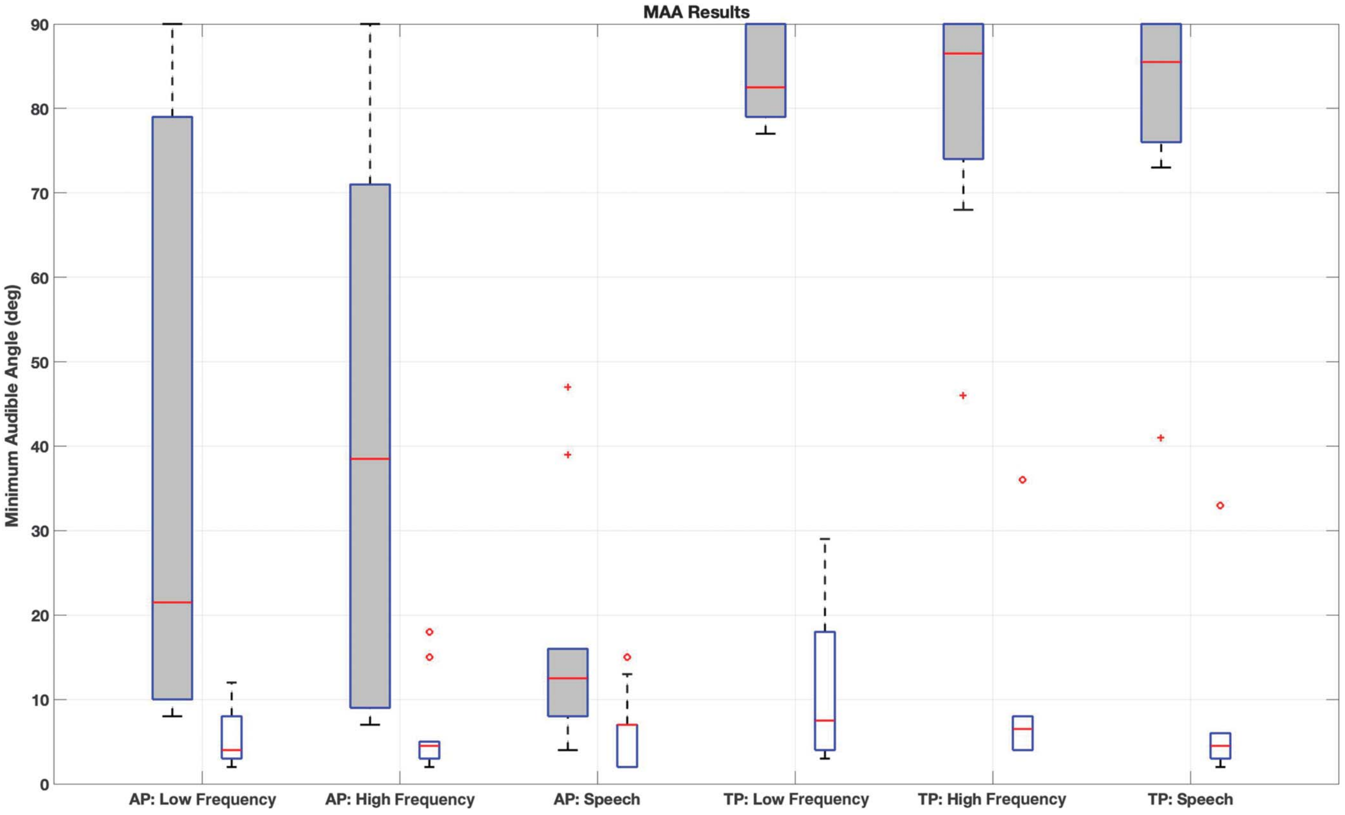
Note. The ">" symbol denotes those results that cannot be accurately measured by the setup utilized in the listening tests and are, therefore, considered outliers and omitted from the median and interquartile range calculations. deg = degrees; CI = cochlear implant.

Table 3. Results of the listening tests for the normal-hearing (NH) group.

