



HAL
open science

Intensive Training of Spatial Hearing Promotes Auditory Abilities of Bilateral Cochlear Implant Adults: A Pilot Study

Aurélie Coudert, Grégoire Verdelet, Karen Reilly, Eric Truy, Valérie Gaveau

► **To cite this version:**

Aurélie Coudert, Grégoire Verdelet, Karen Reilly, Eric Truy, Valérie Gaveau. Intensive Training of Spatial Hearing Promotes Auditory Abilities of Bilateral Cochlear Implant Adults: A Pilot Study. *Ear & Hearing*, 2022, Publish Ahead of Print, 10.1097/AUD.0000000000001256 . hal-03874695

HAL Id: hal-03874695

<https://hal.science/hal-03874695>

Submitted on 28 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Intensive Training of Spatial Hearing Promotes Auditory Abilities of Bilateral Cochlear Implant Adults: A Pilot Study

Aurélie Coudert,^{1,2,3,4} Grégoire Verdelet,^{1,5} Karen T. Reilly,^{4,6} Eric Truy,^{1,2,3,4} and Valérie Gaveau^{1,4}

Objective: The aim of this study was to evaluate the feasibility of a virtual reality-based spatial hearing training protocol in bilateral cochlear implant (CI) users and to provide pilot data on the impact of this training on different qualities of hearing.

Design: Twelve bilateral CI adults aged between 19 and 69 followed an intensive 10-week rehabilitation program comprised eight virtual reality training sessions (two per week) interspersed with several evaluation sessions (2 weeks before training started, after four and eight training sessions, and 1 month after the end of training). During each 45-minute training session, participants localized a sound source whose position varied in azimuth and/or in elevation. At the start of each trial, CI users received no information about sound location, but after each response, feedback was given to enable error correction. Participants were divided into two groups: a multisensory feedback group (audiovisual spatial cue) and an unisensory group (visual spatial cue) who only received feedback in a wholly intact sensory modality. Training benefits were measured at each evaluation point using three tests: 3D sound localization in virtual reality, the French Matrix test, and the Speech, Spatial and other Qualities of Hearing questionnaire.

Results: The training was well accepted and all participants attended the whole rehabilitation program. Four training sessions spread across 2 weeks were insufficient to induce significant performance changes, whereas performance on all three tests improved after eight training sessions. Front-back confusions decreased from 32% to 14.1% ($p = 0.017$); speech recognition threshold score from 1.5 dB to -0.7 dB signal-tonoise ratio ($p = 0.029$) and eight CI users successfully achieved a negative signal-to-noise ratio. One month after the end of structured training, these performance improvements were still present, and quality of life was significantly improved for both self-reports of sound localization (from 5.3 to 6.7, $p = 0.015$) and speech understanding (from 5.2 to 5.9, $p = 0.048$).

Conclusions: This pilot study shows the feasibility and potential clinical relevance of this type of intervention involving a sensorial immersive environment and could pave the way for more systematic rehabilitation programs after cochlear implantation.

Key words: Cochlear implant, Rehabilitation, Spatial hearing, Virtual reality.

Abbreviations: 3D = three-dimensional; CI = cochlear implant; HMD = head-mounted display; ILD = interaural-level differences; ITD = interaural time differences; SSQ, Speech, Spatial, and Qualities of Hearing.

(Ear & Hearing 2022;XX:00–00)

¹ Integrative Multisensory Perception Action & Cognition Team—ImpAct, Lyon Neuroscience Research Center, Lyon, France; ²Department of Pediatric Otolaryngology—Head & Neck Surgery, Femme Mere Enfant Hospital, Hospices Civils de Lyon, Lyon, France; ³Department of Otolaryngology—Head & Neck Surgery, Edouard Herriot Hospital, Hospices Civils de Lyon, Lyon, France; ⁴University of Lyon 1, Lyon, France; ⁵Hospices Civils de Lyon, Neuro-immersion Platform, Lyon, France; and ⁶Trajectoires Team — Lyon Neuroscience Research Center, Lyon, France.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal's Web site (www.ear-hearing.com).

INTRODUCTION

Accurate analysis of the auditory scene is essential for orienting attention in space and necessary for daily behaviors (Kerber & Seeber 2012; Shinn-Cunningham et al. 2017). In complex auditory environments, which represent the majority of everyday situations, this involves extracting the signal of interest (i.e., essentially speech) from noise, so as to correctly localize each of them in relation to the other. This auditory localization is called spatial hearing, and allows the development of an accurate map of the environment where coordinates of each sound source are defined in three dimensional (3D), that is, in azimuth (front, back, left, or right space), in elevation (at, above, or below ear level), and in distance (near or far space). Spatial hearing relies on the brain's accurate interpretation of the binaural and monaural auditory cues reaching each ear. For example, a sound located to the right of a listener arrives first at the right ear and after a small delay at the left ear; this is called the interaural time difference (ITD). Added to this delay, the acoustic wave reaching the left ear is also attenuated by the head, leading to a reduction in its intensity. This creates a second binaural cue called the interaural-level difference (ILD). Moreover, the sound stimulus arriving at each ear is filtered by the pinna and head, and to a lesser extent, by the shoulders, depending on the incidental angle and the sound frequency (Angell & Fite 1901a,b). This filtering creates specific spectral cues called monaural cues. Binaural and monaural cues are complementary, and both are used to localize sounds in 3D. ITD and ILD are necessary for localizing in azimuth, whereas monaural cues are fundamental for elevation perception (e.g., Brungart 1999; Middlebrooks 2015).

It is important to note that spatial hearing is not a passive, but rather an active, multisensory process. Since the pioneering work of Wallach (1940), many studies have highlighted a major role of head movements in spatial hearing: they induce dynamic changes in binaural and monaural cues which increase the information content of the initial auditory inputs, and are thus extremely helpful in difficult hearing situations (e.g., resolving front-back confusions when ITD and ILD are near zero, see Wightman & Kistler 1999, Brimijoin et al. 2010, 2012). Under natural listening conditions, head movements are associated with eye movements, and sound source positions are encoded in retinocentric coordinates (Bulkin & Groh 2006; Pavani et al. 2008). This audiovisual association is essential in complex multisensory environments (Da Silva 1985; Loomis et al. 1998; Bolognini et al. 2005), and many studies have highlighted its benefits in terms of sound localization (Strelnikov et al. 2011) and speech discrimination (e.g., Middelweerd & Plomp 1987; MacLeod & Summerfield 1987; Schwartz et al. 2004).

With the aging population, deafness has become a major public health issue. In cases of severe to profound hearing loss, cochlear implants (CIs) can partially restore hearing. In recent decades, several factors have improved the outcomes of CI users, including the expansion of clinical indications for bilateral cochlear implantation. Indeed, many studies have shown that stereophony restoration with bilateral CI offers better hearing quality than unilateral CI (van Hoesel 2004; Smulders et al. 2016). Some authors have also recorded fewer sound localization errors in a quiet room in bilateral CI users compared to their unilateral peers (van Hoesel & Tyler 2003; Grantham et al. 2007). Moreover, technological advances have led to better sound signal processing, giving patients clearer

voice perception in difficult listening conditions (e.g., noisy environments with competing sounds).

Despite these advances, CIs do not fully restore sound and voice perception. Indeed, even if the quality of life of CI users is significantly improved after surgery (Mo et al. 2005), the Speech, Spatial and Qualities of hearing (SSQ; Gatehouse & Noble 2004) auditory quality of life questionnaire reveals that many users still experience difficulties in everyday situations. A majority of CI patients complain about misunderstanding speech in noisy environments (e.g., restaurant conversations) and poor sound localization (e.g., approaching vehicles) even after 2 years of bilateral hearing experience (e.g., Tyler et al. 2009; Perreau et al. 2014; Zhang et al. 2015; van Zon et al. 2017). As mentioned above, spatial hearing requires good extraction and interpretation of auditory cues, but ITD cues are largely distorted by CIs, which limits their potential usefulness. CI patients must therefore base their sound localization on ILD cues (Seeber et al. 2004; van Hoesel 2004), which partially explains difficulties in front-back judgements, even in bilateral CI users (Pastore et al. 2018; Fischer et al. 2020; Coudert et al. 2021). Furthermore, since the position of most CI microphones is largely outside the external ear, most amplitude and frequency modulations induced by the pinna are eliminated, leading to poor-quality monaural cues. Despite this, bilateral CI users are capable of perceiving small variations in ILD cues induced by head movements and to use these to resolve frontback confusions (Pastore et al. 2018; Coudert et al. 2021).

The existence of spatial hearing difficulties in patients with hearing loss has led some researchers to develop specific training protocols for sound localization. There is now growing evidence that training can promote unisensory learning (Shams et al. 2011), even in a deficient sensory modality (Isaiah & Hartley 2015). These types of training protocols have been explored with a small number of patients with various clinical profiles: single-sided deafness (Firszt et al. 2015), hearing aid users (Kuk et al. 2014), and CI users (Tyler et al. 2010), but to our knowledge, no clinical centers currently offer a standardized spatial hearing rehabilitation protocol.

During the early stages after surgery, auditory rehabilitation focuses on sound detection, with this focus shifting to speech discrimination in quiet and noisy situations several months after surgery. Sound localization, however, is rarely included in routine rehabilitation protocols. Experimental spatial hearing rehabilitation protocols could provide a good starting point for developing training protocols adapted to clinical practice. Indeed, experimental approaches that exploit multisensory interactions to promote learning suggest that rehabilitation of spatial hearing and spatial attention in CI patients is feasible (Tyler et al. 2010; Nawaz et al. 2014). There are, however, several methodological challenges that must be overcome in order to sufficiently control the proposed training sessions. First, since spatial hearing is 3D, it is essential to work in the whole space surrounding the patient, and not only in front space with azimuth (as is commonly the case, e.g., Tyler et al. 2010; Firszt et al. 2015). To do this, using a traditional setup would require many loudspeakers in front and back space at different elevations, which would make the system very bulky and difficult to use in everyday clinical practice. Furthermore, since spatial hearing is a multisensory process, strongly based on visual information

(particularly in hearing impaired patients), control of the visual environment is critical. In most training protocols, participants have visual information about the actual positions of loudspeakers around them, which could bias their sound localization responses (Tyler et al. 2010; Kuk et al. 2014; Nawaz et al. 2014; Firszt et al. 2015).

To overcome the methodological constraints related to 3D space, and the need to take into account the multisensory dimension of auditory perception (i.e., role of head movements and visual information), we developed a spatial hearing training protocol for bilateral CI adults based on a validated virtual reality system combined with real-time 3D motion tracking (Verdelet et al. 2019; Valzolgher et al. 2020a,b; Coudert et al. 2021). This system allows spatial hearing to be studied with (1) very limited constraints on sound source locations and responses (i.e., the whole auditory space can be sampled using a single loudspeaker); (2) continuous recording of head movements; and (3) control of all available visual cues during the experiment.

The pilot study presented here involved bilateral CI users who followed a month-long training protocol in an environment in which visual and auditory stimuli were controlled. We used a multisensory training protocol in the near-field where real sounds and visual information were delivered simultaneously while participants were encouraged to explore their environment using head movements to enrich the auditory information.