Subject	Amplitude panning—pink noise, 20 Hz–1 kHz (deg)	Amplitude panning—pink noise, 1 kHz–8 kHz (deg)	Amplitude panning—speech stimuli (deg)	Time panning—pink noise, 20 Hz–1 kHz (deg)	Time panning—pink noise, 1 kHz–8 kHz (deg)	Time panning—speech stimuli
NH1	4	< 2	< 2	29	4	6
NH2	12	15	7	18	8	3
NH3	4	< 2	7	8	8	6
NH4	< 2	3	6	5	4	4
NH5	3	5	< 2	4	4	6
NH6	8	18	15	25	36	33
NH7	8	3	13	7	8	< 2
NH8	3	5	< 2	4	4	3
NH9	6	4	7	15	7	3
NH10	< 2	5	7	3	6	5
Mdn (control group)	5	5	7	7.5	6.5	5
Interquartile range (control group)	3.1	5.8	3.5	9.4	9.6	9.5

Note. The "<" symbol denotes those results that cannot be accurately measured by the setup utilized in the listening tests and are, therefore, considered outliers and omitted from the median and interquartile range calculations.

Figure 2. Box plot of the results. The shaded boxes represent the results of the bilateral cochlear implant (BiCI) group, while the results of the normal-hearing group are in unshaded boxes. The horizontal line within the box plot denotes the median (excluding outliers) of the results, the ends of the box denote the 75th and 25th percentile results, the whiskers denote the extreme data points, and the outliers are denoted by "+" for the BiCI group and "o" for the NH group. MAA = minimum audible angle; TP = time-panning, AP = amplitude-panning, deg = degrees.



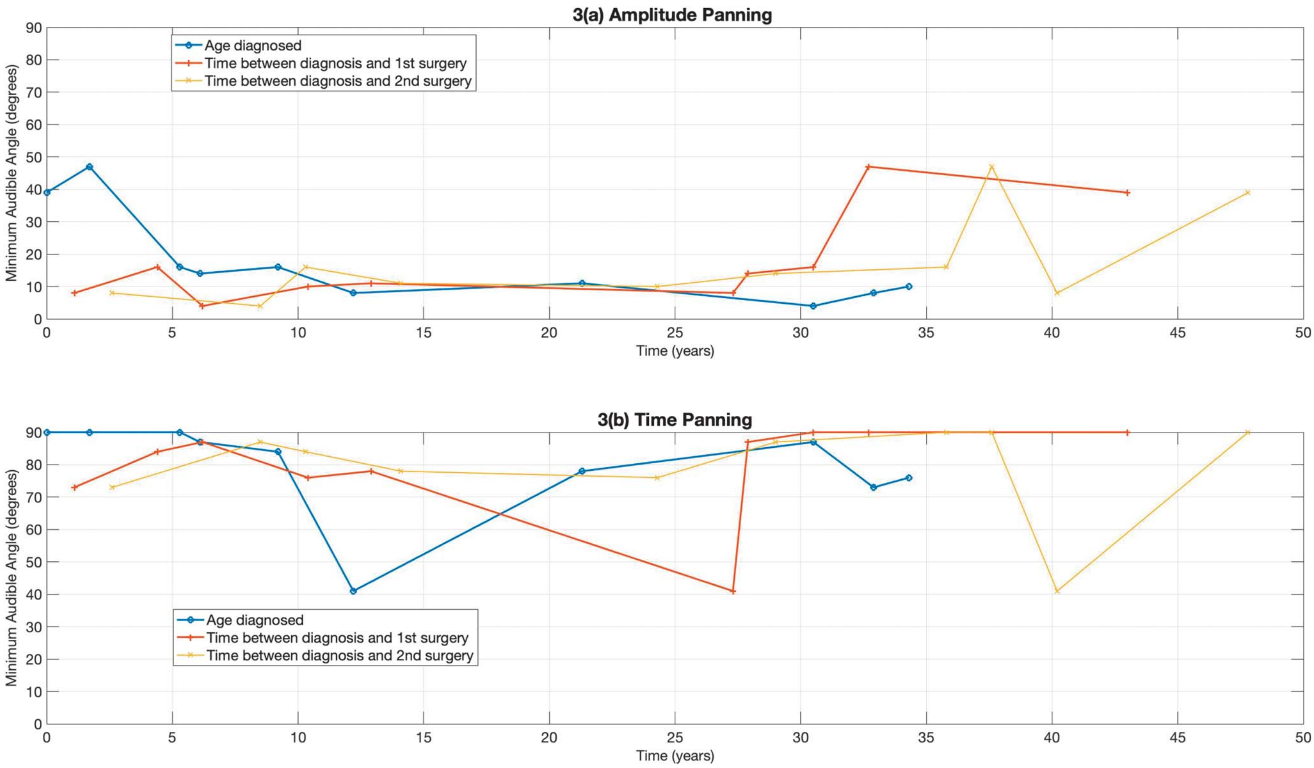
A Friedman test was also carried out on the computed MAAs of the BiCI group for the time-panned stimuli. There was no statistically significant difference found (χ^2 = 1.72, p = .42, df = 29). Similarly, the MAA results for the control group did not show any statistically significant difference in performance across the time-panned stimuli (χ^2 = 3.39, p = .183, df = 29).

The analysis of the patient demographic data was also performed. Figure 3 visualizes the data found to be most statistically significant. Based on the MAA results, the BiCI users who were diagnosed as deafened at a younger age tended to perform better at localizing the amplitude-panned speech stimuli, Spearman rho (r_S) = −0.84, p = .023. Additionally, the BiCI users who had a shorter time period between the date of diagnosis and the date of their first implantation surgery tended to have smaller MAAs for amplitude-panned stimuli (r_S = 0.68, p = .029). The MAA results for time-panned speech stimuli showed similar statistically significant correlations. The BiCI users who were diagnosed as deafened at a younger age tended to have larger MAAs for the time-panned sources, r_S = −0.78, p < .01, and the users with shorter time periods between the date of diagnoses and the first implantation surgery tended to perform better in the localization test, r_S = 0.68, p = .029. It should be noted that no corrections were applied for these analyses.

Discussion

The statistical analyses indicated that the BiCI group's localization performances for the time-panned stimuli were different from their performances for the amplitude-panned stimuli. The control group did not have similar differences in two of the three stimuli cases. From Table 2, it can be seen that the median values for the amplitude-panned stimuli are much smaller than the median values for time-panned stimuli for the BiCI group. Additionally, the MAA values computed for the time-panned stimuli in the case of the BiCI group were noticeably large and close to the upper limit of the test in almost all instances (the exception being subjects CI4 in the high-frequency noise stimulus case and CI6 in the speech stimuli case). These values may have been attained by random guessing and should not be heavily relied upon. The median values of the NH control group were larger than previously reported in literature, but they were, nevertheless, all between 2.5° of each other,

Figure 3. Comparison of the patient demographic data with the computed minimum audible angles. The age at which the bilateral cochlear implant user was diagnosed as deaf is marked with "o" symbols, the time period between the date of diagnosis and the date of the first implantation surgery is marked with "x" symbols, and the time period between the date of diagnosis and the date of the second implantation surgery is marked with "+" symbols. deg = degrees.



indicating the control group were able to localize to the correct hemisphere with a reasonable degree of accuracy. It follows from these results that the BiCI users may be able to localize amplitude-panned sources more accurately than time-panned sources, the latter of which BiCI users may not be able to adequately localize.

The Friedman tests also revealed that the stimuli may have had an effect on the performances of the BiCI group in the case of amplitude panning. The BiCI group had a smaller median value for the speech stimuli than for the noise stimuli. The interquartile ranges were also larger in magnitude for the noise stimuli than for the speech stimuli. The medians of the results for the various amplitude-panned stimuli show a similar trend. This difference in performance across stimuli is not visible in the results for the NH group, and the results of the Friedman tests did not reveal significant differences in the performance across stimuli for both amplitude- and time-panned sources for this group.