The primary objective of this pilot study was to evaluate the feasibility of an intensive spatial hearing training protocol and to examine the impact of this program on sound localization performance, speech comprehension in noise, and quality of life. Based on the multisensory nature of our training, and previous studies demonstrating the benefits of head movements on sound localization (Pastore et al. 2018, Coudert et al. 2021), we predicted an improvement in sound localization after the training protocol. We expected the rehabilitation protocol to have a smaller impact on speech discrimination since the training focused on sound localization and not on word recognition.

The secondary objective of this study was to compare the effect of two different types of spatial feedback: visual (unisensory) or audiovisual (multisensory). To do this, we divided our population into two groups: the first received only visual feedback in the form of a spatial cue of the sound source, as this sensory modality was intact in all CI users. The second group received audiovisual feedback. This group received the same spatial cue as the first group plus audio feedback. Based on evidence suggesting that multisensory training is better than unisensory (Strelnikov et al. 2011), we hypothesized that training-related improvement would be greater in the group that received audiovisual feedback.

MATERIALS AND METHODS

The study was conducted in accordance with the Declaration of Helsinki. The protocol was approved by the French ethics committee Sud-Ouest et Outre-mer III (CPP 2019-A01335-52) and recorded in clinicaltrials.gov (NCT04078763). All participants gave informed consent for inclusion before participating in the study.

Participants

Twelve bilateral CI adults aged between 19 and 69 (mean age \pm SD: 41.4 ± 14.7 years) were recruited from a referral center for cochlear implantation. Inclusion criteria included age at testing between 18 and 75 years old, a minimum of 1 year of bilateral experience (to avoid large variations in CI processor settings), normal vision (with or without correction), and no areflexia (to avoid balance disorders when wearing the virtual reality head-mounted display). Bilateral experience ranged from 18 to 107 months (mean 67.3 ± 32.4 months). Since our long-term goal is to propose this type of rehabilitation program to all interested CI users, we did not target a specific patient profile. As such, no inclusion criteria regarding CI processor brand or settings were applied. The bilateral CI participants were all daily users of their cochlear implants and had excellent monosyllabic word recognition performance at 50 dB HL (mean with left CI: $87.5 \pm 9.3\%$, and with right CI: $91.8 \pm 8.8\%$).

Additional information about patient demographics and device settings (i.e., internal parts of CI, sound processors, programming parameters, sound coding strategies) are summarized in Table 1. Importantly, participants wore their own cochlear implant processors and no parameter adjustments were made before testing. During the sound localization test no adjustments of microphone position were required, as the head-mounted display did not press on the processor, leaving the microphones unimpeded, whether behind or off the ear. All participants had a fitting with an audiologist less than 1 month before inclusion to check their processors were functioning properly. Since neither the training nor the evaluation conditions included any background noise, those participants with directional microphones used omnidirectional settings.

Experimental Protocol

The sound localization rehabilitation protocol was conducted across 10 weeks and consisted of eight training sessions interspersed with evaluation sessions (Fig. 1). Since our objective was to test the feasibility of this type of program, the sessions followed the same rhythm as speech rehabilitation sessions following surgery (2 sessions/week). The evaluation sessions allowed us to estimate the impact of training on sound localization and speech perception abilities and were performed at different times throughout the rehabilitation protocol: before (evaluations E0 and E1), after four (E2) and eight training sessions (E3), and 1 month after the eighth training session (E4).

All sessions (training and evaluation) took place in a reverberant room ($3.6 \text{ m} \times 3.9 \text{ m} \times 2.7 \text{ m}$, reverberation time RT60: 0.32 s) within a hospital. All participants followed the same training protocol regardless of their level of performance at inclusion.

Material for Delivery of Auditory and Visual Stimuli

The idea behind the training sessions was to promote spatial hearing rehabilitation by providing participants with trial-by-trial feedback in the form of real sounds and visual information congruent with sound sources, or with visual information only. Visual information was delivered using a virtual reality (VR) apparatus and sounds were delivered using a single loudspeaker (mini speaker

model JBL GO Portable from HARMAN International Industries, Northridge, California USA; $68.3 \times 82.7 \times 30.8$ mm, Output Power 3.0 W; frequency response: 180 Hz–20 kHz) that was moved in space by the experimenter. For each sound localization trial, the experimenter stood next to the participant (who was sitting on a rotating chair) and silently moved the loudspeaker by hand to the desired sound position guided by a weak echo radar signal for elevation placement and visual feedback on a computer screen for direction placement. Further details of the experimental setup are available in Coudert et al. (2021).

Visual stimuli were delivered using a VR system validated in previous studies for use in behavioral research (Verdelet et al. 2019; Valzolgher et al. 2020a,b; Coudert et al. 2021). Our VR system consisted of a head-mounted display (HMD; HTC VIVE System, resolution: 1080×1200 px, Field Of View (FOV): 110° , Refresh rate: 90 Hz) with integrated eye-tracking technology (SensoriMotoric Instruments, Berlin, Germany; www.smivision.com; 60 Hz frequency and 0.5° spatial precision). The main advantages of using the virtual reality apparatus for training sessions were (1) to control the visual environment (i.e., avoid access to real visual cues that could help sound localization) and (2) to give trial-by-trial visual feedback in VR on sound localization performance. It also enabled us to use a handheld VIVE controller that participants positioned in space to indicate sound localization (during evaluation sessions), and to track the 3D position of the loudspeaker using a second VIVE device (see Coudert et al. 2021 for more details). All training and evaluation sessions were performed in a quiet environment (i.e., background noise at 33.7 dB SPL). Tested positions were predetermined and controlled using custom-made software (Unity, 2017.4.10f1).

To avoid any transfer of sensorimotor learning from the training to the evaluation sessions, we used different stimuli and responses. During training, stimuli were spoken words (see details below) and participants responded by orienting their head to the direction of the sound. During evaluation, stimuli were 3-second white noise bursts and participants indicated sound sources using a handheld pointer. The spatial positions of sound sources also differed between training and evaluation sessions, as did the visual environments—a visual, structured immersive display during training and a uniform, gray environment during evaluation.

Training Sessions

The rehabilitation protocol consisted of eight 45-minute training sessions (T1 to T8) distributed across 4 weeks (Fig. 1). During training, participants wore an HMD and a sound was played for 15 seconds through a loudspeaker positioned at a predetermined position in their near-field space (i.e., less than 1 m from the head).

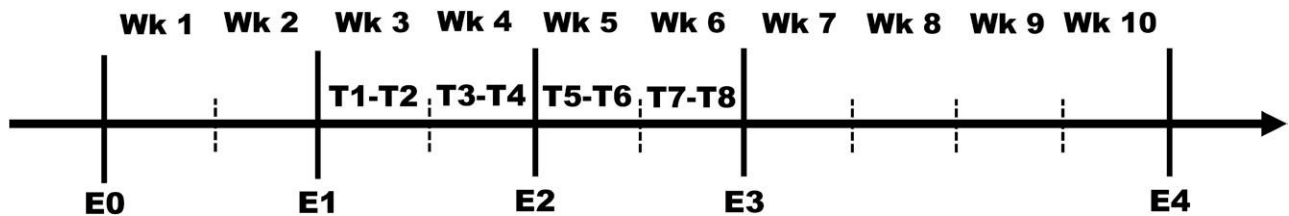


Fig. 1. Flowchart of the rehabilitation training protocol conducted over 10 weeks (Wk 1 to Wk 10), with five evaluation sessions (E0 to E4) distributed across the 10 weeks and 8 training sessions (T1 to T8) distributed across 4 weeks (weeks 3 to 6).

The auditory stimulus delivered during the training sessions was a Lafon list (1964) of three disyllabic words repeated throughout the 15-second trial. Six different lists were used in each training session so that participants did not habituate to the words. During training the participant's task was to localize the source of the sound by orienting their head toward the sound's perceived location. The loudspeaker's distance from the center of the head remained fixed at 90 cm, and depending on the training session, its position varied in azimuth and/or elevation (see Fig. 2).

At the beginning of each session, participants were informed of the portion of space that would be trained. Tested stimulus locations and session difficulty were varied across training sessions to maintain motivation and to challenge participants throughout the training protocol. The level of difficulty was increased across the eight training sessions by training first in only one area of space (front or back) and altering only one dimension (azimuth or elevation), while in the final two sessions, participants were tested in both front and back space with alterations in either azimuth (session 7) or elevation (session 8). The number of trials per session and session duration was fixed (72 trials; 45 minutes) since our goal was to test the feasibility of using this type of standardized spatial hearing rehabilitation protocol in everyday clinical practice.

When participants entered the training room, they saw the room's layout and furniture but had no visual information about loudspeaker position. They were introduced to the apparatus and the training protocol by watching a short video. They were then invited to wear the HMD and were immersed in virtual reality where the room's layout and furniture were represented accurately. Participants underwent a 5-minute familiarization session during which they had to localize sounds by orienting their heads to the perceived sound location. Actual sound locations were 10 different possible positions anywhere around them, similar to the T7/T8 training sessions (see Fig. 2E).

Each training trial was divided into two phases: a sound localization phase and a feedback phase. During the sound localization phase, the experimenter silently moved the loudspeaker to a predetermined position (unknown to the participant) and a sound was emitted for 15 seconds. The participant's task was to identify the direction from which the sound was emitted by orienting their head to the perceived sound source location. Scanning the environment with head movements was encouraged during this phase. They then validated their head position (which was represented by a white cross in the HMD, see videos in Supplemental Digital Content 1, <http://links.lww.com/EANDH/B30>, and Supplemental Digital Content 2, <http://links.lww.com/EANDH/B31>) by clicking a button on a handheld VIVE controller. This initiated the feedback phase during which they received immediate feedback on their performance.

If the participant did not respond during the 15-second trial duration feedback appeared automatically at the end of the trial. Half of the participants (group A) received both visual and auditory feedback, in the form of an image of the loudspeaker in virtual reality in a position spatially congruent with its actual position in space, while the sound stimulus continued playing from the loudspeaker's unchanged position. The other half (group B) received exactly the same visual feedback as group A (indicating the spatial position of the sound source) but without any sound (i.e., no spatial auditory information). The added value of providing multisensory information was investigated by comparing audiovisual feedback with visual-only feedback because the visual system dominates over the auditory system in perception of the environment (Witten & Knudsen 2005).

Both groups saw the image of the loudspeaker plus an arrow at the position at which they indicated that they had perceived the sound (with their head orientation). The distance between the actual and perceived locations informed participants about the size of the error. To reinforce this feedback, this information was also communicated via arrow thickness—thicker arrow = greater error. Thus, they received direct and informative visual feedback about their localization error. Both groups were instructed to use the feedback to correct their sound localization response by reorienting their head such that the arrow overlapped with the image of the loudspeaker, and to maintain this position for 5 seconds, after which time the next trial began. Videos of an audiovisual feedback trial and a visual-only feedback trial are available in Supplemental Digital Content 1, <http://links.lww.com/EANDH/B30>, and Supplemental Digital Content 2, <http://links.lww.com/EANDH/B31>, respectively.

Participants were allocated to one of the two groups in order of inclusion (first patient to visuospatial, second to visual, third to visuospatial, and so on).