The difference in localization between the noise stimuli and speech stimuli for the BiCI group may be attributed to the short duration of the stimuli, since the speech stimuli utilized for the test were longer in duration compared to the noise stimuli. It may have been difficult for the BiCI group to perceive the short noise stimuli due to the processing performed within their CI devices. Senn et al. (2005) utilized 1,000-ms white noise stimuli in listening tests, while Litovsky et al. (2006) reported that preliminary tests using noise burst stimuli were found to have unreliable results. Future work could, therefore, involve gathering more data on the performance of the BiCI group in a similar listening test utilizing longer noise stimuli and determine if these results are reliable and comparable to MAAs for real sound sources.

It may be noted from Table 3 that the range of MAAs for the amplitude-panned speech stimuli for the BiCI group was 4°–47°, while that of the NH group was 2°–15°. These ranges were larger than the MAA ranges of real sound sources reported by Litovsky et al. (2006), which were 5°–40° for the BiCI group and 1°–2° for the NH group, and Senn et al. (2005), which was 3°–8° for the BiCI group. However, the participants in the study of Litovsky et al. (2006) were 13 pediatric patients, while Senn et al. (2005) had only five adult participants, so the difference in MAAs may also be due to the difference in participant demographics. The greater ranges may have also been due to a difference in auditory cues present in the amplitude-panned sources produced over the limited loudspeaker setup.

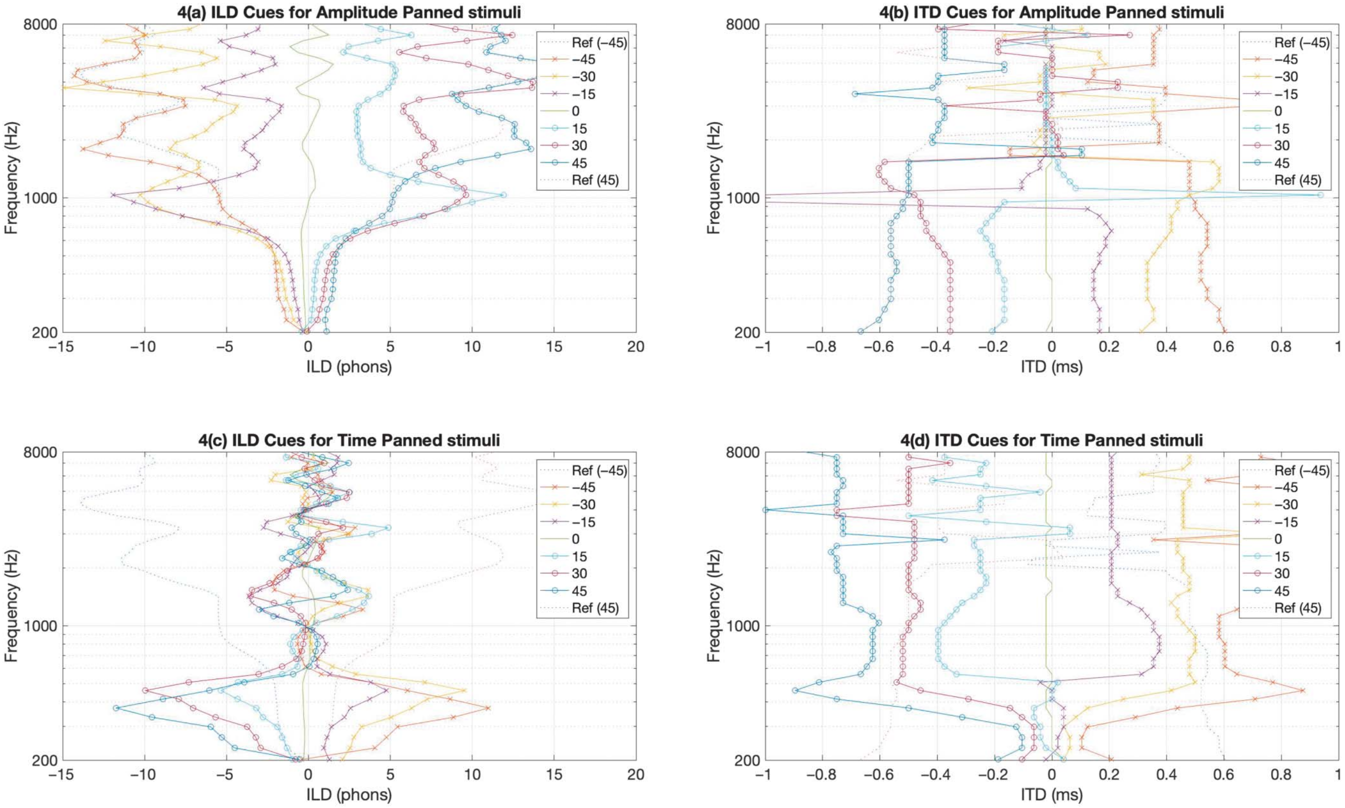
Therefore, to better determine how the panning methods impacted the cues, the ILDs and ITDs of the amplitude- and time-panned sources were simulated for the stimuli utilized in the tests. The ear canal signals were modeled by filtering the stimuli with HRTFs measured from the position of a mock-up hearing-aid frontal microphone placed on a dummy head (Sivonen, 2011). A simple binaural auditory model comprising a gammatone filter bank, loudness model, and specific binaural difference computation was utilized to estimate the ITDs and ILDs (Pulkki, 2002).

Figure 4 depicts the simulated ILD and ITD results for different source directions, when utilizing the amplitude- and time-panning methods implemented for the listening test. For amplitude-panned stimuli, it can be observed in Figure 4(a) that the simulated ILD cue is relatively stable and in the correct hemisphere above 1.1 kHz. At frequencies below 500 Hz, the simulated ILD cue is relatively small in magnitude, yet it still indicates the correct hemisphere. In the frequencies between these thresholds, it is comparatively more ambiguous, but still corresponds to the correct direction. It is also noted that the magnitude of these ILD cues vary from the cues given by the reference real sound sources at −45° and 45°. The ITD cues provided by the amplitude-panned stimuli, as shown in Figure 4(b), are comparatively stable at low frequencies and indicate the same hemisphere as θ; therefore, they correspond to the correct hemisphere. At frequencies above 1 kHz, the ITD cue is comparatively more ambiguous and unstable, as the instabilities caused by the short wavelength, relative to the size of the human head, are further exacerbated by the interference of the signals arriving from each loudspeaker, thus causing spectral cancellations and phase modifications to the sound signals arriving at each ear.

The deviation in magnitude of the ILD cues (when compared to the ILD cues of a real source) may be responsible for the wider MAA range achieved by the BiCI users and the NH control group with virtual sound sources, in comparison to the MAA ranges reported in previous studies that employed real sound sources. However, while the MAA range for amplitude-panned speech stimuli was larger in magnitude compared to the results published in previous literature, they do suggest that virtual sound sources synthesized over a limited loudspeaker setup using amplitude panning are still perceived in the correct hemisphere by BiCI users. Thus, an avenue for future work could involve further exploring the utilization of amplitude panning in clinics as a testing tool in the evaluation of BiCI users' localization abilities by comparing their performance of these sound sources against their performance with real sound sources. The technique utilized for these tests can be conveniently implemented in clinics that already have a suitable loudspeaker setup, such as the loudspeaker arrangement required for speech-in-noise tests and, therefore, can be implemented during routine clinic visits and follow-ups with patients. There is also the possibility of training BiCI users to improve their ability to localize sound by utilizing virtual sound sources (Majdak et al., 2010), and, therefore, this method of production over a limited loudspeaker setup may be useful for future clinical or even at-home training.