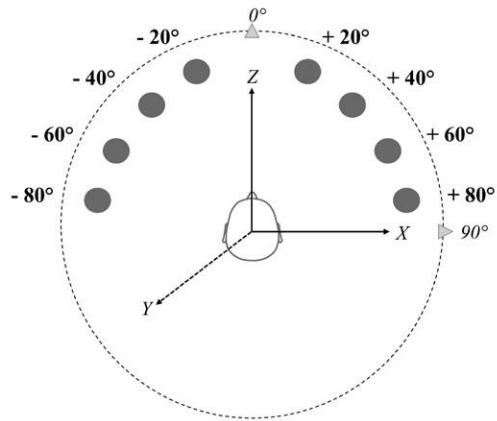
Evaluation Sessions

To control for the possibility that the evaluation session itself led to performance improvements, two evaluation sessions (E0 and E1) were performed before the start of training. The two evaluations were separated by 2 weeks and performance in the two sessions was compared (see Results). To evaluate the immediate benefits of the training sessions we performed an evaluation session in the middle of the training sessions (E2) and after the last training session (E3). A final evaluation (E4) was performed 1 month after the end of the last training session to evaluate whether any observed effects of training persisted over time. Each evaluation session

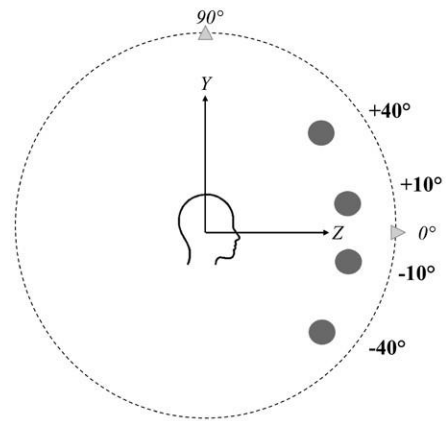
consisted of a quality of life questionnaire and two objective tests: a 3D sound localization test and a speech comprehension test in noise.

Quality of Life Questionnaire • A French language short form of the Speech, Spatial and Qualities of hearing scale (called the SSQ questionnaire) was selected among the large panel of quality of life questionnaires that exist for CI patients (Moulin et al 2019). We chose this questionnaire because it

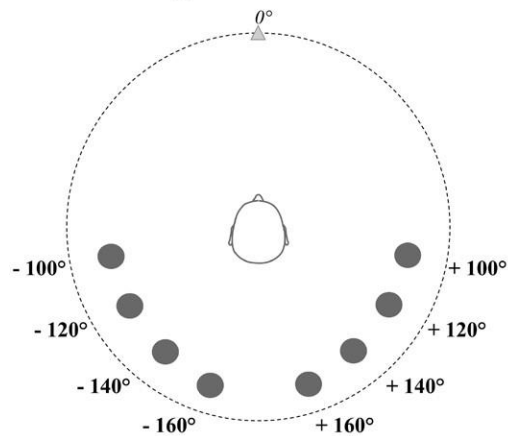
A Training 1



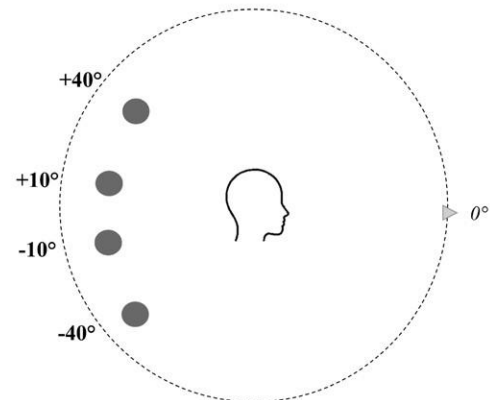
B Training 2 and 3



C Training 4



D Training 5 and 6



E Training 7 and 8

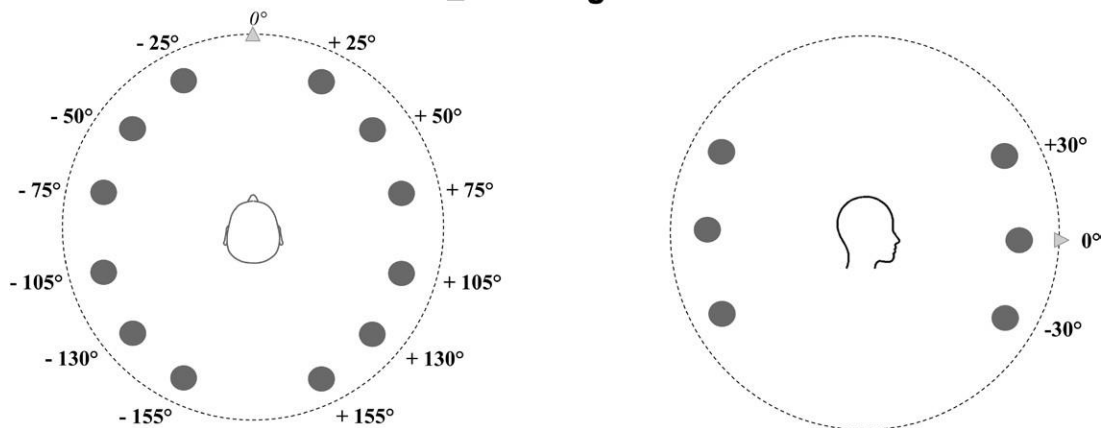


Fig. 2. Configuration of loudspeaker positions during the 8 training sessions (T1 to T8). A and B, Training sessions with sound sources placed in front space; (C) and (D) in back space; (E) in both front and back space with different azimuthal and elevation sound sources. Three axes were defined according to the reference-frame (i.e.

participant head-centered): X, azimuth; Y, elevation; and Z, distance. Note that for trainings T1 and T4 elevation remained fixed at ear level; for trainings T2 and T3 azimuthal stimulations were fixed at -45° and $+45^\circ$; and for trainings T5 and T6 azimuthal stimulations were fixed at -135° and $+135^\circ$.

contains specific questions on sound localization and speech comprehension in noise. The short form of the SSQ has 15 items, divided into three subscales: speech perception; spatial hearing; and other qualities of hearing. Each question is scored from 0 (not at all) to 10 (perfectly) and higher scores represent greater perceived ability in everyday situations. The questionnaire was administered during a face-to-face interview with an ENT doctor.

3D Sound Localization Task • The 3D sound localization task was performed in the same room as the training sessions. During this task, participants wore the HMD but received no visual information about their environment (the screen was gray). Participants were told that sounds would be played from various positions around their body and that their task was to place a handheld pointer at the exact position at which they perceived the sound. They had no risk of colliding with the loudspeaker which was rapidly removed after the sound ended. They were also told they could reach each sound source without leaving their chair. Unbeknownst to them, eight predetermined sound positions at a constant distance of 55 cm from the center of the head (within reaching space) were used. The loudspeaker could be located at $+30^\circ$, $+70^\circ$, $+120^\circ$, $+160^\circ$, -30° , -70° , -120° , and -160° in azimuth with respect to the participant's head looking straight ahead (see Fig. 3, positive values indicating right space and negative values left space). Two elevations were evaluated; $+25^\circ$ and -25° . Each position was repeated eight times, for a total of 64 trials. It is important to note that all of these positions differed from those tested in the training sessions. We recorded responses in 3D, considering azimuth, elevation and distance. During this task, the sound stimulus was 3 seconds of white noise, amplitude-modulated at 2.5 Hz (modulation depth at 80%) and delivered at 73 dB SPL.

Speech Comprehension in Noise Task • Speech comprehension in noise was evaluated using the French Matrix Test (Jansen et al. 2012). This test was performed in the sound field in a calibrated room ($4.1\text{ m} \times 2.8\text{ m} \times 2.7\text{ m}$, reverberation time RT60: 0.39 seconds). Two loudspeakers were placed at ear level, at -45° and $+45^\circ$ azimuth angle, and at 1 m from the participant's head. One loudspeaker delivered speech sentences (all based on the same syntactic structure: name-verb-numberobject-color), while the other delivered noise (i.e., a stationary long-term average speech spectrum made by Oldenburg Measurement Application, www.hoertech.de). Sentences were always presented on the side of the better ear. The test consisted of three blocks: a training block of 10 sentences (results were not taken into account in the statistical analyses), followed by two 20-sentence blocks. For each block, we used an adaptive procedure (see Jansen et al. 2012 for details) where noise was fixed at 60 dB SPL and speech level varied to obtain a signal-to-noise ratio (SNR) with 50% correct word recognition: used to determine the speech recognition threshold (SRT). The lower the SRT, the better the ability to discriminate speech in the presence of high background noise.

Data Analysis

To assess the impact of the training sessions on auditory performance and perceived quality of life, for each participant we analyzed performance during each of the four evaluation sessions (E0 to E4). For the sound localization test, we calculated two separate values: (1) the percentage of trials on which they made front-back confusions in azimuth and (2) the percentage of trials on which they confused up and down sound sources in elevation. Performance at chance level was defined as a score of $>50\%$ of confusions as in this task the participant's response was either in the correct hemi-field or not. For the matrix test, we averaged the two SRT values (one per 20-sentence block) to obtain a mean SRT score. For the SSQ questionnaire, for each patient we calculated the mean SSQ score per subscale (speech perception, spatial hearing, and other qualities of hearing), as well as the mean total SSQ score across the three subscales. There are no norms for changes across time on SSQ scores in CI patients. Thus, to evaluate the individual benefits of the training on quality of life, we used a categorization system based on Noble's work which assessed the benefits of hearing aids on SSQ scores (Noble & Gatehouse 2006). A score increase of one to two points corresponds to a moderate effect, while an increase of more than two points corresponds to a large benefit.

Performance improvement was defined as a decrease in sound localization confusions, and/or a decrease in SRT score, and/or an increase in total SSQ score during the rehabilitation protocol (at E2), immediately after at the end of rehabilitation (at E3), and/or one month later (at E4).

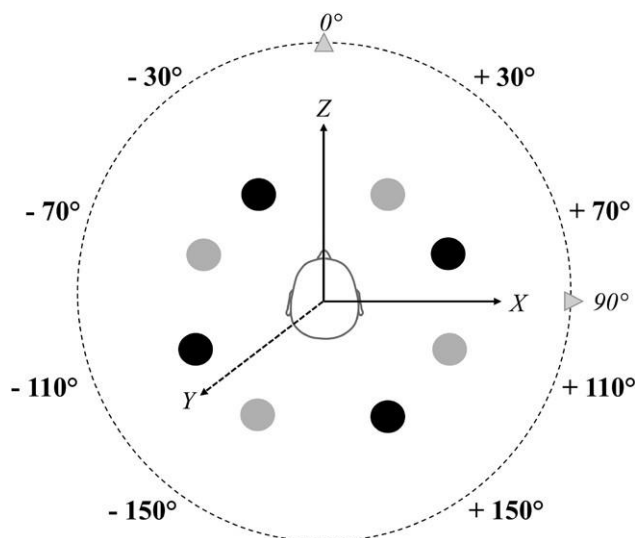


Fig. 3. Setup during sound localization evaluation sessions. A total of eight positions were tested. Black and gray circles indicate two target elevations in the near-field, at $+25^\circ$ and -25° , respectively. Distance was fixed at 55 cm. Three axes were defined according to the reference-frame (i.e. patient head-centered): X, azimuth; Y, elevation; and Z, distance.

Statistical analyses and data visualization were performed using the R-studio environment (www.rstudio.com). Since Shapiro–Wilk tests revealed that not all data were normally distributed, we used nonparametric tests for all statistical analyses. A Friedman test followed by a Wilcoxon test with Bonferroni correction was performed to compare results between E1 and E2, E1 and E3, E1 and E4.

RESULTS

Training Feasibility

All 12 participants attended all evaluation and 45-minute training sessions, demonstrating the feasibility of implementing this type of rehabilitation protocol in everyday clinical practice.

Auditory Performance Before Training

To investigate the possibility that participation in the evaluation sessions had a beneficial effect on auditory performance, we compared performance at E0 and E1. The delay between these two sessions was 2 weeks and both were conducted before training began (see Fig. 1). Three separate Wilcoxon signedrank tests indicated no significant performance differences between E0 and E1 on any of the tests: sound localization test ($V = 29.5$, $p = 0.121$), matrix test ($V = 56$, $p = 0.051$), and SSQ questionnaire ($V = 40$, $p = 0.563$). These results suggest that the content of the evaluation sessions did not induce any learning. For all subsequent analyses, baseline performance was defined as performance at E1, as this session was the closest in time to the training sessions.

Figure 4 shows data from all 12 participants, for all eight sound sources for each evaluation session. Each colored symbol represents the average hand pointing position for a given participant for a given sound source. The left column shows performance as a function of front-back (A. Bird's eye view), and the right column shows up-down performance (B. Lateral view). The top two panels of Figure 4 show average hand pointing positions during E1. These figures reveal that prior to training bilateral CI users mainly pointed along the interaural axis (around their cochlear implants) and in front space. This resulted in a large number of front-back confusions (median of 32.0%, E1, Fig. 5A). For elevation discrimination, the median up-down confusion was 49.2%, revealing that they were near-chance level for discriminating sounds delivered in the upper or lower part of space (E1, Fig. 5B). Figure 6 shows median SRT scores on the matrix test. Before training the median SRT was 1.5 dB SNR and between-participant variability was very high (range -4.7 to 6.8). The median for the total SSQ score was 5.9 of 10.

Auditory Performance Mid-training and Immediately After Training

Evaluation sessions E2 (after 4 weeks of training) and E3 (after 8 weeks of training) allowed us to assess the immediate benefits of the rehabilitation protocol, whereas E4 allowed us to establish whether any benefits were still present 4 weeks after the end of the protocol. Performance at each of these three time points was compared to baseline performance obtained during the pretraining assessment E1.

The four middle panels of Figure 4 show pointing responses half-way through (E2) and immediately after the end of the training (E3), while the bottom panels show responses 4 weeks after the end of training (E4). The left-hand panels show that, compared with E1, participants' performance for back sound sources progressively improved, with the median percentage of frontback confusions decreasing from 25.8% during E2 to 14.8% during E3 and stabilizing at 14.1% during E4 (Fig. 5A). The group-level improvement at

E3 was driven by eight of 12 bilateral CI users, all of whom had fewer than 25% of confusions, although it is important to note that at E4 nine of 12 participants also had fewer than 25% of confusions. A Friedman test including data from all four sessions found a significant decrease in front-back confusions between sessions ($\chi^2(3) = 10.9, p = 0.012$, effect size $W = 0.31$). Pairwise comparisons (Wilcoxon test with Bonferroni adjustment) showed that this improvement was significant between E1 and E3 ($p = 0.017$) and between E1 and E4 ($p = 0.023$) but not between E1 and E2 ($p = 0.169$).

The right-hand panels in the middle of Figure 4 show that even after 8 weeks of training up-down discrimination remained difficult, and the bottom right-hand panel shows that this difficulty was still present 4 weeks after the end of the protocol. This can also be seen in Figure 5B, which shows only a slight decrease in up-down confusions across the four evaluation sessions (from 46.1% during E2 to 42.2% during E4). Interestingly, even if the median number of up-down confusions at E4 was similar to that at E3, several participants continued to improve between E3 and E4, and the improvements in two participants (A01 and B03) put them below the chance level (see Table 2). A Friedman test revealed no significant differences between the four evaluation sessions ($\chi^2(3) = 6.54, p = 0.088$).

Together, the data in Figures 4 and 5 show that training induced a moderate improvement in front-back confusions, and that this improvement was still present 4 weeks after the end of training. In contrast, although up-down confusions decreased slightly with training, this decrease was not significant, and confusions remained at the chance level at all evaluation sessions.

One participant (A05) was never able to repeat the sentences in the Matrix test even after 8 weeks of training. Thus, Figure 6 shows SRT scores from all four evaluation sessions for 11 of 12 participants. The median SRT score decreased from 1.5 dB at E1 to 0.2 dB SNR at E2 and -0.5 dB at E3 then stabilized at -0.7 dB SNR at E4. In addition to clearly showing the gradual decrease in SRT scores across evaluation sessions, Figure 6

also shows a large decrease in between-participant variability across sessions. SRT scores ranged from 6.8 to -4.7 dB SNR at E1 and from 1.3 to -4.8 dB SNR at E4. SRT scores improved between E1 and E2 for eight of 11 participants and between E2 and E3 for 10 of 11 participants. Some participants continued to improve between E3 and E4, although these improvements were generally small and were observed in only five of 11 participants (see Table 3 for individual data). A Friedman test revealed a significant decrease in SRT scores across all 4 evaluation sessions ($\chi^2(3) = 12.6, p = 0.006$, effect size $W = 0.39$). Pairwise comparisons (Wilcoxon test with Bonferroni adjustment) showed that this improvement was significant between E1 and E3 ($p = 0.029$), and between E1 and E4 ($p = 0.018$) but not between E1 and E2 ($p = 0.194$). It is important to note that at the end of training, eight bilateral CI users were able to discriminate sentences with a *negative* SNR, meaning that they were able to correctly repeat sentences with a higher noise than the speech level. In summary, the data from the Matrix test show that SRT improved moderately with training and that this improvement was still present 4 weeks after the end of training.

The median total SSQ score was stable from E1 to E2 (5.9 and 5.8), then increased to 6.7 at E3, and stabilized at 6.7 at E4. A Friedman test revealed a significant increase in median total SSQ score across all four evaluation sessions ($\chi^2(3) = 22.3, p < 0.001$, effect size $W = 0.62$). Pairwise comparisons (Wilcoxon test with Bonferroni adjustment) showed that this improvement was significant between E1 and E3 ($p = 0.003$), and between E1 and E4 ($p = 0.015$) but not between E1 and E2 ($p = 0.205$). To investigate whether this improvement concerned all three subscales of the SSQ questionnaire, we compared the scores on each subscale across all four evaluation sessions (Fig. 7). The median scores for speech perception and spatial hearing at E1 were similar (5.2 and 5.3), whereas the other qualities of hearing score was higher (7.9). Speech perception scores progressively improved from 5.5 at E2 to 5.6 at E3 finishing at 5.9 at E4. This improvement was significant (Friedman test, $\chi^2(3) = 8.78, p = 0.032$, effect size $W = 0.24$), but only between E1 and E4 ($p = 0.048$). Spatial hearing scores steadily increased over the four evaluations; 5.8, 6.6, and 6.7 from E2 to E4. This improvement was significant (Friedman test, $\chi^2(3) = 14.7, p = 0.002$, effect size $W = 0.41$) between E1 and E3 ($p = 0.015$), and E1 and E4 ($p = 0.032$). Other qualities of hearing remained constant around 8 throughout the evaluations ($\chi^2(3) = 6.85, p = 0.077$).

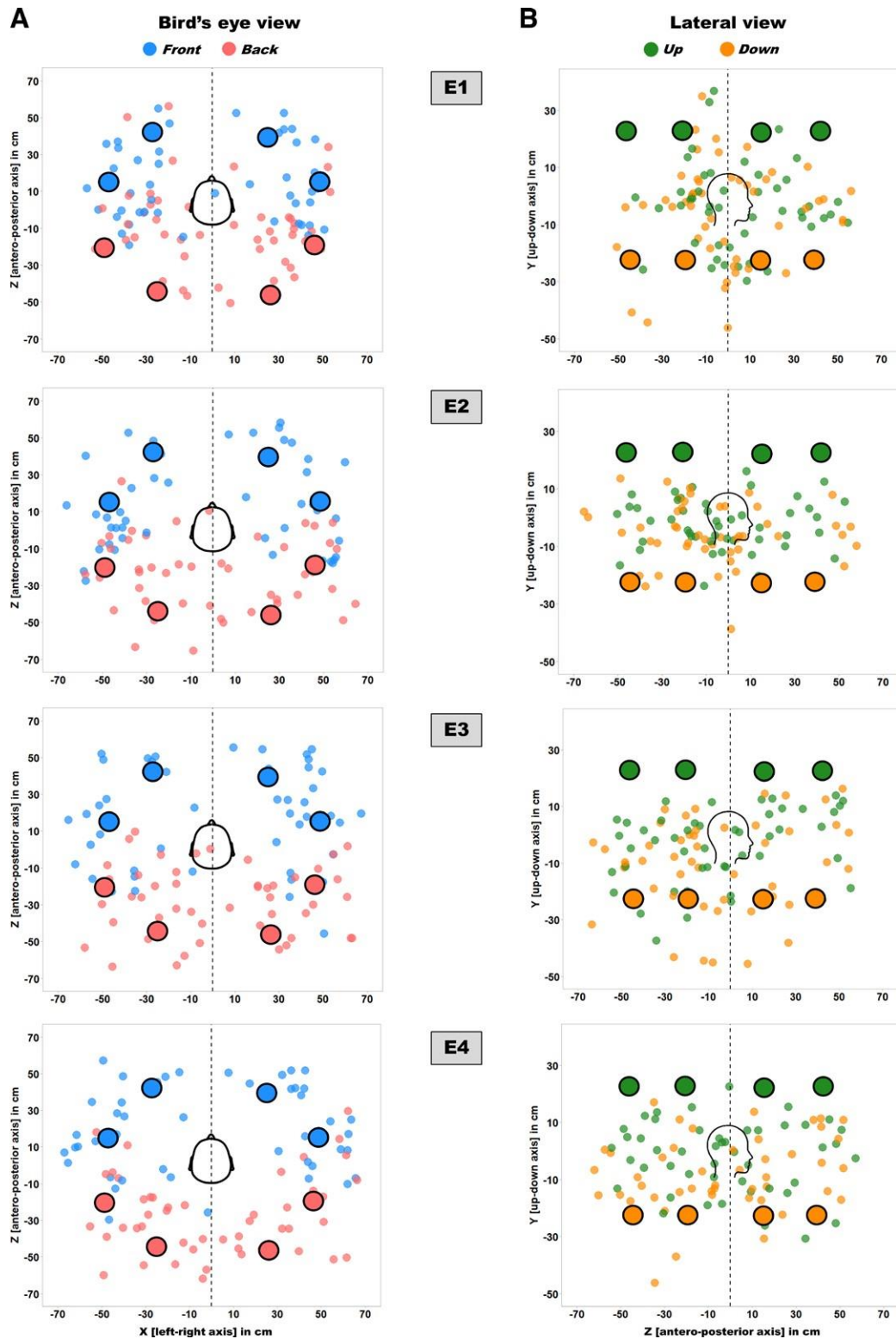


Fig. 4. Sound localization performance of all 12 bilateral cochlear implant users during each evaluation session: before training (E1), after 4 (E2), and 8 (E3) training sessions, and 1 month after the end of the training sessions (E4). Black outlined symbols represent the sound sources and colored symbols represent the mean response for each participant per target. A, Bird's eye view showing sound localization indicated by hand position (see methods) as a function of front-back sound sources (blue and red circles for front and back sound sources). B, Lateral view showing sound localization based on hand position as a function of up stimulation (green circles) and down stimulation (yellow circles).