For the time-panned stimuli, in Figure 4(c), the simulated ILDs are relatively stable at low frequencies, but

Figure 4. The interaural time difference (ITD) and interaural level difference (ILD) cues of the amplitude- and time-panned stimuli. The dotted lines are the stimuli for real sources located at positions −45° and +45° azimuth. Sources at positions to the left of the listener (−45° to −1°) are marked with "x" symbols, while sources at positions to the right of the listener are marked with "o" symbols.



indicate the source position is located in the direction opposite to the intended source position. At frequencies larger than 500 Hz, the ILD cue for the time-panned stimuli is comparatively ambiguous. The ITD cues, shown in Figure 4(d), are relatively stable and indicate the same hemisphere as θ, and are also larger in magnitude when compared to the ITD cues for amplitude-panned stimuli.

The difference in performance of the NH group for both panning methods in the case of the low-frequency noise stimulus may have been due to the misleading ILD cue in combination with the low magnitude of the ITD cues in the time-panned stimuli for frequencies below 1 kHz. For the BiCI users, one may infer that due to their reliance on ILD cues, the ILD cue present in time-panned sources would mislead BiCI users; however, the computed MAA results for the BiCI group, as well as the percentage of correct responses achieved at the angles tested, do not provide strong evidence in favor of this hypothesis. Whereas some members of the BiCI group were not able to localize the time-panned stimuli, other test participants were able to localize the stimuli panned at larger θ values, but struggled with smaller values of θ. This indicates that they may not solely depend on ILD cues when localizing time-panned speech stimuli. For larger θ values, it is possible that the envelope ITD cue was sufficiently large enough to be perceived by BiCI users, as studies have reported that envelope ITD cues were capable of being detected by BiCI users (Seeber & Fastl, 2004; Senn et al., 2005; van Hoesel et al., 2009). It is possible that the ability of individual BiCI users to localize the time-panned speech stimuli is linked to their ability to perceive envelope ITD cues. Therefore, it could be beneficial to test this hypothesis in a future study.

Conclusions

This article investigated the localization ability of BiCI users, when they are presented with virtual sound sources produced over a limited loudspeaker arrangement. The MAAs for a group of BiCI users and a control group of NH listeners were calculated for amplitude- and time-panned speech and noise stimuli in the horizontal plane via a 2AFC listening test.

The calculated MAAs for the NH control group were found to be marginally larger than the MAAs found for real sound sources, as reported previously in the literature; however, the NH listeners were able to localize the panned stimuli in an otherwise reliable manner. For the BiCI group, the range of calculated MAAs for the amplitude-panned speech stimuli was found to be slightly larger than the ranges of MAAs for real sound sources reported in previous literature (Litovsky et al., 2006; Senn et al., 2005). A simulation of the localization cues delivered by both amplitude-panned virtual sources and real sources suggests slightly altered ILD cues for amplitude-panned sources when compared to the cues generated by a real sound source at the same position. The MAAs derived for the amplitude-panned noise stimuli, however, were not consistent with the MAA values previously reported in literature. The Friedman test carried out on the MAAs computed for the amplitude-panned stimuli also indicated that there was a significant difference in performance across stimuli. This is theorized to have been caused by the short time duration of these stimuli, as previous studies of a similar nature had utilized noise stimuli of longer duration. Regarding time-panned stimuli, the calculated MAAs for the BiCI group were found to be close to 90°, which was the highest possible value supported by the test apparatus. The median MAAs of the BiCI group for all three time-panned stimuli were larger in magnitude than those calculated for the amplitude-panned stimuli. The statistical analysis concluded that there was a statistically significant difference in the performances of the BiCI group in localizing the amplitude-panned sources and the time-panned sources across all three stimuli, thus indicating that BiCI users are not able to localize time-panned stimuli as well as they were able to localize the amplitude-panned stimuli. Additionally, the computed MAAs and the patient demographic data were compared, and the resultant Spearman correlation coefficients implied that the age at which the BiCI user was diagnosed as deafened, and the time period between the date of diagnosis and the date of the first implantation surgery were statistically significant factors affecting their performance in localizing the amplitude-panned speech stimuli.