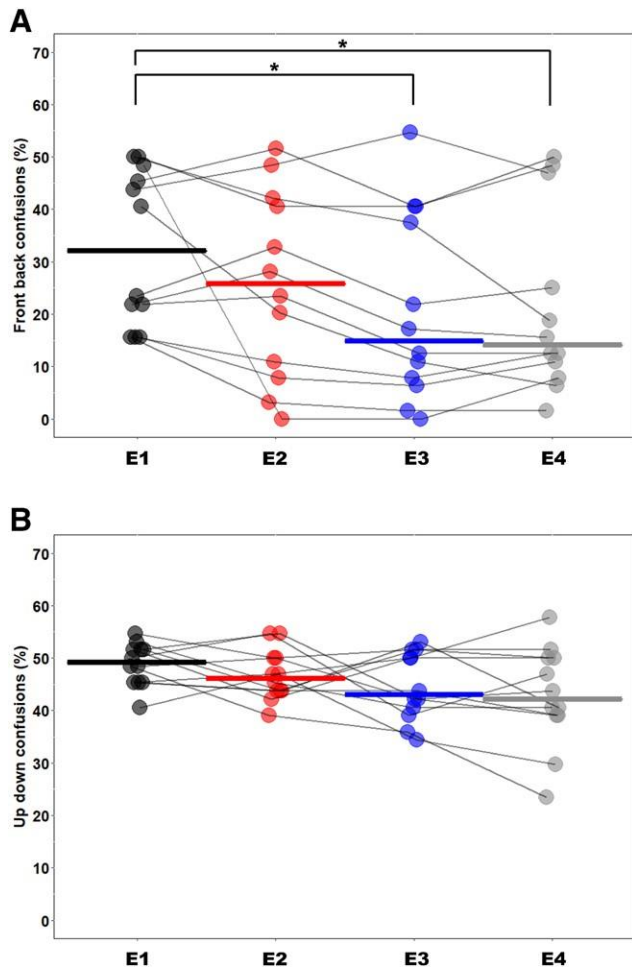


Fig. 5. Sound localization performance during evaluation sessions E1 to E4. Thick lines represent the median percentage of confusions for all 12 participants: (A) front-back confusions in azimuth and (B) up-down confusions in elevation. Colored dots correspond to the percentage of confusions for each participant per evaluation session (black, red, blue and gray for E1, E2, E3, and E4, respectively). Asterisks indicate significant differences (Wilcoxon test, $*p < 0.05$).

It is interesting to note that the individual participant data in Table 4 reveal substantial variability in SSQ scores. Some participants reported no difficulties in their daily lives on any of the subscales (e.g., B03, B05), whereas others still reported sound localization problems after the training (e.g., A02, A06, B01).

Overall, the questionnaire revealed that self-reported quality of life improved substantially with training and that this improvement was still present 4 weeks after the end of training. Four participants reported significant benefits for the “speech perception subscale”: two had moderate benefits (i.e., a score increase between 1 and 2 points) and two had large benefits (i.e., a score increase between 2 and 4 points). Eight participants reported benefits on the “spatial hearing subscale”: three had moderate benefits and five large benefits.

Auditory Performance with Audiovisual or Visual-only Feedback During Training

Since our main goal was to examine the feasibility of this novel rehabilitation protocol and to obtain preliminary data concerning its potential benefits, the data presented above include

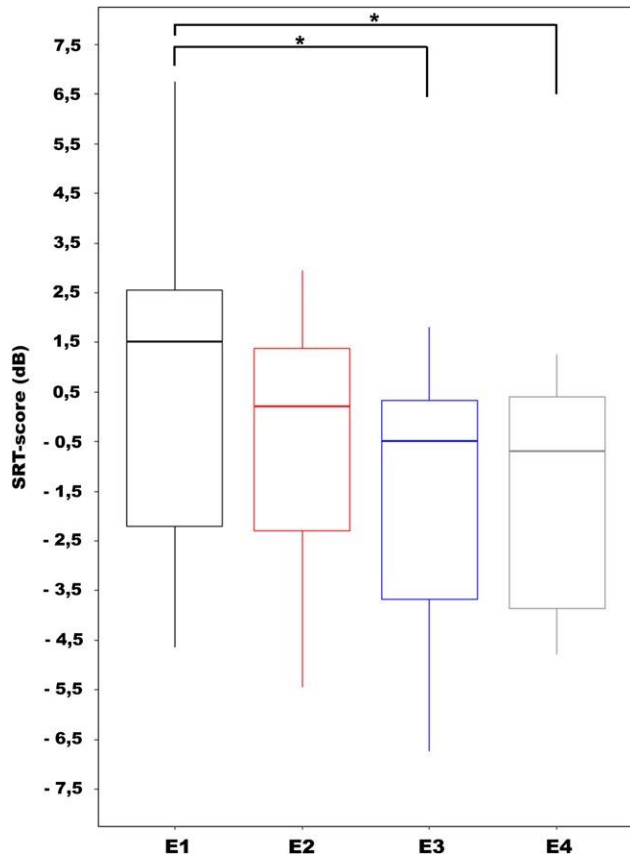


Fig. 6. SRT scores during evaluation sessions E1 to E4. Lines are medians, box limits 25th–75th percentiles and error bars 95% confidence limits ($n = 11$ for each session). Asterisks indicate significant differences (Wilcoxon test, $*p < 0.05$). SRT, speech recognition threshold.

all participants, regardless of the type of feedback they received. Despite the small number of participants in each group ($n = 6$), we were also interested to know whether there was an added benefit of training for those participants who received feedback in two sensory modalities. To assess this, we first ensured that performance in the two groups was similar before training. A Mann–Whitney test revealed no significant differences between groups for the sound localization test ($V = 13$, $p = 0.466$), the matrix test ($V = 22$, $p = 0.247$), or the SSQ questionnaire ($V = 17$, $p = 0.937$). We then compared performance in the two groups at E3 at the end of the eight training sessions. No significant differences on any of the evaluation tests emerged between groups at E3 (Mann–Whitney tests): the sound localization test ($V = 21$, $p = 0.699$), the matrix test ($V = 20$, $p = 0.429$), and the SSQ questionnaire ($V = 15.5$, $p = 0.748$). For these two small groups, there was no difference between visual feedback alone and receiving both visual and auditory feedback.

DISCUSSION

In this pilot study, 12 bilateral CI adults were included in an intensive spatial hearing rehabilitation protocol. The study lasted 10 weeks as it included evaluation sessions 2 weeks before and 4 weeks after eight biweekly training sessions. Development of the training sessions was based on three observations for which there is increasing clinical and scientific support: (1) spatial hearing is a multisensory process, (2)

TABLE 2. Individual patient data for sound localization errors in terms of front-back and up-down confusions across evaluation sessions (E1 to E4)

Group	Id	Front-back Confusions (%)				Up-down Confusions (%)			
		E1	E2	E3	E4	E1	E2	E3	E4
A	A01	15.6	3.1	1.6	1.6	48.4	39.1	35.9	23.4
	A02	21.9	28.1	17.2	15.6	50.0	54.7	39.1	46.9
	A03	43.8	48.4	54.7	46.9	51.6	42.2	50.0	50.0
	A04	50.0	42.2	37.5	18.8	45.3	46.9	50.0	57.8
	A05	15.6	10.9	7.8	12.5	54.7	50.0	51.6	51.6

	A06	23.4	32.8	21.9	25.0	53.1	45.3	40.6	40.6
	Mean all	28.4	27.6	23.5	20.1	50.5	46.4	44.5	45.1
B	B01	45.3	51.6	40.5	48.4	48.4	50.0	42.2	43.8
	B02	50.0	40.6	40.6	50.0	45.3	43.8	51.6	50.0
	B03	15.6	7.8	6.3	10.9	40.6	46.9	34.4	29.7
	B04	40.6	20.3	10.9	6.3	51.6	54.7	42.2	39.1
	B05	21.9	23.4	12.5	12.5	45.3	43.8	43.8	39.1
	B06	48.4	0.0	0.0	7.8	51.6	43.8	53.1	40.6
	Mean all	37.0	24.0	18.5	22.7	47.1	47.2	44.6	40.4

Group A, audiovisual feedback; Group B, visual feedback; Id, patient identification.

exploration of the auditory environment with head movements improves sound localization, and (3) training-induced learning is better when several sensory modalities are combined. Our study demonstrates the feasibility of using a standardized training protocol with bilateral CI patients, who are known to have large clinical variability.

In the present study, we evaluated the benefits of eight 45-minute training sessions spread over 4 weeks by examining performance on two auditory tests and a questionnaire before, during and after the training sessions. We observed that four training sessions spread across 2 weeks were insufficient to induce significant performance changes, whereas performance on both auditory tests and the questionnaire improved after eight training sessions spread across 4 weeks. It is important to note that all participants benefited from the training, regardless of their clinical profile, CI device brand or setting, or duration of bilateral experience. Furthermore, whatever their performance before training, all participants improved their performance, suggesting that this protocol can be offered to all participants, regardless of their preinclusion performance.

Until now, most spatial hearing protocols have been developed with normally hearing (NH) adults with little homogeneity between protocols: some protocols used real sounds (Strelnikov

TABLE 3. Individual patient data for the speech reception threshold in the matrix test across evaluation sessions (E1 to E4)

Group	Id	E1	E2	E3	E4
A	A01	2.1	1.6	0.9	-0.4
	A02	-1.3	-1.3	-1.1	-2.3
	A03	3.0	2.6	-0.3	0.9
	A04	4.0	1.2	-0.2	1.3
	A06	1.5	0.4	-0.5	0.7
	B	B01	6.8	3.0	1.7
B02		1.6	0.2	1.8	0.2
B03		-3.9	-5.5	-6.8	-4.4
B04		-1.2	-1.1	-2.3	-3.5
B05		-4.7	-3.7	-5.2	-4.2
B06		-3.2	-3.4	-5.1	-4.8

Group A, audiovisual feedback; Group B, visual feedback; Id, patient identification.

et al. 2011), others virtual sounds (Mendonça et al. 2013, Steadman et al. 2019), some performed the training in a laboratory setting while others implemented at-home training (Tyler et al. 2010). Strelnikov et al. (2011) were one of the first groups to propose training sessions spread over several days, and this was done with NH adults wearing a monaural plus. Studies conducted with hearing impaired people are rare, with only three published studies of note. These included either patients with unilateral hearing loss (Firszt et al. 2015), bilateral moderate deafness (Kuk et al. 2014), or a single case study of a bilateral CI user (Tyler et al. 2010). The present study is thus

the first to evaluate the feasibility and efficacy of a standardized rehabilitation protocol spread over 4 weeks for a group of bilateral CI users.