While the small number of participants in the listening tests (10) prevents a generalizable conclusion, the results do indicate that amplitude-panning, utilizing a limited loudspeaker arrangement, is indeed capable of delivering localization cues to BiCI users in a manner similar to that of real sound sources when speech stimuli are utilized. This is of particular interest, as similar loudspeaker arrangements are already available in the majority of clinics, since they are often employed for speech-in-noise tests. The generation of virtual sound sources may, therefore, be used to evaluate the localization abilities of BiCI users, without the need for additional and potentially expensive setups. Furthermore, utilization of amplitude panning opens up the possibility for training BiCI users, since previous literature involving NH listeners has indicated that localization abilities can improve with training.

References

Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization*. MIT Press.

Bronkhorst, A. W. (1995). Localization of real and virtual sound sources. *The Journal of the Acoustical Society of America, 98*(5), 2542–2553. https://doi.org/10.1121/1.413219

Buell, T. N., & Hafter, E. R. (1991). Combination of binaural information across frequency bands. *The Journal of the Acoustical Society of America, 90*(4), 1894–1900. https://doi.org/10.1121/1.401668

Dorman, M. F., Loiselle, L. H., Cook, S. J., Yost, W. A., & Gifford, R. H. (2016). Sound source localization by normal-hearing listeners, hearing-impaired listeners and cochlear implant listeners. *Audiology and Neurotology, 21*(3), 127–131. https://doi.org/10.1159/000444740

Dunn, C. C., Tyler, R. S., Oakley, S., Gantz, B. J., & Noble, W. (2008). Comparison of speech recognition and localization performance in bilateral and unilateral cochlear implant users matched on duration of deafness and age at implantation. *Ear and Hearing, 29*(3), 352–359. https://doi.org/10.1097/AUD.0b013e318167b870

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., & Haynes, D. S. (2007). Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. *Ear and Hearing, 28*(4), 524–541. https://doi.org/10.1097/AUD.0b013e31806dc21a

Jauhiainen, T. (1974). *An experimental study of the auditory perception of isolated bi-syllable Finnish words* [Doctoral dissertation, University of Helsinki].

Lee, H. (2017). *Perceptually Motivated Amplitude Panning (PMAP) for accurate phantom image localization*. Audio Engineering Society.

Lee, H., & Rumsey, F. (2013). Level and time panning of phantom images for musical sources. *Journal of the Audio Engineering Society, 61*(12), 978–988.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America, 49*(2B), 467–477. https://doi.org/10.1121/1.1912375

Litovsky, R. Y., Johnstone, P. M., Godar, S., Agrawal, S., Parkinson, A., Peters, R., & Lake, J. (2006). Bilateral cochlear implants in children: Localization acuity measured with minimum audible angle. *Ear and Hearing, 27*(1), 43–59. https://doi.org/10.1097/01.aud.0000194515.28023.4b

Litovsky, R. Y., Parkinson, A., Arcaroli, J., Peters, R., Lake, J., Johnstone, P., & Yu, G. (2004). Bilateral cochlear implants in adults and children. *Archives of Otolaryngology—Head & Neck Surgery, 130*(5), 648–655. https://doi.org/10.1001/archotol.130.5.648

Majdak, P., Goupell, M. J., & Laback, B. (2010). 3-D localization of virtual sound sources: Effects of visual environment, pointing method, and training. *Attention, Perception, & Psychophysics, 72*(2), 454–469. https://doi.org/10.3758/APP.72.2.454

Majdak, P., Goupell, M. J., & Laback, B. (2011). Two-dimensional localization of virtual sound sources in cochlear-implant listeners. *Ear and Hearing, 32*(2), 198–208. https://doi.org/10.1097/AUD.0b013e3181f4dfe9

Moore, B. C. J. (2012). *An introduction to the psychology of hearing*. Brill.