Spatial Hearing Improvement

Compared to our previous pilot study where normal hearing participants performed at ceiling at a similarly localization task, in the present study all CI users had significant spatial hearing difficulties before training. Spatial hearing improvement was most noticeable for front-back confusions, the median percentage of confusions decreased by more than half (from 32% before training to 14.8% 4 weeks after training). While this improvement could be due to a transfer of motor learning from the training to the evaluation sessions we believe this unlikely. First, because the response modality differed between the training and evaluation (head versus hand), and second because there is no published evidence for transfer of motor learning from the head to the upper limb. We suggest instead, that our rehabilitation protocol gave bilateral CI users an opportunity to train their auditory skills and that this training transferred to non-trained locations. The improvement we observed was likely facilitated by the fact that the training was carried out in the near-field (less than 1 m from the participant), where auditory cues are more readily available compared to the far-field. Indeed, the closer the sound is to the listener the larger the low-frequency ILD, and the better the accuracy in azimuth and distance (Brungart & Rabinowitz 1999; Kolarik et al. 2016).

Despite this improvement, front-back discrimination remained difficult for CI users, partly due to the lack of salient

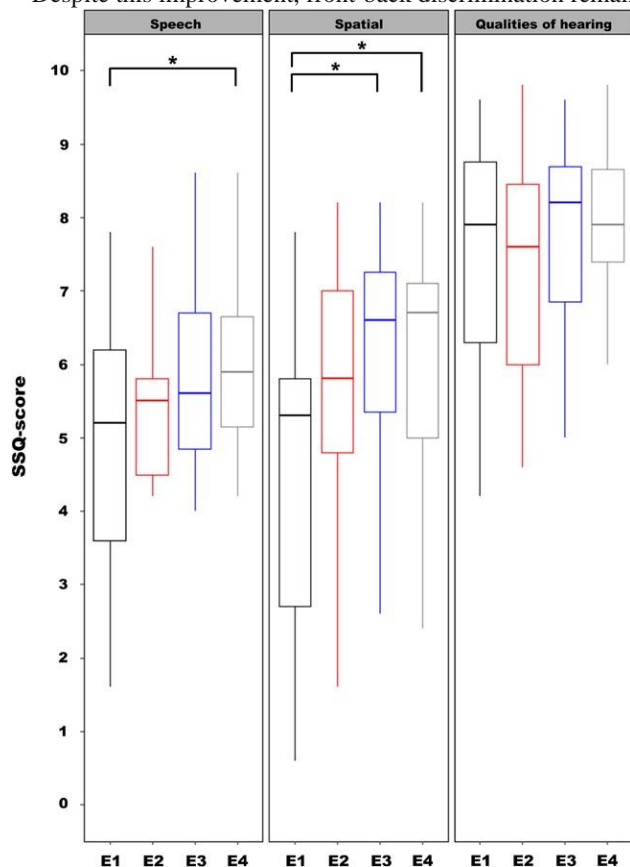


Fig. 7. SSQ-score during evaluation sessions E1 to E4 shown separately for each subscale (SSQ). Lines are medians, box limits 25th–75th percentiles and error bars 95% confidence limits ($n = 12$ for each session). Asterisks indicate significant differences (Wilcoxon test, $*p < 0.05$). SSQ, speech, spatial, and qualities of hearing.

ITD cues (Aronoff et al. 2010; see Laback et al. 2015 for review). Some authors suggest that the fine-structure temporal processing implemented in some processors (e.g., FS4 in MED-EL) improves ITD cues and helps patients with some auditory skills (e.g., music perception, Roy et al. 2015). In accordance with a recent study showing that patients with a fine-structure ITD processing did not perform better on a spatial localization task (Ausili et al 2020), participants in our sample fitted with this technology were not those who showed the most improvement in sound localization performance. Bilateral CI users also experience difficulties extracting ILD cues because the magnitude of ILDs is decreased by the compression applied by the CI processor. In more detrimental situations the automatic gain control can even lead to inverted ILDs (Dorman et al. 2014; Archer-Boyd & Carlyon 2019). All these elements contribute to explaining the presence of large sound localization errors in azimuth in bilateral CI users (e.g., Kerber & Seeber 2012). Importantly, this can be partially compensated for by head movements, which create dynamic binaural cues and naturally increase the sound level differences arriving at each ear. The benefits of head movements for resolving front-back ambiguities have already been noted in previous CI studies (Mueller et al. 2014; Pastore et al. 2018; Fischer et al. 2020; Coudert et al. 2021), and in the present study,

head-movement-induced ILD variations could have been large enough to be usable by the bilateral CI users. This reinforces the importance of focusing

TABLE 4. Individual patient data for the SSQ questionnaire as a function of subscale (A- speech perception; B- spatial hearing, and C- other qualities of hearing) across evaluation sessions (E1 to E4)

Group	Id	Subscale	E1	E2	E3	E4
A	A01	A	4.2	4.2	4	4.2
		B	5.8	5	7.2	5.4
		C	8.8	7	9	9.2
	A02	A	2.4	4.6	5.2	5
		B	0.6	1.4	2.2	2.4
		C	8.8	8	8.6	8
	A03	A	6.2	5.8	6.4	6.8
		B	5.6	6.8	6.4	6.6
		C	7.4	7.4	7.8	7.8
	A04	A	3.6	5.6	5.6	5.2
		B	7	8.2	7.4	7
		C	8.4	8.4	8	7.8
	A05	A	6.2	5.8	6.6	6.2
		B	3	5.8	6.8	6.8
		C	4.2	6	6.4	6.6
A06	A	5.6	4.8	5.6	6.6	
	B	1.8	1.6	2.6	3.8	
	C	9.6	9.8	9.6	9.8	
B	B01	A	1.6	1.8	0.8	1.4
		B	0.6	4.2	4	3.8
		C	5.4	6	5	6
	B02	A	4.8	5.4	5	5.4
		B	5.4	5.8	5.8	6.2
		C	6.6	6	7	7.6
	B03	A	7.8	8.6	8.6	8.6
		B	7.8	8	8.2	7.4
		C	8.4	8.6	9	8.6
	B04	A	3.6	4.2	4.4	5.6
		B	4.8	5	6.2	7.6
		C	4.8	4.6	6.2	6.8
	B05	A	7.8	7.6	8.2	8.2

	B	5.8	7.6	8.2	8.2
	C	8.8	8.6	8.6	8.8
B06	A	6	5.8	7	6.4
	B	5.2	6.2	6.8	6.8
	C	7.4	7.8	8.4	8

SSQ, speech, spatial, and qualities of hearing.

auditory rehabilitation on the dynamic interaction between the two CIs that occurs naturally in everyday behavior during which the head moves freely.

There is little data available on sound localization in elevation in CI users, but one study by Majdak et al. (2011) reported results similar to our pretraining results. That is, near-chance sound localization accuracy in elevation, large inter-individual variability, and localization clustered around the level of the CI. While sound localization in azimuth relies on binaural cues, detecting a sound in elevation is based on monaural spectral indices from upper body filtration (i.e., the pinna and to a lesser extent the head and shoulders) above 3 kHz (e.g., Musicant & Butler 1984; Perrett & Noble 1997). CI processors do not correctly provide these spectral cues, since all incident sound waves are directly caught by the microphone and the range of upper frequencies encoded by the processor is limited. It is therefore not surprising that unlike our results for front-back performance, only two participants improved their up-down discrimination after training, while the other 10 remained at the chance level. The improvement shown by these two participants raises the question of whether the rehabilitation protocol might have trained them to identify and extract auditory information that enabled them to learn new monaural coordinate cues. This would be consistent with the suggestion by Algazi et al. (2001) that sound localization in elevation is possible from low spectral cues when sound sources are positioned laterally. It is also possible that the position of the microphone (i.e. behind or off the ears) could affect up-down discrimination by modifying monaural cues. However, the two best performers had their processors behind the ears, which is the most detrimental situation compared to a microphone placed inside the external auditory canal. Further investigations are clearly needed to better understand the effect of microphone position on spatial hearing performance.

Rehabilitation Benefits Beyond Spatial Hearing

This study is the first demonstration that a spatial hearing rehabilitation protocol focused on training sound localization can improve other hearing qualities. Indeed, during the dichotic matrix test the mean SRT score decreased from 1.5 to -0.5 dB SNR after eight training sessions, meaning that after training most participants were able to repeat 50% of words when the level of noise was higher than that of speech. Given that the intrinsic variability of this test between sessions is 0.4 dB, this result is noteworthy. Some participants even managed to reach the mean score of NH peers, that is, -6 dB SNR (Jansen et al. 2012). All participants reported tiredness after the 50 sentences and one participant was consistently unable to perform the test, revealing the reality of the everyday difficulties and fatigue experienced by CI users when attempting to understand speech in noise.

In dichotic situations (i.e., when speech and noise sources are spatially separated), the listener has to be able to finely analyze spectral information coming from speech and noise in order to correctly segregate the two and focus on speech decoding (Anderson & Kraus 2010). As mentioned above, however, CI devices have poor spectral cue resolution and limited capacity to convey the fine temporal structure of speech that is essential for word perception (e.g., Moore 2008; Won et al. 2012; D'Alessandro et al. 2018). Improvement on the Matrix test after spatial localization training could be due to the fact that participants learned to better exploit spatial hearing cues necessary for segregating speech from noise, which in turn facilitated their speech understanding. Indeed, the large interindividual variability in this ability cannot be explained by technological constraints alone, but is likely to be partially explained by central factors like semantic knowledge linked to the age of deafness, working memory, and non-verbal intelligence (O'Neill et al. 2019, Zaltz et al. 2020).

While the 3D sound localization test and the matrix test were important for objectively assessing performance, they were performed under artificial experimental conditions. For this reason, we decided to evaluate the impact of the training sessions on participants' daily life using a validated quality of life questionnaire; the short form of the SSQ (Moulin et al. 2019). In line with the content of the training protocol, bilateral CI users' scores on questions about sound localization improved (from 5.3 to 6.6). In line with their improvement on the Matrix test, their scores for speech comprehension in noise also improved (from 5.2 to 5.6), but less than for sound localization. It is difficult to assess whether these changes are clinically/behaviorally relevant, as there are currently no norms for CI users on the SSQ questionnaire. This questionnaire has most often been used to evaluate the benefits of cochlear implantation in the first 2 years after surgery (Hasepass et al. 2013; Zhang et al. 2015), and no data exist at longer delays, when performance and device settings are stable. Based on the categorization system developed by Noble and Gatehouse (2006), we found that one third of participants significantly improved their "speech perception score" and two thirds their "spatial hearing score." Moreover, when compared to the results of a large cohort of patients suffering from moderate hearing loss, that is, a maximum loss of 55 dB HL (Moulin et al. 2019), our mean scores per subscale one month after the end of training were close to theirs (speech perception: 5.9 and 6.7, spatial hearing: 6.7 and 6.5, other qualities of hearing: 7.9 and 8.2, respectively, in our study and theirs). This suggests that our

rehabilitation protocol allowed bilateral CI users to reach a similar hearing-related quality of life to that reported by patients with a less disabling hearing deficit.

Persistent Benefits

One month after the end of the rehabilitation training protocol, performance improvement was maintained for all tests, and some patients even had superior performance at 1-month followup than at the end of training. Previous studies of spatial hearing rehabilitation did not include a follow-up evaluation of the benefits in hearing impaired patients (Kuk et al. 2014; Firszt et al. 2015), nor did any studies in NH plugged-participants, as the plug was removed immediately after training (e.g., Strelnikov et al. 2011; Mendonça et al. 2013; Steadman et al. 2019). As such, this study is the first to demonstrate that the benefits of a rehabilitation protocol can continue beyond the training period. This result raises two questions: (1) do participants maintain their performance because they indirectly continue to train using the multisensory stimuli of everyday life? (2) what is the minimum number of training sessions necessary to see persistent benefits over months or even years?

Multisensory Stimulation and Feedback

A secondary objective of this pilot study was to investigate whether the nature of the feedback (i.e., unisensory versus multisensory) influenced training improvement. The multisensory feedback group received visual and auditory information, similar to real-world situations in which localizing a sound in the environment mostly involves these two sensory systems. When the visual and auditory sound sources are spatially congruent the information from these two systems largely overlaps, which allows the brain to develop an optimal spatial map of the environment. This improves localization accuracy (Bulkin & Groh 2006) and can be useful for resolving confusing situations (e.g., when background noise masks the sound source of interest). Visual and auditory information can also be complementary; for example, when the stimulus is outside the visual field or when there is a sensory deficit (e.g., hearing loss).

Recent studies have shown that multisensory training can promote subsequent unisensory learning (Shams et al. 2011; Isaiiah & Hartley 2015), and that adding redundant information from other intact sensory modalities (e.g., vision) does not make the task too easy but instead reduces the effort involved and promotes better learning (Strelnikov et al. 2011; Isaiiah & Hartley 2015). In everyday situations, patients with hearing deficits rely heavily upon the visual system to compensate for the lack of information from the auditory system (Rouger et al. 2007). This compensation leads to a high level of fatigue at the end of the day (Alhanbali et al. 2017, Hughes et al. 2018). Based on these ideas, we predicted that training with multisensory feedback would lead to greater performance improvement than training with visual feedback alone. We found, however, that training-related performance improvement was similar in the two groups. This finding should be interpreted cautiously, however, as the two groups were small and were not matched for age, hearing history, or any other demographic variables. If, however, the results were not due to uncontrolled clinical variables or the small sample size, the similarity in performance raises several hypotheses. First, that the nature of the feedback was less important for performance improvement than the multisensory interactions and information available during the search for the loudspeaker position. Indeed, during this search phase, patients were encouraged to actively move their heads to help them perceive differences in binaural cues and they all received visual and auditory inputs that were temporally and spatially congruent. It is possible that the training-related learning was linked to this process and not to the nature of the feedback. Second, based on data from imaging research, the absence of a difference between the two groups could be due to the predominance of the visual system in hearing impaired patients. Giraud et al (2001) found that an auditory task activated both visual and auditory primary cortex in normal hearing subjects and in CI users, but that visual cortex activation was greater in CI users, even 3 years after surgery. A final explanation could be that the feedback is not a necessary part of the training protocol, and providing patients with an opportunity to practice spatial localization abilities is sufficient to induce learning that transfers beyond spatial hearing performance. We think this is unlikely, however, as the position and type of the stimuli, as well as the response modality differed between the training and evaluation sessions. Furthermore, since all patients had at least 18 months of bilateral experience it is unlikely that the feedback was not important and that performance improvement can be explained simply by the additional listening experience provided by eight 45-minute training sessions.

Clinical Implications

Speech understanding is at the center of hearing rehabilitation after cochlear implantation. Given its importance in everyday life this makes sense. Spatial hearing is also important but is often neglected, even several years after surgery when patients are comfortable understanding speech but spatial localization remains difficult. The promising results from the training protocol used in this pilot study suggest that spatial hearing training could be systematically proposed to a range of patients regardless of implantation age or duration of bilateral experience. Intensive training is feasible in clinical practice, and relies largely on patients actively seeking care to improve their hearing quality and being motivated to attend rehabilitation sessions. We did not test any patients less than 1 year after surgery, but the success of this initial pilot study suggests that it would be interesting to investigate the possible benefits of adding spatial hearing training to the speech understanding rehabilitation that begins just after surgery. Since hearing impaired patients routinely face challenging situations when attempting to understand speech (e.g., interfering background noise, competing speakers, and reverberant environments), a next step in developing our training protocol could be the addition of more complex stimuli that simulate real-life situations (e.g., adding background noise and varying its sound level). This type of rehabilitation protocol could also be proposed to patients wearing hearing aids, who also experience spatial hearing difficulties.

CONCLUSION

This pilot study demonstrated the feasibility and benefits of a new approach to spatial hearing rehabilitation in CI users. Our 4-week training protocol led to substantial improvement in resolving front-back confusions, in understanding speech in noise, and in hearing-related quality of life. All patients adhered to the training sessions over the 4 weeks and attended all five evaluation sessions across the 10-week study duration. The ease of use of the virtual reality system regardless of the participants' age, as well as the fun and engaging aspects of the technology make it a tool of choice for wider clinical use. Future studies including control groups are needed to determine whether the feedback is an essential aspect of the protocol, and if so, the nature of the feedback that leads to the greatest improvement in performance.

ACKNOWLEDGMENTS

The authors are grateful to all participants who took part in this study. We thank the IMPACT team administrative staff for their administrative and technical support.

A.C. was supported by a grant from the Hospices Civils de Lyon; G.V., F.P., E.T., and V.G. were supported by a grant from the Agence Nationale de la Recherche (ANR-16-CE17-0016, VIRTUALHEARING3D, France).

All the authors contributed significantly to this work. A.C., V.G., and F.P. designed the experiment; A.C., V.G. performed experiments and analyzed the data; A.C., K.R., and V.G. wrote the paper. E.T. provided cochlear implant participants. G.V. implemented the experiments. All authors discussed the results, reviewed the manuscript and approved the final version of the revised manuscript.

The authors have no conflicts of interest to disclose.

Address for correspondence: Aurélie Coudert, Department of Pediatric Otolaryngology – Head & Neck Surgery, Femme Mere Enfant Hospital - Hospices Civils de Lyon, 59 Boulevard Pinel, 69677 Bron, France. E-mail: aurelie.coudert2@chu-lyon.fr

Received August 10, 2021; accepted May 13, 2022

REFERENCES

- Algazi, V. R., Avendano, C., Duda, R. O. (2001). Elevation localization and head-related transfer function analysis at low frequencies. *J Acoust Soc Am*, *109*, 1110–1122.
- Alhanbali, S., Dawes, P., Lloyd, S., Munro, K. J. (2017). Self-reported listening-related effort and fatigue in hearing-impaired adults. *Ear Hear*, *38*, e39–e48.
- Anderson, S., & Kraus, N. (2010). Sensory-cognitive interaction in the neural encoding of speech in noise: A review. *J Am Acad Audiol*, *21*, 575–585.
- Angell JR, Fite W. (1901a). The monaural localization of sound. *Psychol Rev*, *8*, 225–246.
- Angell JR, Fite W. (1901b). Further observations on the monaural localization of sound. *Psychol Rev*, *8*, 449–458.
- Archer-Boyd, A. W., & Carlyon, R. P. (2019). Simulations of the effect of unlinked cochlear-implant automatic gain control and head movement on interaural level differences. *J Acoust Soc Am*, *145*, 1389.
- Aronoff, J. M., Yoon, Y. S., Freed, D. J., Vermiglio, A. J., Pal, I., Soli, S. D. (2010). The use of interaural time and level difference cues by bilateral cochlear implant users. *J Acoust Soc Am*, *127*, EL87–EL92.
- Ausili, S. A., Agterberg, M. J. H., Engel, A., Voelter, C., Thomas, J. P., Brill, S., Snik, A. F. M., Dazert, S., Van Opstal, A. J., Mylanus, E. A. M. (2020). Spatial hearing by bilateral cochlear implant users with temporal finestructure processing. *Front Neurol*, *11*, 915.
- Bolognini, N., Frassinetti, F., Serino, A., Làdavas, E. (2005). “Acoustical vision” of below threshold stimuli: Interaction among spatially converging audiovisual inputs. *Exp Brain Res*, *160*, 273–282.
- Brimijoin, W. O., McShefferty, D., Akeroyd, M. A. (2010). Auditory and visual orienting responses in listeners with and without hearing-impairment. *J Acoust Soc Am*, *127*, 3678–3688.
- Brimijoin, W. O., McShefferty, D., Akeroyd, M. A. (2012). Undirected head movements of listeners with asymmetrical hearing impairment during a speech-in-noise task. *Hear Res*, *283*, 162–168.
- Brungart, D. S. (1999). Auditory localization of nearby sources. III. Stimulus effects. *J Acoust Soc Am*, *106*, 3589–3602.
- Brungart, D. S., & Rabinowitz, W. M. (1999). Auditory localization of nearby sources. Head-related transfer functions. *J Acoust Soc Am*, *106*(3Pt 1), 1465–1479.
- Bulkin, D. A., & Groh, J. M. (2006). Seeing sounds: Visual and auditory interactions in the brain. *Curr Opin Neurobiol*, *16*, 415–419.
- Coudert, A., Gaveau, V., Gatel, J., Verdelet, G., Salemme, R., Farne, A., Pavani, F., Truy, E. (2022). Spatial hearing difficulties in reaching space in bilateral cochlear implant children improve with head movements. *Ear Hear*, *43*, 192–205.
- Da Silva, J. A. (1985). Scales for perceived egocentric distance in a large open field: Comparison of three psychophysical methods. *Am J Psychol*, *98*, 119–144.
- Dincer D'Alessandro, H., Ballantyne, D., Boyle, P. J., De Seta, E., DeVincentiis, M., Mancini, P. (2018). Temporal fine structure processing, pitch, and speech perception in adult cochlear implant recipients. *Ear Hear*, *39*, 679–686.
- Dorman, M. F., Loisselle, L., Stohl, J., Yost, W. A., Spahr, A., Brown, C., Cook, S. (2014). Interaural level differences and sound source localization for bilateral cochlear implant patients. *Ear Hear*, *35*, 633–640.
- Firszt, J. B., Reeder, R. M., Dwyer, N. Y., Burton, H., Holden, L. K. (2015). Localization training results in individuals with unilateral severe to profound hearing loss. *Hear Res*, *319*, 48–55.
- Fischer, T., Schmid, C., Kompis, M., Mantokoudis, G., Caversaccio, M., & Wimmer, W. (2020). Pinna-imitating microphone directionality improves sound localization and discrimination in bilateral cochlear implant users. *Ear and Hearing*, *42*, 214–222.
- Gatehouse, S., & Noble, W. (2004). The speech, spatial and qualities of hearing scale (SSQ). *Int J Audiol*, *43*, 85–99.
- Giraud, A. L., Price, C. J., Graham, J. M., Truy, E., Frackowiak, R. S. (2001). Cross-modal plasticity underpins language recovery after cochlear implantation. *Neuron*, *30*, 657–663.
- 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 Hear*, *28*, 524–541.
- Hassepass, F., Schild, C., Aschendorff, A., Laszig, R., Maier, W., Beck, R., Wesarg, T., Arndt, S. (2013). Clinical outcome after cochlear implantation in patients with unilateral hearing loss due to labyrinthitis ossificans. *Otol Neurotol*, *1278*–1283.