Moua, K., Kan, A., Jones, H. G., Misurelli, S. M., & Litovsky, R. Y. (2019). Auditory motion tracking ability of adults with normal hearing and with bilateral cochlear implants. *The Journal of the Acoustical Society of America, 145*(4), 2498–2511. https://doi.org/10.1121/1.5094775

Neuman, A. C., Haravon, A., Sislian, N., & Waltzman, S. B. (2007). Sound-direction identification with bilateral cochlear implants. *Ear and Hearing, 28*(1), 73–82. https://doi.org/10.1097/01.aud.0000249910.80803.b9

Nopp, P., Schleich, P., & D'haese, P. (2004). Sound localization in bilateral users of MED-EL COMBI 40/40+ cochlear implants. *Ear and Hearing, 25*(3), 205–214. https://doi.org/10.1097/01.AUD.0000130793.20444.50

Perrott, D. R., & Saberi, K. (1990). Minimum audible angle thresholds for sources varying in both elevation and azimuth. *The Journal of the Acoustical Society of America, 87*(4), 1728–1731. https://doi.org/10.1121/1.399421

Pulkki, V. (1997). Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society, 45*(6), 456–466.

Pulkki, V. (2002). Microphone techniques and directional quality of sound reproduction. *Audio Engineering Society,* 112. Retrieved from http://www.aes.org/e-lib/browse.cfm?elib=11353

Sechler, S. D., Valdes, A. L., Waechter, S. M., Simoes-Franklin, C., Viani, L., & Reilly, R. B. (2017). Virtual reality sound localization testing in cochlear implant users. In *Neural Engineering (NER), 2017 8th International IEEE/EMBS Conference On* (pp. 379–382). IEEE.

Seeber, B., & Fastl, H. (2004). Localization cues with bilateral cochlear implants investigated in virtual space—A case study. *Proc Joint Congress CFA/DAGA, 4,* 213–214. http://mediatum.ub.tum.de/doc/1138418/file.pdf

Seeber, B. U., & Fastl, H. (2008). Localization cues with bilateral cochlear implants. *The Journal of the Acoustical Society of America, 123*(2), 1030–1042. https://doi.org/10.1121/1.2821965

Senn, P., Kompis, M., Vischer, M., & Haeusler, R. (2005). Minimum audible angle, just noticeable interaural differences and speech intelligibility with bilateral cochlear implants using clinical speech processors. *Audiology and Neurotology, 10*(6), 342–352. https://doi.org/10.1159/000087351

Shaw, E. (1974). Transformation of sound pressure level from the free field to the eardrum in the horizontal plane. *The Journal of the Acoustical Society of America, 56*(6), 1848–1861. https://doi.org/10.1121/1.1903522

Sivonen, V. P. (2011). Binaural directivity patterns for normal and aided human hearing. *Ear and Hearing, 32*(5), 674–677. https://doi.org/10.1097/AUD.0b013e31821a481f

Tyler, R. S., Noble, W., Dunn, C., & Witt, S. (2006). Some benefits and limitations of binaural cochlear implants and our ability to measure them: Algunos beneficios y limitaciones de los implantes cocleares binaurales y nuestra capacidad para medirlos. *International Journal of Audiology, 45*(Suppl. 1), 113–119. https://doi.org/10.1080/14992020600783095

van Hoesel, R. J. M., Jones, G. L., & Litovsky, R. Y. (2009). Interaural time-delay sensitivity in bilateral cochlear implant users: Effects of pulse rate, modulation rate, and place of stimulation. *JARO, 10*(4), 557–567. https://doi.org/10.1007/s10162-009-0175-x

van Hoesel, R. J. M., & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *The Journal of the Acoustical Society of America, 113*(3), 1617–1630. https://doi.org/10.1121/1.1539520

Yost, W. A., & Dye, R. H., Jr. (1988). Discrimination of interaural differences of level as a function of frequency. *The Journal of the Acoustical Society of America, 83*(5), 1846–1851. https://doi.org/10.1121/1.396520