- Hughes, S. E., Hutchings, H. A., Rapport, F. L., McMahon, C. M., Boisvert, I. (2018). Social connectedness and perceived listening effort in adult cochlear implant users: A grounded theory to establish content validity for a new patient-reported outcome measure. *Ear Hear*, *39*, 922–934.
- Isaiah, A., & Hartley, D. E. (2015). Can training extend current guidelines for cochlear implant candidacy? *Neural Regen Res*, *10*, 718–720.
- Jansen, S., Luts, H., Wagener, K. C., Kollmeier, B., Del Rio, M., Dauman, R., James, C., Fraysse, B., Vormès, E., Frachet, B., Wouters, J., van Wieringen, A. (2012). Comparison of three types of French speech-in-noise tests: A multi-center study. *Int J Audiol*, *51*, 164–173.
- Kerber, S., & Seeber, B. U. (2012). Sound localization in noise by normalhearing listeners and cochlear implant users. *Ear Hear*, *33*, 445–457.
- Kolarik, A. J., Moore, B. C., Zahorik, P., Cirstea, S., Pardhan, S. (2016). Auditory distance perception in humans: A review of cues, development, neuronal bases, and effects of sensory loss. *Atten Percept Psychophys*, *78*, 373–395.
- Kuk, F., Keenan, D. M., Lau, C., Crose, B., Schumacher, J. (2014). Evaluation of a localization training program for hearing impaired listeners. *Ear Hear*, *35*, 652–666.
- Laback, B., Egger, K., Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hear Res*, *322*, 138–150.
- Lafon, J. C. (1964). *Le test phonétique et la mesure de l'audition*, éditions centrex –eindhoven, 144–146.
- Loomis, J. M., Klatzky, R. L., Philbeck, J. W., Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Percept Psychophys*, *60*, 966–980.
- MacLeod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *Br J Audiol*, *21*, 131–141.
- Majdak, P., Goupell, M. J., Laback, B. (2011). Two-dimensional localization of virtual sound sources in cochlear-implant listeners. *Ear Hear*, *32*, 198–208.
- Mendonça, C., Campos, G., Dias, P., & Santos, J. A. (2013). Learning auditory space: Generalization and long-term effects. *PLOS ONE*, *8*, e77900.
- Middelweerd, M. J., & Plomp, R. (1987). The effect of speechreading on the speech-reception threshold of sentences in noise. *J Acoust Soc Am*, *82*, 2145–2147.
- Middlebrooks, J. C. (2015). Sound localization. *Handb Clin Neurol*, *129*, 99–116.
- Mo, B., Lindbaek, M., Harris, S. (2005). Cochlear implants and quality of life: A prospective study. *Ear Hear*, *26*, 186–194.
- Moore, B. C. J. (2008). The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people. *J Assoc Res Otolaryngol: JARO*, *9*, 399–406.
- Moulin, A., Vergne, J., Gallego, S., & Michey, C. (2019). A new speech, spatial, and qualities of hearing scale short-form : Factor, cluster, and comparative analyses. *Ear and Hearing*, *40*, 938–950.
- Mueller, M. F., Meisenbacher, K., Lai, W. K., Dillier, N. (2014). Sound localization with bilateral cochlear implants in noise: How much do head movements contribute to localization? *Cochlear Implants Int*, *15*, 36–42.
- Musicant, A. D., & Butler, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. *J Acoust Soc Am*, *75*, 1195–1200.
- Nawaz, S., McNeill, C., Greenberg, S. L. (2014). Improving sound localization after cochlear implantation and auditory training for the management of single-sided deafness. *Otol Neurotol*, *35*, 271–276.
- Noble, W., & Gatehouse, S. (2006). Effects of bilateral versus unilateral hearing aid fitting on abilities measured by the speech, spatial, and qualities of hearing scale (SSQ). *Int J Audiol*, *45*, 172–181.
- O'Neill, E. R., Kreft, H. A., Oxenham, A. J. (2019). Cognitive factors contribute to speech perception in cochlear-implant users and age-matched normal-hearing listeners under vocoded conditions. *J Acoust Soc Am*, *146*, 195.
- Pastore, M. T., Natale, S. J., Yost, W. A., Dorman, M. F. (2018). Head movements allow listeners bilaterally implanted with cochlear implants to resolve front-back confusions. *Ear Hear*, *39*, 1224–1231.
- Pavani, F., Husain, M., Driver, J. (2008). Eye-movements intervening between two successive sounds disrupt comparisons of auditory location. *Exp Brain Res*, *189*, 435–449.
- Perreau, A. E., Ou, H., Tyler, R., Dunn, C. (2014). Self-reported spatial hearing abilities across different cochlear implant profiles. *Am J Audiol*, *23*, 374–384.
- Perrett, S., & Noble, W. (1997). The effect of head rotations on vertical plane sound localization. *J Acoust Soc Am*, *102*, 2325–2332.
- Rouger, J., Lagleyre, S., Fraysse, B., Deneve, S., Deguine, O., Barone, P. (2007). Evidence that cochlear-implanted deaf patients are better multisensory integrators. *Proc Natl Acad Sci U S A*, *104*, 7295–7300.
- Roy, A. T., Carver, C., Jiradejvong, P., Limb, C. J. (2015). Musical sound quality in cochlear implant users: A comparison in bass frequency perception between fine structure processing and high-definition continuous interleaved sampling strategies. *Ear Hear*, *36*, 582–590.
- Schwartz, J. L., Berthommier, F., Savariaux, C. (2004). Seeing to hear better: Evidence for early audio-visual interactions in speech identification. *Cognition*, *93*, B69–B78.
- Seeber, B. U., Baumann, U., Fastl, H. (2004). Localization ability with bimodal hearing aids and bilateral cochlear implants. *J Acoust Soc Am*, *116*, 1698–1709.
- Shams, L., Wozny, D. R., Kim, R., Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Front Psychol*, *2*, 264.
- Shinn-Cunningham, B., Best, V., Lee, A. K. (2017). Auditory object formation and selection. In *The Auditory System at the Cocktail Party*. Springer, J.C. Middlebrooks, J.Z. Simon, A.N. Popper, R.R. Fay. 7–40.
- Smulders, Y. E., van Zon, A., Stegeman, I., Rinia, A. B., Van Zanten, G. A., Stokroos, R. J., Hendrice, N., Free, R. H., Maat, B., Frijns, J. H., Briaire, J. J., Mylanus, E. A., Huinck, W. J., Smit, A. L., Topsakal, V., Tange, R. A., Grolman, W. (2016). Comparison of bilateral and unilateral cochlear implantation in adults: A randomized clinical trial. *JAMA Otolaryngol Head Neck Surg*, *142*, 249–256.
- Steadman, M. A., Kim, C., Lestang, J. H., Goodman, D. F. M., Picinali, L. (2019). Short-term effects of sound localization training in virtual reality. *Sci Rep*, *9*, 18284.
- Strelnikov, K., Rosito, M., Barone, P. (2011). Effect of audiovisual training on monaural spatial hearing in horizontal plane. *PLoS One*, *6*, e18344.
- Tyler, R. S., Perreau, A. E., Ji, H. (2009). Validation of the spatial hearing questionnaire. *Ear Hear*, *30*, 466–474.
- Tyler, R. S., Witt, S. A., Dunn, C. C., Wang, W. (2010). Initial development of a spatially separated speech-in-noise and localization training program. *J Am Acad Audiol*, *21*, 390–403.
- Valzolgher, C., Alzhaler, M., Gessa, E., Todeschini, M., Nieto, P., Verdelet, G., Salemm, R., Gaveau, V., Marx, M., Truy, E., Barone, P., Farnè, A., Pavani, F. (2020a). The impact of a visual spatial frame on real soundsource localization in virtual reality. *Current Research in Behavioral Sciences*, *1*, 100003.
- Valzolgher, C., Verdelet, G., Salemm, R., Lombardi, L., Gaveau, V., Farnè, A., Pavani, F. (2020b). Reaching to sounds in virtual reality: A multisensory-motor approach to promote adaptation to altered auditory cues. *Neuropsychologia*, *149*, 107665.
- van Hoesel, R. J. M. (2004). Exploring the benefits of bilateral cochlear implants. *Audiol. Neurootol*, *9*, 234–246.
- van Hoesel, R. J., & Tyler, R. S. (2003). Speech perception, localization, and lateralization with bilateral cochlear implants. *J Acoust Soc Am*, *113*, 1617–1630.
- van Zon, A., Smulders, Y. E., Stegeman, I., Ramakers, G. G., Kraaijenga, V. J., Koenraads, S. P., Zanten, G. A., Rinia, A. B., Stokroos, R. J., Free, R. H., Frijns, J. H., Huinck, W. J., Mylanus, E. A., Tange, R. A., Smit, A. L., Thomeer, H. G., Topsakal, V., Grolman, W. (2017). Stable benefits of bilateral over unilateral cochlear implantation after two years: A randomized controlled trial. *Laryngoscope*, *127*, 1161–1168.
- Verdelet, G., Desoche, C., Volland, F., Farnè, A., Coudert, A., Hermann, R., Truy, E., Gaveau, V., Pavani, F., Salemm, R. (2019). *Assessing Spatial and Temporal Reliability of the Vive System as a Tool for Naturalistic Behavioural Research*. 2019 International Conference on 3D Immersion (IC3D), 1–8. <https://doi.org/10.1109/IC3D48390.2019.8975994>
- Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *J. Exp. Psychol*, *27*, 339–368.

- Wightman, F. L., & Kistler, D. J. (1999). Resolution of front-back ambiguity in spatial hearing by listener and source movement. *J Acoust Soc Am*, *105*, 2841–2853.
- Witten, I. B., & Knudsen, E. I. (2005). Why seeing is believing: Merging auditory and visual worlds. *Neuron*, *48*, 489–496.
- Won, J. H., Lorenzi, C., Nie, K., Li, X., Jameyson, E. M., Drennan, W. R., & Rubinstein, J. T. (2012). The ability of cochlear implant users to use temporal envelope cues recovered from speech frequency modulation. *The Journal of the Acoustical Society of America*, *132*, 1113–1119.
- Zaltz, Y., Buganim, Y., Zechoval, D., Kishon-Rabin, L., Perez, R. (2020). Listening in noise remains a significant challenge for cochlear implant users: Evidence from early deafened and those with progressive hearing loss compared to peers with normal hearing. *J Clin Med*, *9*, E1381.
- Zhang, J., Tyler, R., Ji, H., Dunn, C., Wang, N., Hansen, M., Gantz, B. (2015). Speech, spatial and qualities of hearing scale (ssq) and spatial hearing questionnaire (shq) changes over time in adults with simultaneous cochlear implants. *Am. J. Audiol*, *24*, 384–397.